May 8, 1974

To the Graduate Council:

I am submitting herewith a dissertation written by Robert Allen Gregory entitled "Aptitude-Treatment Interaction and Retention: A Study of the Differential Effects of Two Teaching Methods." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Education with a major in Educational Psychology and Guidance.

William H. Cormier

William H. Cormier, Major Professor

We have read this dissertation
and recommend its acceptance:

Accepted for the Council:

Hilton A. Smith

Vice Chancellor
Graduate Studies and Research

APTITUDE-TREATMENT INTERACTION AND RETENTION: A STUDY OF THE DIFFERENTIAL EFFECTS OF TWO TEACHING METHODS

A Dissertation

Presented for the

Doctor of Education

Degree

The University of Tennessee

Robert Allen Gregory

June 1974

ACKNOWLEDGEMENTS

Great appreciation is expressed to the Doctoral Committee in guiding this research effort: Dr. William H. Cormier, Chairman, and members: Dr. L. M. DeRidder, Dr. S. W. Huck, and Dr. O. Milton.

Sincere thanks go to Elizabeth L. Mattice for her assistance and encouragement in this endeavor.

Acknowledgement is hereby made to The University of Tennessee Learning Research Center (Dr. Ohmer Milton, Head), for a mini-mini grant which funded part of this study.

1146696

ABSTRACT

This study investigated the possibility and nature of the occurrence of aptitude-treatment interactions at the college level and the degree of retention of material due to two different methods of instruction. The aptitude investigated was represented by prior achievement as measured by grade point average (GPA), and the treatment was two types of instruction: traditional lecture/discussion and self-pacing. Dependent measures were examination scores.

It was hypothesized that lower achieving subjects (Ss) would achieve better under the highly structured less threatening self-paced condition than in the intense discussion wherein implication, applications, and logical extentions of material covered were stressed with personal involvement in the discussions required. Higher achieving Ss should do better under the latter condition. It was further hypothesized that, disregarding achievement level, students would retain more of what they learned under the self-paced condition.

Fifty-four females at The University of Tennessee, predominantly in their second and third year of college, who were enrolled in a basic educational psychology course required of all elementary education majors were divided into two groups based upon GPA. They were then randomly assigned either to class or self-pacing in a counter-balanced design wherein they would receive one treatment for one-half the quarter and the other treatment the second half. All Ss were given study guides containing pre-tests, objectives, activities, and post-tests for ten

weeks (one unit per week). Students in the traditional class condition were given check-ups (quizzes) each week while those on self-pacing were encouraged to take these check-ups at their convenience but attended no formal class sessions.

The first hypothesis of differential achievement was not substantiated. A Lindquist Type I analysis of variance showed the class condition to be superior to self-pacing for both achievement levels ($p < .05$). To check for retention a follow-up examination on material covered by the mid-term examination was done and also the final (which was independent--content-wise--of the mid-term). A Lindquist I analysis of covariance (GPAs for the two groups differed) was done on each of these. There was a significant interaction ($p < .05$) for the mid-term examination and its follow-up for the two treatment conditions. Though the class condition had scored higher at the mid-term, the self-paced condition scored higher three months later. Analysis for simple main effects showed both treatment groups to drop significantly from the mid-term to the follow-up, but those in the class condition retained less. Results from the final and follow-up were not statistically significant but were of a similar pattern (the Ss who returned for the follow-up had scored higher originally in the self-paced condition--unlike the mid-term--and did not go down as much).

A further analysis was done to investigate the type of material retained best, by whom, and learned under what condition. Two Lindquist Type III analyses of covariance (follow-up scores from mid-term and final), with original scores being the covariates, were done on the two types of material tested upon--applied, problem-situation and association

items or recall information. No statistically significant interactions were found from these analyses. Across treatments and achievement levels, however, Ss scored significantly better on applied material than on association material in both analyses.

Overall, the findings suggest that initial support for traditional class might not hold up when retention is examined three months later. What is learned under self-paced conditions is retained better. Under all conditions, applied material seems to be learned better than recall (or association) material.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	1
	Attitude-Treatment Interactions	2
	Personological Variables	4
	Grade Point Average	5
	Retention	9
	Instructional Method	10
	Environmental Press and Academic Experience Questionnaire	11
	Purpose of the Study	13
II.	METHOD	15
	Subjects	15
	Procedures	15
	Dependent Measures	18
	Attitude Measures	20
	Instructor Behavior	21
	Experimental Design	22
III.	RESULTS	24
	Mid-Term and Final Examinations	24
	Retention Examination	30
	Differential Retention of Material	35
	Attitude Measures	36
IV.	DISCUSSION	44

CHAPTER PAGE

Aptitude-Treatment Interactions 44

Retention 46

Press of Class 50

V. SUMMARY 52

Recommendations for Further Research 54

LIST OF REFERENCES 56

APPENDIX 61

VITA 63

LIST OF TABLES

TABLE PAGE

1. Summary of Results of the Analyses of Variance for All Ss, 40 Extreme Ss, and 20 Extreme Ss 26
2. Cell Means for the Ss Used in the Analyses of Variance for the Major Treatment Effects 28
3. Summary of the Results of the Analysis of Covariance for Sequence as the Dependent Variable with GPA Used as the Covariate 29
4. Analyses of Covariance of Class Versus Self-paced on Mid-term and Final and Respective Retention Examinations . . . 33
5. Analyses of Covariance of High GPA and Low GPA Subjects in Traditional Class Versus Self-paced on Mid-term and Final Examinations Broken Down into Applied- Versus Association-Types of Test Items 37
6. Adjusted Cell Means and Within Group Regression Slopes; Mid-term and Final Portions of the Follow-up Examination . 38
7. Composite Results of the Press of Class Questionnaire . . . 40
8. Composite Results of Academic Experiential Questionnaire Time Spent in Class 41
9. Composite Results of Academic Experiential Questionnaire Time Spent Out of Class 43

CHAPTER I

INTRODUCTION

In a summary of research on teaching as recently as 1973, Rosenshine and Furst have declared:

> Despite the doctrine of individual differences in education, studies still have been designed as if one set of instructional behaviors were effective for all students. Almost all the experimental and correlational studies have focused upon the relationship of teacher behavior to the class mean. Few investigators have focused on the personality or learning style of subgroups of learners, or have stratified classes according to the initial knowledge or aptitude of the students (p. 172).

The present study is designed to examine individual differences of students using a broad aptitude and two treatment conditions. The possibility of differential treatment effects on groups of learners is the major concern. Also, this study seeks to determine if levels of aptitude (grade point average) interact with different levels of treatment given the learners. In addition, both immediate learning or achievement and retention of what is learned are examined.

In this chapter the literature on aptitude-treatment interaction (ATI) is reviewed as it relates to the present study. Subsequently, literature concerning personological variables (aptitudes) is examined. Following this general review, more specific studies relating to the aptitude chosen for study along with variables thought to interact with grade point average (GPA) are reviewed. The limited amount of material on retention in relation to the present study is examined, followed by a brief review of research on instructional conditions (treatments)

which are somewhat similar to this present study. A description of the environmental press as this concept relates to this study is presented at the conclusion of this chapter.

Attitude-Treatment Interactions

Cronbach (1967) made a plea for more research aimed toward establishing which kinds of instructional methods would result in the biggest payoff for different kinds of students. The search for an optimal teaching method is futile he contends; we must search for " . . . a theory whose propositions would state the best conditions for pupils of certain types, both conditions and types being described in terms of fairly broad dimensions" (p. 90). This plea has been echoed by several others (Siegel, 1968; Siegel and Siegel, 1966; Jensen, 1967; Koran, 1972; Levin, Rohwer, and Cleary, 1971; Reynolds and Balow, 1972; Nunney and Hill, 1972; Newson, Eischens, and Looft, 1972; Atkinson and Paulson, 1972; Haskell, 1970, 1971; McNally, 1973; DiVesta, 1973; Snow, 1970; Mitchell, 1969; Randhawa and Fu, 1973; and Milton, 1972). Much of this research has been related to individualization of instruction; implicit in any effort to individualize instruction is the recognition that one learning situation or one treatment (broadly defined) is not optimal for everyone.

Snow (1970) and Koran (1972) state the purpose of aptitude-treatment interaction research as finding the manner in which learner characteristics affect the extent to which the learner attains desired outcomes from each treatment being considered; i.e., what treatment is best for each learner. The definition of ATI (Snow, 1970) comes from

the method of analysis of variance: "An interaction is present when an effect found for one kind of subject or in one kind of setting is not found under other conditions" (p. 64). In describing population validity, Bracht and Glass (1968) accent the importance of the interaction of personological variables and treatment effects. They say: "If the superiority of one experimental treatment over another would be reversed when subjects at a different level of some variable descriptive of persons are exposed to the treatments, there exists an interaction of treatment effects and personological variable" (p. 438). Bracht and Glass go on to discuss both population and ecological validity and convey the impression that the degree of ecological validity of an experiment is, to a large degree, proportional to the degree to which ATIs are discovered.

The emphasis of ATI research is generally related to disordinal interactions and not to ordinal ones (Bracht, 1970). An ordinal interaction is generally seen as having lines which have different slopes but do not intersect within the aptitude range being considered (Bracht and Glass, 1968). One treatment might well be superior across all aptitude levels. In a disordinal interaction, however, the lines of the treatments cross somewhere within the range of the aptitude variable under consideration. Bracht and Glass indicate that lines crossing, however, is not sufficient evidence for concluding that the interaction is disordinal, and they stress the need for statistical tests to determine the nature of a significant interaction when an interaction is found.

Mitchell (1969) describes the state of the ATI research as being in its infancy, and the value of such research can be determined only

after much more effort has been expended. Milton (1972) makes the pertinent observation that "So many factors are inextricably bound together that to discuss one in isolation from others is to grossly oversimplify and invite invalidity" (p. 6).

Personological Variables

"Aptitude" is generally defined quite broadly as a personological variable (e.g., I.Q., scientific interest, anxiety, etc.) upon which people can be reliably and differentially measured. Walberg (1970) defines "aptitude" as " . . . characteristic of the individual or his experiences that predict a given learning criteria" (p. 185). He feels that learning is a function not only of instruction, aptitude, and environment but also of the various interactions of these three dimensions. Old research questions have dealt with only the instruction and the aptitude; the third variable (environment) and the four possible interactions (three first order and one second order) have not been considered or dealt with adequately. Usually either the personological variables, the environment, or the treatments have been considered alone; less frequently an interaction of two of these dimensions has been investigated.

Milton's (1972) above statement accentuates the complexity of the type of research undertaken by this study. For example, Snow (1970) emphasizes the lack of organization or unification of study in the area of personological variables. Snow is "staggered" by the huge number of possible dimensions but "stopped by the recognition that few . . . dimensions have been sufficiently validated . . ." (p. 68). Snow cites

as an example the factor-analytic study of intellectual abilities performed by Guilford (1956). The fine distinctions made by Guilford are not justifiable to Snow. General ability, Snow feels, regardless of components, is clearly and strongly related to learning in many instructional situations. Snow suggests a hierarchical conception of ability organization (as opposed to Guilford's three-dimensional model) for the purpose of using ability variables in research or practice regardless of the "real" organization of human abilities. Ability constructs, in other words, should be based upon utility rather than some all-encompassing theoretical basis.

Grade Point Average

Cofer (1967) and Cronbach (1967) strongly emphasize the need to look at broader dimensions which will be of utility. The practicality of using a variable that discriminates is of utmost importance. Schultz (1966) urges the criterion of utility also: "To attain practical significance one must first insure that the dependent variable being studied is of practical interest, that it is a condition one would like to predict and control" (p. 151). Several studies have attempted to define a "type" of student or personality which is best for a particular method of instruction.

There is considerable evidence to indicate that the broad dimension of grade point average (GPA) is more accurate in prediction of success in school than is any other single measure (Stroup, 1970; Stroup and Eft, 1969; Barton, Dielman, and Cattell, 1971; Miller and O'Connor, 1969; and Entwistle, 1972). It is further suggested by

Entwistle (1972) that GPA incorporates the personality variables associated with traditional school learning; therefore, much of the variance in academic achievement can sensibly and easily be accounted for since these personality variables are incorporated in this relevant dimension. A logical extrapolation of this would be to divide students according to GPA in order to examine the efficacy of this as a practical blocking variable to check for ATIs. Obviously, this would only work where GPA was an indicator of traditional school achievement. Any major and continued interference with its meaning--such as changing to individualized instruction based upon prior GPA--would change its meaning, and the variables previously leading to its discriminatory ability would have to be identified. GPA, in other words, should reflect many personological or aptitude variables and provide a broad dimension along which to separate people for different treatments or instruction. Much of the variance associated with achievement should be accounted for and incorporated into GPA. Domino (1968) forcefully presents the view that there is a quite distinct and understandable interaction between achievement and demands of the environment. Other variables that might be incorporated into GPA are instructional method and factors in the classroom environment. A model for research combining these variables was suggested and tested by Siegel and Siegel (1966).

The work of Siegel and Siegel (1966) which incorporates these three dimensions investigated the effects of instructional processes on environmental and personological variables. Their work was done exclusively in university settings. After a complex analysis and synthesis of a variety of variables, they concluded that the results could be

described in two broad "organizing concepts"--instructional press and idiosyncratic drive patterns.

The concept of instructional press is derived from the way in which the class is conducted. If student performance is highly dependent upon the instructor's behavior and the environmental surroundings then it is referred to as extrinsic instructional press. In contrast, intrinsic instructional press refers to situations wherein student performance seems to reflect more the characteristics of the learner than of the instructor and the environmental conditions. Siegel and Siegel (1966) use the phrase "mixed instructional press" to refer to the many graduations between the extremes.

The other variable of importance to Siegel and Siegel (1966) is the idiosyncratic drive patterns. The drive patterns discussed (though not proported to be the only two) are quite broad and seem to represent both the student's academic ability and his educational set. The drive pattern is supposed to represent "the current end products of a student's past history of successes and failures for cognitive attainment. Thus they are the givens with which he enters every course" (p. 151). Based upon the above research, it would appear that GPA reflects and is a practical manifestation of much of the variance due to the two types of drive patterns and instructional press.

Siegel and Siegel (1966) hypothesize based upon their data that:

> Performance by low ability students may be facilitated when the instructional environment insulates them from intellectual threat and/or provides them with academic guidance. The converse of these conditions tends to inhibit performance by low ability students. High ability students tend to perform best when the instructional environment offers appropriate intellectual challenges, whereas their performance may be inhibited when the environment presents inappropriately easy or unchallenging material (p. 153).

Domino (1968, 1970) has been quite interested in the interactive effects of certain aptitudes of students and the differential effects they have on achievement. In his 1968 publication, he reports differences in achievement reflecting conforming and independent behavior. Domino states there is great heuristic value in separating students based upon GPA into categories reflective of the diverse demands placed upon them. Domino did stress the differential effects of other variables, however. Intellectual ability and environmental demands were two major factors which might differentially effect such results as he obtained.

The present study is quite similar in certain respects to one done by Webb and Baird (1968). Webb and Baird discuss a continuous progress plan of instruction used at Brigham Young University in which students in an individualized study class did better on final examinations than did those taught by a more traditional lecture-discussion method. These results were unusual since most studies have shown no overall differences when different methods are employed. In order to look at student characteristics which might interact with method, they took the students with the ten highest and ten lowest GPAs and compared their post-test scores under the two methods of instruction. The low GPA group who had continuous progress instruction performed better than their counterparts who were in the traditional class. The higher GPA students performed only slightly better on the post-test than did students of similar ability taught by the more traditional method. This study strongly suggests the differential effects of a method of instruction on students with varying characteristics.

Retention

Very little work has been done on retention in the past decade. Gruber and Weitman (1962) claim as a result of several different studies, most of them similar in purpose, that retention is better when independent study is used. Wessell (1967) views a major objective of a college education to be " . . . development of lifetime habits of independent study and continuing intellectual curiosity" (p. 208). Wessell questions the assessment of the development of these qualities on a short-term basis.

Travers (1973) reports on a study done by Baskin in which there was little difference between experimental students and control students in achievement on examinations or in retention of knowledge two years later where independent study was used as the experimental condition and traditional instruction was the control condition.

A rather recent study by Corley, Valente, and Shamow (1971) showed greater achievement resulting from a personalized system on instruction with considerably greater retention ten months later. They used a group of introductory psychology students and tested part of them ten months after course completion. The students on a personalized system scored better initially and retained more of the material.

In a recent study Moore, Hauck, and Gagné (1973) report that in a college physics course taught by continuous progress and traditional instruction, the continuous progress not only showed greater acquisition and transfer of knowledge but also greater retention as measured one year later. This study by Moore et al. revealed that a lower correlation existed between acquisition and retention of material for the continuous

progress group than for the control group. This finding indicates that retention may be a result of factors other than the amount initially learned. Moore et al. suggest that increased retention may be a result of the students using a mixture of learning strategies (related to original learning) some of which favor long-term retention. In addition, these authors suggest that some of the continuous progress students might have formed greater conceptual organization while others in the same group relied on memorization.

Instructional Method

Research on various teaching methods has probably been going on since before the methods were named. Milton (1972) discusses the equivocal nature of the results of this kind of research. As of yet, no one best teaching method has been found. A good deal of research has been done on independent study in higher education in recent years (Travers, 1973), and much of this research has been on the type of instruction popularized by Keller (1968) as the "Keller Plan" or "Personalized System of Instruction" (PSI).

Though the present study has not been strictly patterned after the Keller Plan, it is done in the same spirit as the PSI research. An attempt is being made to see what type of instruction is more appropriate for what type of student so that some personalization can occur. In one study done (Moore, Mahan, and Ritts, 1969) several different classes were taught on a continuous progress program patterned after Keller. Achievement was higher, and attitudes toward instructional procedures were better than with traditional instruction. The researchers concluded

that their results support a continuous progress program independent of academic discipline.

No research was found that investigated the possible interactions of different levels of aptitudes in the context of PSI research. Travers (1973) emphasizes the lack of standardization in measurement of results of independent study experimentation; most writers favor independent study, he states, but bases for favoring it are quite varied. Attitudes are generally the basis of comparison rather than achievement and retention.

Environmental Press and Academic Experience Questionnaire

The importance of the press of the environment to the present study has been mentioned previously. A good deal of research on press of environments has been conducted. An entire body of research, for example, has developed around the College Characteristics Index (discussed by Stern, 1970). The development of this instrument--theory behind it, research surrounding it, etc.--by George G. Stern and C. Robert Pace is well reviewed in Stern's (1970) publication: People in Context: Measuring Person-Environment Congruence in Education and Industry. The ultimate goal of these researchers is to find the relevance of the characteristics of college environments, once defined, "to the prediction of behavior and performance of any institutional incumbent . . ." (p. 3). Stern and Pace contend that their tools can reliably describe the characteristics of students and collegiate environments in terms of comparable psychological dimensions. They " . . . have found that colleges differ systematically in the kinds of students they attract and in the experiences to which they are exposed" (p. 168).

Another broad general summarization of the studies done of college environments has been made by Kenneth A. Feldman and Theodore M. Newcomb (The Impact of College on Students, 1969). Feldman and Newcomb are strong supporters of interactional studies and are of the same orientation as Pace and Stern in looking at unique characteristics which are relevant in fitting people to environments.

Another researcher who has devoted a great deal of effort to environmental study is Alexander W. Astin (The College Environment, 1968). Astin's most important goal has been the identification of major differences in undergraduate college environments and developing sets of objective measures for these environmental differences.

Siegel and Siegel (1966) examined the college environment and made a strong point of the importance of instructional press to GPA. Siegel and Siegel feel that the instructor is the key determinant for the instructional press operating in a class:

> He communicates an attitude to his students by such behavior as his enthusiasm for the subject matter, the organization of his lectures, his rapport with the class, and so on. By his actions, the instructor implies to his students that he regards them as intellectually mature or immature: i.e., as capable of coping with appropriately difficult subject matter and assuming appropriate responsibility for their own learning, or as requiring an over-simplified presentation and continual supervision in order to learn (p. 150).

They go on to say:

> An extrinsic instructional press sensitizes students to the potentially punitive and threatening (inhibiting) aspects of the instructional environment. An intrinsic instructional press sensitizes students to the potentially supportive (facilitative) aspects of the instructional environment (p. 151).

Siegel and Siegel (1966) were careful to point out that inhibitors and facilitators "may be" effective as opposed to "will be" effective;

whether these potential inhibitors and facilitators become such, in fact, depends on the instructional press. The facilitators are effective only when the press is intrinsic, and the inhibitors actually inhibit performance only when the instructional press is extrinsic. Performance by low ability students, therefore, can be maximized by placing them in protective educational settings (where the instructor cannot probe in an intellectually embarrassing fashion and where the student does not feel unduly threatened by his failure to comprehend the instructional content). High ability students perform best in classes wherein their intellectual resources are stretched such as in a probing and intense discussion.

Purpose of the Study

The purpose of the present study was twofold: first, it investigated the possibility and nature of the occurrence of ATIs at the college level. Secondly, the study examined the degree of retention due to different methods of instruction. The aptitude investigated was achievement as measured by GPA and the treatment was two types of instruction roughly categorized as to degree of overt classroom involvement. GPA is felt to reflect the effect of many personological variables on achievement. High GPA students have obviously done well with the traditional structure; however, different structure might be a productive alternative for the lower GPA students. Though ATI research, as has been cited, is primarily concerned with disordinal interactions, this study did not anticipate that result. By using the broad aptitude of GPA, higher GPA students would be expected to do better across both

conditions, and ordinal interactions might more logically be anticipated.

This study constitutes a test of parts of the Siegel and Siegel (1966) model and is a more thorough study than (but similar in purpose to) that of Webb and Baird (1968). The independent variables were GPA and method of instruction; the dependent variable was test achievement. An intrinsic instructional press was created for both methods of instruction, and a good deal of guidance was provided for all students but with much flexibility. The expectation was that lower GPA students would do best with little overt classroom involvement and high GPA students would do best with more involvement. Another expectation was that students would have higher retention of the material learned in the more self-paced condition. As stated above, the second purpose of the study was the examination of possible differential retention of material. If one group of students retained more, the two types of subject matter (association material or problem-situations) tested on should be compared to determine if there is interaction with the other variables.

CHAPTER II

METHOD

Subjects

The subjects (Ss) were students at The University of Tennessee registered for Educational Psychology 2430 (Child Study) for Fall Quarter 1973. Fifty-four female undergraduates, most of whom were in their second or third year of college, enrolled in two sections of this course. The age range was from 19 to 49 with only two Ss above the age of 26; the mean age was 21.2, and the median was 21. The course is required of all elementary education majors and deals with various aspects of psychology as applied to the education of children.

Students signing up for the two sections were assigned through the normal processes of registration and scheduling and could, therefore, be considered a reasonably representative sample of females who take this course at The University of Tennessee.

After explanation to each class of the procedures to be used in the course, anyone desiring to drop or change sections was given the opportunity to do so. In an attempt to avoid differential achievement because of the time of class, only Monday, Wednesday, and Friday morning classes were used.

Procedures

The course content was divided into ten units; five units comprised half the course. A learning packet or study guide was prepared

for each unit. Treatment A was the most intensive classroom involvement. Under this condition, Ss met in a group twice per week for lecture and/or discussion-questioning about the content for that session. This treatment was as much as possible like the condition described by Siegel and Siegel (1966) under which high ability students were seen to thrive (intensive questioning, etc.). Under this condition the instructor lectured occasionally but was open to interruption by students for questions and comments. The class primarily consisted of intense discussion led by the instructor. Students were probed for opinions, logical inferences, and applications of the material. Students were called upon by first names and were encouraged to take stands on controversial issues and then defend their stands. They took a unit check-up at the end of each week.

Treatment B had no class discussion, but Ss were encouraged to come in for one (or more) unit check-up(s) each week. A room was made available for two hours, three days per week (their class time plus one hour) where they could take the check-up(s) when they desired. An assistant was available at these times for proctoring and/or helping. Both the instructor and assistant were available on an appointment basis. Study guides for the units were given out at the check-up session each week under treatment A condition and when desired by the B treatment condition. The self-paced Ss could take as many unit check-ups as they desired, up to the maximum--five plus extra credit--when they desired. The entire procedure was explained on the second day of class when initial assignments were made.

Ss were given points for taking check-ups (a minimum number of such points were required for certain grades; e.g., four out of the possible seven of these check-ups had to be taken under both treatment conditions for an "A"); the grades did not count toward the final course grade. If Ss made less than 70 percent on these check-ups they were strongly encouraged to study the material further and, if in the self-paced condition, given an opportunity to retake the check-up. The check-ups were graded by the Ss themselves and then turned in after being reviewed by them. Examinations were graded by a teaching assistant and according to pre-established criteria.

In addition to the regular classroom or self-pacing, all Ss were required to be involved in a field experience. They were assigned to teachers in the public schools in their particular area of interest and/or at the level they were planning to teach. These experiences reportedly varied considerably from teacher to teacher, and some Ss only graded papers while others taught classes. In conjunction with the experience, all Ss were required to attempt a behavior modification project on a child and write-up their steps taken and results. These requirements were completed by at least a two-hour session in the public school per week for the duration of the quarter.

Both methods utilized the same books, study guides, check-ups, and tests. The primary textbooks were W. C. Becker, S. Engleman, and D. Thomas' Teaching: A Course in Applied Psychology (1971) and B. J. Wadsworth's Piaget's Theory of Cognitive Development (1971). Both sections were taught by the same instructor. The study guide for each weekly unit consisted of objectives for the unit, a pre-test, activities

to complete (a part of the semi-programmed textbook and/or other readings), and a post-test or self-test to see if the Ss were ready for the check-up.

Ss were classified or blocked as to "high" or "low" based upon relative standing by GPA. They were then randomly assigned to sequence of treatments in a counter-balanced design. One-half of the Ss, with an equal number of high and low GPAs represented, received one treatment the first half of the quarter while the other half received the other treatment; at the mid-term the treatments were switched.

GPAs for all Ss were obtained from their college record. Transcripts of Ss were obtained from the records office of the University, and only students with similar course histories were utilized. Permission to use this data was granted by the University provided the information was kept confidential, that the anonymity of the Ss was insured, and that the material was reported only as grouped data.

Dependent Measures

Mid-term and final examinations. The mid-term and final examinations were the dependent measures and were required. The final examination did not duplicate content of the mid-term examination. These examinations were composed primarily of the items from the check-ups or similar items. The points made on these examinations did count toward the final course grade; a minimum number of points was required for an "A" or a "B." The overall course grade was criterion-referenced. Numerous activities, such as optional units, were provided on which Ss could make points so that all Ss could conceivably have made an "A."

The mid-term and final examinations were made up of two kinds of questions: association items and problem-situations (see Appendix for examples). These kinds of items were included to control for any "test-wiseness" which had developed by the Ss on the multiple choice college examinations which predominate at lower level courses. With these more unique measures it was hoped that scores would be more representative of learning rather than of ability to take multiple choice examinations. The association-type examination is described by Jennings (1968). Jennings says the word-association test is a highly reliable and valid testing procedure. It is viewed as a device to measure concept formation which is better than multiple choice tests with respect to the requirement for recall of information learned as opposed to the recognition required by multiple choice tests. Though more difficult to grade, the association examination is easier to construct and, as stated, requires a different level of learning than multiple-choice examinations.

The problem-situation items were designed in hopes of tapping higher levels of learning and application than is possible using multiple-choice examination items. Students are required to use learned concepts and to combine them in new and unique ways after recognizing particular stimuli in a situation and come up with the appropriate recombination of elements to fit the new situation. If differential retention of course material were to be found under the two treatment conditions, examination of possible interactions between these levels of test material and the other variables would prove productive.

Retention examination. The retention examination, given two and one-half months following the course, consisted of equally weighted numbers of items similar to the mid-term and final examinations given during the quarter. Many of the items were duplicated.

The retention examination was given the next quarter but was mentioned near the end of the course. Students addresses were obtained so they could be contacted. They were encouraged to plan on attending. This examination was given in a night (early evening) session three weeks prior to the end of the next quarter. Ss were paid a sum significant enough to encourage their attendance ($5.00), and alternate times to take the retention examination were offered. Forty-four Ss returned to take this examination. Forty of these completed acceptable examinations.

Attitude Measures

Press of class. A questionnaire regarding the press of the class was administered near the end of the course and asked the students to rate the class and instructor relative to other classes and instructors on certain variables (described under "Instructor Behavior," below) which make up an intrinsic instructional press.

Academic experience questionnaire (AEQ). One other type of measure was utilized to see what other kinds of experiences the Ss had during the quarter which might have enhanced or retarded the effect of the experimental conditions. The AEQ (patterned after a similar instrument of Gruber and Weitman, 1962) was given after the course was well

under way (just after the middle of the quarter) and asked for descriptions of the other academic experiences the Ss were having during the quarter to get an overall picture of how significant this one course was as a "drop in the bucket" of total experiences.

Instructor Behavior

The instructor tried to create an intrinsic instructional press for both conditions. This press was operationally defined in terms of Siegel and Siegel's (1966) definition. Criterion-referenced grading with the student determining her own grade was used. The instructor attempted to deal with the students openly and honestly concerning the study and their part in it and solicit their cooperation. A number of optional activities were offered for the student to have some choices in regard to what she did for the course. Ss were encouraged, not forced, to remediate if they made low scores on unit check-ups. As required by Siegel and Siegel (1966), the instructor and assistant saw themselves as helpers, facilitators, and resource persons and not disciplinarians. Attempts were made to plan and organize material well in advance and leave room for individual variation.

For this intrinsic press, Siegel and Siegel see the instructor as being of greatest importance; an attitude of enthusiasm for the subject matter must be communicated. Good rapport is essential; tests must be seen as fair; students must be given some responsibility for their own learning; and the instructor should be perceived as a facilitator. Attempts were made by the instructor to meet these criteria.

Experimental Design

Figure 1 shows the design used to be a 2 X 2 repeated measures design. The blocking variable was GPA which was split at the mid-point of the distribution to give the two levels. In some analyses, GPA was used as a covariate. Assignment to sequence was random, and an analysis to test for a sequence effect was done. In the first analyses the repeated measures were the mid-term and final examinations. The retention examination which was given after these initial portions of the study, was used as one of the repeated measures in later analyses.

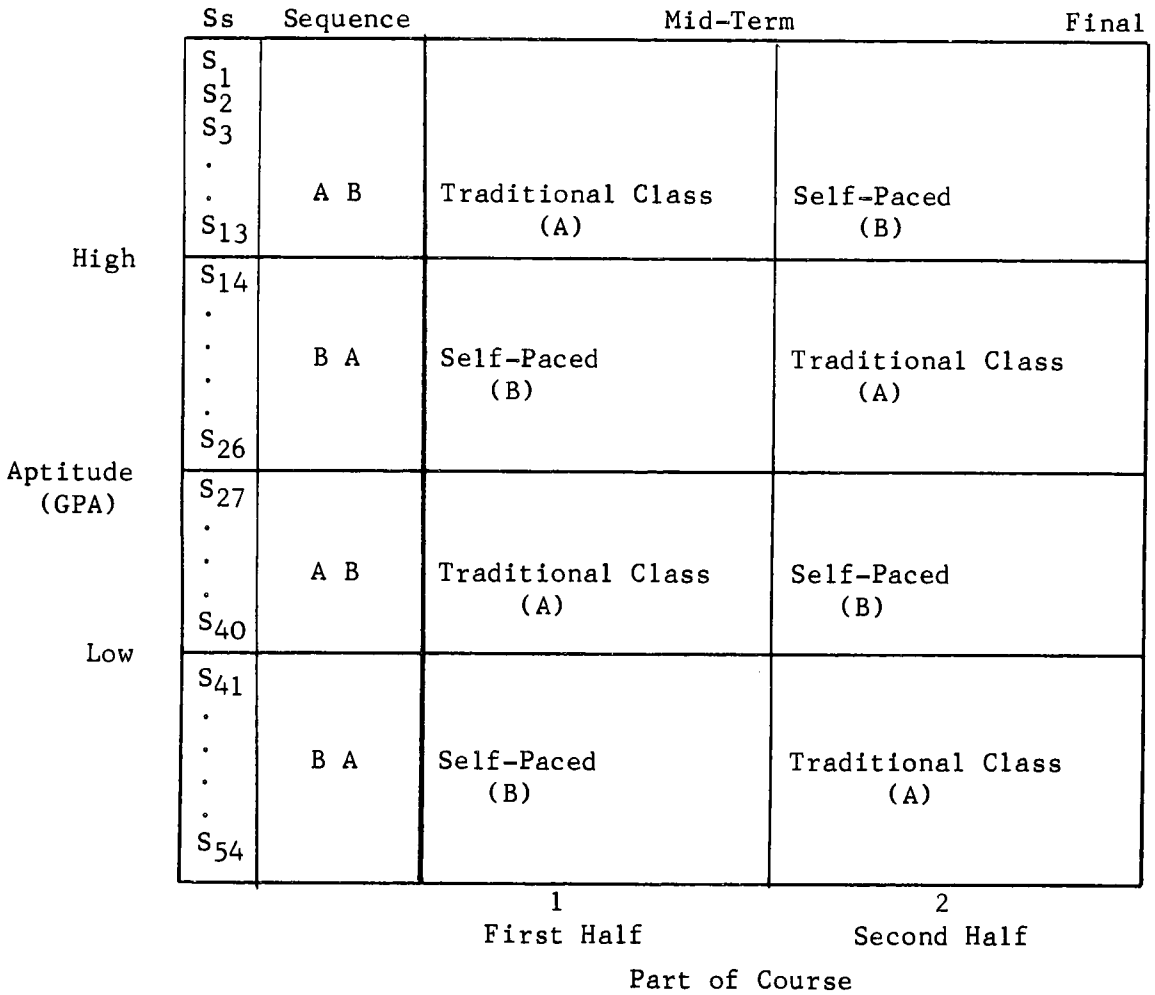


FIGURE 1. Design of the study.

CHAPTER III

RESULTS

Several analyses and surveys were used to examine the data obtained from the study. Because of the nature of the data, some analyses, of necessity, overlap slightly. In some cases, one analysis raised questions leading to further analysis to refine the answer sought (for example, the first analysis for retention of material was felt to potentially mask possible differences, so two separate analyses were done in addition to it). The reliability of the examinations is reported with the major findings. The possible sequence effect is examined followed by the results of the retention examination and analyses thereon. The results of the surveys on the press of the class and the "Academic Experience Questionnaire" conclude the chapter.

Mid-Term and Final Examinations

Reliability of examinations. The mean of the mid-term examination was 77.18 out of a possible 100 points with a standard deviation of 11.77. The internal reliability was calculated (using number of questions correct rather than points made) to be .796 by Kuder-Richardson Formula 21. The mean number of points on the final examination was 78.69 out of a possible 100 with a standard deviation of 14.45. Reliability (again using number of questions correct rather than points) according to the same reliability formula (KR 21) was .900. A t-test of difference scores revealed the tests to be nonsignificantly different

at the .05 level (t = .925; d.f. = 40). The tests were, therefore, considered equivalent (or at least not different) in terms of mean score.

Analyses of variance. Lindquist Type I (Lindquist, 1953) analyses of variance were performed on the dependent variable--achievement under the two treatment conditions--with GPA as the blocking variable. All Ss were used in one analysis (the top one-half versus the bottom one-half) and, in a manner similar to Webb and Baird (1968), the more extreme cases on the blocking variable were used in the two following analyses--the extreme forty (twenty at each end) and then the extreme twenty. Table 1 presents a summary of the analyses.

It is to be noted, again, that the repeated measure is not the same as it will be in the analysis for a sequence effect. The sequence analysis uses the mid-term and final scores as the repeated measure. In the present analysis the repeated measure is "treatment" (traditional class or self-paced); the first measure for one-half the Ss is the mid-term examination, and for the other half it is the final examination.

In all three analyses there is a significant F for "Groups." Higher GPA students achieved better than lower GPA in all cases, as was expected.

In the first two of the three analyses, the traditional class condition was significantly superior to the self-paced condition. This was not the case when only the ten top and ten bottom Ss were used; these twenty Ss did not differ significantly in their performance under the two conditions. The same results are found using the

TABLE 1. Summary of Results of the Analyses of Variance for All Ss, 40 Extreme Ss, and 20 Extreme Ss

Source	df	SS	MS	F
All Ss				
Between Subjects	53	14873.88		
A (Groups)	1	3641.25	3641.25	16.86**
Error B	52	11232.62	216.01	
Within Subjects	54	3726.50		
B (Treatment)	1	366.69	366.69	5.72*
A X B	1	26.69	26.69	<1.00
Error W	52	3333.12	64.10	
Total	107	18600.38		
Extreme 40 Ss				
Between Subjects	39	13641.06		
A (Groups)	1	4836.12	4836.12	20.87**
Error B	38	8804.94	231.71	
Within Subjects	40	3099.00		
B (Treatment)	1	460.88	460.88	6.75**
A X B	1	44.88	44.88	<1.00
Error W	38	2593.25	68.24	
Total	79	16740.06		
Extreme 20 Ss				
Between Subjects	19	7867.94		
A (Groups)	1	6300.12	6300.12	72.33**
Error B	18	1567.81	87.10	
Within Subjects	20	1178.00	58.90	
B (Treatment)	1	152.12	152.12	2.73
A X B	1	22.43	22.43	<1.00
Error W	18	1003.44	55.75	
Total	39	9045.94		

*$p < .05$.

**$p < .01$.

Geisser-Greenhouse conservative degrees of freedom procedure (Dayton, 1970) to avoid having to test for homogeneity of variance and covariance matrices. Table 2 shows the cell means for the various analyses.

In all cases the F ratios for the interaction were less than 1.00 and, therefore, nonsignificant. When plotted on a graph, lines for the groups were almost parallel though they were further apart as the more extreme cases were used.

Sequence effect. After blocking for high and low GPA, Ss were randomly assigned to a sequence; the mean GPAs were not the same for the two groups (2.63 versus 2.70 with standard deviations of .44 and .65, respectively), however. Thus to test for an effect for sequence, an analysis of covariance was necessary using GPA as the covariate.

A Lindquist Type I analysis of covariance was used to test for this sequence effect; i.e., does the traditional class condition result in better achievement if it comes in one position or the other and/or does the self-paced condition result in differences due to its position in the sequence. Table 3 shows the results of this analysis. There is no sequence effect, nor is there a significant difference between the mid-term and final examination scores. A significant interaction is observed. This is attributed to the fact that the traditional class condition did significantly better than the self-paced condition for both ability levels. Adjusted cell means were: cell 1,1 (mid-term, class) 79.73 and cell 1,2 (final, self-paced) 77.03; cell 2,1 (mid-term, self-paced) 76.12 and cell 2,2 (final, class) 80.79.

TABLE 2. Cell Means for the Ss Used in the Analyses of Variance for the Major Treatment Effects

			Treatment Condition	
			Traditional Class	Self-paced
GPA	High	Top 10	91.20	88.80
		Top 20	87.65	84.35
		All	85.77	83.12
	Low	All	75.14	70.50
		Bottom 20	73.60	67.30
		Bottom 10	67.60	62.20

TABLE 3. Summary of the Results of the Analysis of Covariance for Sequence as the Dependent Variable with GPA Used as the Covariate

Source	df	SS	MS	F
Between Subjects	52	8674.39		
A (Test)	1	0.13	0.13	<1.00
Error B	51	8674.26	170.08	
Within Subjects	53	3726.50		
B (Sequence)	1	26.01	26.01	<1.00
A X B	1	366.68	366.68	5.61*
Error W	51	3333.81	65.37	
Total	105	12400.89		

*p <.05.

Follow-up examinations for simple main effects showed only one significant difference. The self-paced condition at mid-term was significantly different from the final for class (76.12 and 80.79, respectively; $p < .05$).

The importance of the statistical assumption of homogeneity of regression slopes is recognized as a prerequisite to doing the preceding analyses. A survey of several standard statistical textbooks (Winer, 1971; Kirk, 1968; Dayton, 1970), however, failed to reveal a clear-cut procedure for this calculation in a covariance repeated measures design. As a substitute for, or estimate of, this calculation within group regression slopes were calculated in this analysis and in similar ones to follow. Since these slopes are close, the analysis is continued as if the assumption were met. Regression slopes for the cells were: cell 1,1--.09; cell 1,2--.17; cell 2,1--.14; and cell 2,2--.09.

Retention Examination

The retention examination was split according to whether the half was covered in traditional class or in the self-paced condition. Care was taken in construction to weight them equivalently. Reliability of the total test was calculated to be .85 (KR 21; again, based upon the number of items correct as opposed to the number of points). The mean points made was 51.85 with a standard deviation of 15.46. The mean of the portion based upon the mid-term was 27.75 out of a possible 50 with a standard deviation of 9.42 (KR 21 reliability of this half of the examination was .47). The mean of the portion of the examination based upon the final or last half of the course was 24.1 with a standard

deviation of 8.72 (KR 21 reliability on this half of the examination was .44). For convenience in analyzing and to have a common basis of comparison to other scores, these scores were multiplied by two for the subsequent analyses.

In answering the question "Is there a differential retention of material learned due to the method of instruction by which it was learned?" three major initial analyses were done. The first analysis was an overall measure of traditional class versus self-paced irrespective of whether the treatment was given the first half or the last half of the course. In other words, the scores of those having class the first half of the quarter (test score on material up to mid-term--first half of follow-up examination) were combined with the scores of the Ss having class the second half of the quarter and considered as one measure which was compared to the similarly combined scores (some mid-term "halves" and others final "halves") on the self-paced portion. A related samples t-test was done to test significance of the difference between these two means (class--52.90; self-paced--50.68; standard deviations were 15.79 and 17.42, respectively). The difference was found to be nonsignificant ($t = 1.12$; d.f. = 35, $p > .05$). This test might be misleading in that any differential effect for the time difference (which is collapsed across and ignored in this test) or initial differences in scores is lost.

Due to the possibility of confounding from a number of factors (different time lapses between tests, possibility of a fatigue effect causing scores to be different on the two halves of the retention examination, the possibility of one-half being more difficult, the very strong

possibility of a "wash-out" effect for results on the halves should they be in opposite directions, and a number of other minor considerations), it was decided that the retention examination should be broken down and the two halves analyzed separately. The prior counter-balancing could possibly correct for some of these factors but not for all of them; for example, the fatigue effect should be accounted for by it but not the possibility of a "wash-out" effect; in fact, the latter would be masked by this counter-balancing procedure.

The next analysis compared, therefore, the original mid-term scores--achievement at that time--to the scores on the mid-term portion of the retention examination. Twenty-one Ss who had had class and nineteen who had had self-pacing had handed in completed follow-up examinations. Two of the traditional class Ss were randomly selected out in order to have equal n for the Lindquist Type I analysis of covariance on these retention scores. Covariance was used here for the same reason and with the same covariate as in the analysis for the sequence effect--Ss in the two groups had different GPAs (2.66 versus 2.82). The first portion of Table 4 shows the results of this analysis. Within group regression slopes for this analysis were: cell 1,1 (class, mid-term) .11; cell 1,2 (class, retention) .12; cell 2,1 (self-paced, mid-term) .14; and cell 2,2 (self-paced, retention) .10.

Overall, there was no difference in traditional class versus self-paced as had been found at the mid-term. As might be expected, a good deal had been forgotten, so there was a highly significant difference between the mid-term scores and the scores on the part of the retention examination based upon the mid-term. The interaction, which

TABLE 4. Analyses of Covariance of Class Versus Self-paced on Mid-term and Final and Respective Retention Examinations

Source	df	SS	MS	F
Mid-term				
Between Subjects	36	5424.53		
A (Class vs. Self-paced)	1	.32	.32	<1.00
Error B	35	5424.21	154.98	
Within Subjects	37	13787.50		
B (Test Scores)	1	10352.22	10352.22	118.09**
A X B	1	366.96	366.96	4.19*
Error W	35	3068.32	87.67	
Total	73	19212.03		
Final				
Between Subjects	36	9332.76		
A (Class vs. Self-paced)	1	492.97	492.97	1.95
Error B	35	8839.79	252.57	
Within Subjects	37	24841.50		
B (Test Scores)	1	21322.75	21322.75	226.16**
A X B	1	218.96	218.96	2.32
Error W	35	3299.79	94.28	
Total	73	34174.26		

*p <.05.

**p <.01.

is the comparison of most interest to this study, is found to be significant at the .05 level. If graphed, this interaction would have the lines crossed: the traditional class mean (adjusted) on the mid-term being 82.18 and the retention mean of 54.45 while the self-paced mean was 77.92 and the retention mean for this was 58.97. Covariates were 2.65 and 2.82 for class and self-paced, respectively. Though the Ss in the self-paced condition during the first half of the course learned less originally, they retained more of what they had learned and more material in an absolute sense three months later.

Follow-up examination for simple main effects (due to the significant interaction) showed two major significant differences. Both the traditional class and self-paced groups declined significantly from the mid-term to the retention examination ($p < .01$). Class and self-paced were not, however, significantly different from each other at either mid-term or retention time. Both groups retained significantly less, but Ss who had had self-pacing retained more than the traditional class group, resulting in the significant interaction.

Results of the analysis of the data concerning the final examination and retention examination are somewhat different (Table 4). It may be observed that there is no significant overall difference between traditional class and self-paced. Again, there is a very significant loss of material, and retention scores are much lower than original scores. Though the interaction is not significant, the trend is in the same direction as on the mid-term. The self-paced Ss (last part of the course) who returned for the retention examination had made higher on the final originally (83.8 versus 82.04 made by the ones who had had

class), and their retention was greater than the traditional class section on the retention examination (45.15 for class and 53.69 for self-paced Ss). The covariates for these adjusted means were 2.65 (for class) and 2.82 (for self-paced). The follow-up analysis for simple main effects were quite similar to the results for the mid-term (these were analyzed even though there was no significant interaction). Retention scores were significantly lower than the final scores for both conditions ($p < .01$). The traditional class mean was not significantly different on the final from the self-paced final mean. The difference in retention means approaches significance (45.15 for class and 53.69 for self-paced where the critical $F = 4.17$ and the calculated $F = 4.00$). Within group regression slopes for this analysis were: cell 1,1 (class, final) .14; cell 1,2 (class, retention) .10; cell 2,1 (self-paced, final) .11; cell 2,2 (self-paced, retention) .14.

Differential Retention of Material

Since there was a significant interaction found with the mid-term examination and retention examination with the two treatment conditions, further analyses were undertaken. The two halves of the retention examination (mid-term and final portions) were divided into two parts based upon the association item scores and the applied, problem-situation items. Each half of the examination had a possible 50 points; this being divided into 32 points for association items and 18 points for applied items. In order to equate these parts, the latter was multiplied by 1.778, thereby making each portion of each half of the examination have a possible 32 points.

A Lindquist Type III analysis of covariance was used to do these analyses. GPA was used as the blocking variable and the initial examination score was used as the covariate in order to equate the scores on the amount of knowledge originally learned. The repeated measure factor was association items versus applied items on the examination. The third dimension was self-paced versus traditional class. After dropping Ss from some cells to obtain equal n, eight Ss were left in each cell.

Table 5 shows the results of this analysis. None of the interactions were statistically significant. The only main effect which was statistically significant was the type of item or material. Overall, Ss scored significantly better on applied material than they did on association material on both the mid-term and the final portions of the follow-up examination. Means adjusted for the covariate (original examination score) were 16.5 and 20.1 for the mid-term and 14.1 and 17.1 for the final, association items and applied items, respectively. The adjusted means and within group regression slopes for the cells in the analysis are shown in Table 6.

Attitude Measures

Press of class. The "Press of Class Questionnaire" indicates that an intrinsic instructional press was created. A Likert scale was used with "Strongly Agree" at one end and "Strongly Disagree" at the other; numerical values of 5 and 1 were assigned, respectively, with 4,3, and 2 being intermediate values. A 3 indicated relative neutrality. The Ss were asked to place "X" on each continuum to indicate their perception of the experimental condition (educational psychology course)

TABLE 5. Analyses of Covariance of High GPA and Low GPA Subjects in Traditional Class Versus Self-paced on Mid-term and Final Examinations Broken Down into Applied- Versus Association-Types of Test Items

Source	df	SS	MS	F
Mid-term				
Between Subjects	30	742.15		
A (Class)	1	45.81	45.81	1.78
B (GPA)	1	0.31	0.31	<1.00
A X B	1	1.82	1.82	<1.00
Error B	27	694.21	25.71	
Within Subjects	32	718.52		
C (Association vs. Applied)	1	199.52	199.52	11.09**
A X C	1	0.02	0.02	<1.00
B X C	1	0.02	0.02	<1.00
A X B X C	1	15.02	15.02	<1.00
Error W	28	503.94	18.00	
Total	62	1460.67		
Final				
Between Subjects	30	892.16		
A (Class)	1	12.83	12.83	<1.00
B (GPA)	1	0.05	0.05	<1.00
A X B	1	21.58	21.58	<1.00
Error B	27	857.70	31.77	
Within Subjects	32	634.50		
C (Association vs. Applied)	1	147.02	147.02	8.51*
A X C	1	2.64	2.64	<1.00
B X C	1	0.76	0.76	<1.00
A X B X C	1	0.39	0.39	<1.00
Error W	28	483.69	17.27	
Total	62	1526.66		

*$p < .05$.

**$p < .01$.

TABLE 6. Adjusted Cell Means and Within Group Regression Slopes; Mid-term and Final Portions of the Follow-up Examination

Group	Mid-term		Final	
	Adjustment Means	Regression Slope	Adjustment Means	Regression Slope
Traditional Class				
High GPA				
Association items	20.00	.319	11.22	.022
Applied items	24.50	.265	14.72	.173
Low GPA				
Association items	12.63	.318	17.90	.333
Applied items	15.12	.258	21.28	.227
Self-paced				
High GPA				
Association items	24.55	.017	11.16	.287
Applied items	27.18	-.238	14.16	.752
Low GPA				
Association items	8.95	.073	16.08	.234
Applied items	13.45	.188	18.34	.430

and the instructor(s) and an "O" to represent experiences with other courses and instructors taken collectively.

Table 7 shows the average values assigned to each question and the difference between the experimental course and their other courses. All questions show the press of the experimental course to be more intrinsic than other courses. Though some differences are fairly small, they are all in the desired direction.

Academic Experience Questionnaire. Students were asked to divide up the time, by percentages, they spent on academic tasks for the experimental course and on their other courses taken collectively. The questionnaire was in two parts: "Time Spent in Class" and "Time Spent Out of Class." The first category was broken down into four major areas with various subdivisions in each area. Table 8 shows the mean time, in percentages, spent by students in the various categories and subdivisions thereof, for the experimental course and for the student's other courses taken collectively ("In Class").

Large differences can be noted in the first three categories and their subdivisions. In the first area, for example, students reportedly spend 21.8 percent of their class time in the experimental course "listening to the instructor" while an average of 43.8 percent of class time in other courses is spent in that manner. Of the 21.8 percent listening time, 9.7 percent in the experimental course is spent trying to follow the instructor and 12.0 percent is spent in thinking critically and actively about what is being presented. This is compared to 26.7 of the 43.8 percent in other courses being spent trying to follow the

TABLE 7. Composite Results of the Press of Class Questionnaire

Item	X Ed Psyc	O Other Courses	X-O Difference
1. The student has control of determining his/her own grade.	4.183	3.216	.967
2. The criteria for "knowing where I stand" are quite clear.	3.983	3.216	.767
3. A number of ways of reaching the goals of the course(s) is (are) provided.	3.716	2.800	.916
4. The instructor(s) I have is/are humane; he/she (they) treat(s) me with respect as a person--an adult capable of making my own decisions--not an object to be manipulated.	4.616	3.666	.950
5. The instructor(s) is(are) honest and open with students.	4.516	3.633	.883
6. The instructor(s) is(are) helper(s), facilitator(s), and resource person(s) and not a disciplinarian(s) who focus(es) on continual supervision and punitiveness.	3.983	3.316	.667
7. The instructor(s) has(have) enthusiasm for the subject.	4.516	3.883	.663
8. The instructor(s) has(have) good rapport with the class.	4.610	3.372	1.238
9. Exams are fair and just.	3.728	3.203	.525
10. Subject matter of the course is not oversimplified but, generally, of an appropriate difficulty level for me.	3.816	3.733	.083
11. Students are seen as intellectually mature; capable of coping with appropriately difficult subject matter and assuming appropriate responsibility for their own learning.	4.450	3.866	.584
AVERAGE	4.124	3.4459	.7465

TABLE 8. Composite Results of Academic Experiential Questionnaire Time Spent in Class

Item	Ed Psyc	Other Courses
1. Listening to instructor	21.8[a]	43.8
a. Trying to follow instructor	(9.7)	(26.7)
b. Thinking actively and critically about the material being presented	(12.0)	(17.1)
2. Discussion and recitation	24.6	16.2
a. Answering questions and doing exercises to show that you have done assignments	(11.8)	(8.1)
b. Real discussion that involves thinking actively and critically	(12.9)	(8.1)
3. Laboratory work (include field work)	27.8	15.2
a. Follow instructions with little or no opportunity to make decisions	(5.5)	(5.1)
b. Carrying out assignments with some important decisions about what to do and how to do it left up to student	(10.6)	(4.8)
c. Independent work with most of the important decisions left up to the student	(11.7)	(5.3)
4. Taking tests	25.9	24.9
a. Objective tests (true false, multiple choice, completion, etc.)	(7.1)	(9.6)
b. Problems involving application of methods applied by instructor or textbook	(10.8)	(7.1)
c. Essays and problems in which major emphasis is on your own thought	(8.0)	(8.1)

[a]All values are percentages.

instructor, and the remaining 17.1 percent is being spent thinking about the material being presented.

The experimental course has considerably more discussion and recitation and considerably more laboratory or field work. The students see much more of the laboratory and field work in the experimental course as being done independently than in other courses.

"Test taking" appears to be about equal in the percentage of time consumed for the experimental course as compared to other courses.

Table 9 illustrates the distribution of the percentages of time spent "Out of Class" in the experimental course in comparison to other courses, again, with the latter taken collectively. There is very little difference in the distributions. Students seem to spend about equal percentages of time in doing out of class work in the various categories for the experimental course as they do for the out of class work for other courses.

About 40 percent of out of class academic work is spent reading; most of this being in required textbook(s). About 16 percent of time is spent discussing, 15 percent for writing, 16 percent for perfecting skills, and 13 percent for research and creative work. In most of the subdivisions, the distributions are quite similar in pattern.

TABLE 9. Composite Results of Academic Experiential Questionnaire
Time Spent Out of Class

Item	Ed Psyc	Other Courses
1. Reading	40.0[a]	37.6
a. Reading assigned textbooks	(25.1)	(24.0)
b. Assigned reading other than textbooks	(9.1)	(9.0)
c. Reading unassigned material connected with subject matter of course(s)	(5.8)	(4.4)
2. Discussing	15.6	15.8
a. Review and clarifying notes, textbooks, classwork	(8.0)	(8.4)
b. Thinking actively and critically about subject matter explored in class	(7.5)	(7.4)
3. Writing	16.3	14.8
a. Routine themes and reports	(6.7)	(7.5)
b. Papers in which major emphasis is on your own thought and investigation	(9.6)	(7.3)
4. Perfecting skills	15.5	17.6
a. Doing problems	(5.4)	(6.2)
b. Practicing (for example: music)	(3.9)	(4.8)
c. Memorizing	(6.2)	(7.1)
5. Research and creative work	12.8	13.8
a. Scientific (experimental and theoretical)	(7.3)	(6.4)
b. Aesthetic (creative writing, composing, artistic creation, etc.)	(5.5)	(7.4)

[a]All values are percentages.

CHAPTER IV

DISCUSSION

The present study was begun with two major purposes. The first purpose was to attempt to describe some of the complexity of the teaching-learning process in the college classroom by looking at (or for) aptitude-treatment interactions. The second purpose was to examine the amount of retention students had of material learned under the two different treatment conditions chosen; in other words, not only was the study looking at the immediate result of differential treatments, but a major interest was in the lasting (and possibly differential) effects of the treatments. More specifically, the two treatments were a more or less traditional class and a self-paced condition which were examined in relation to two aptitude levels as demonstrated by GPA.

Aptitude-Treatment Interactions

The major interest of the study was in the ATIs which might be discovered. Initially, the interest was in the mid-term and final examinations. Based upon the results of these analyses, the retention of material learned was investigated to determine the nature and extent of possible interactions.

Mid-term and final examinations. The hypothesized ATIs were not found. Neither the research of Siegel and Siegel (1966) nor that of Webb and Baird (1968) was supported by this study. Both of these studies would predict higher achievement of lower ability (or achieving) students

in the self-paced condition with either higher achievement for the higher ability students in the traditional class or little difference for the higher students under the two treatments. All Ss did better under the traditional class condition.

Both Milton (1972) and Siegel and Siegel (1966) have emphasized the complexity of the many forces interacting in the teaching-learning environment. Possible reasons for the results obtained abound. One of the most obvious possibilities might be the amount of time spent in class. The Ss in the class section spent (probably) two hours each week more on the content of the course than did their counterparts in the more self-paced condition. The rather highly structured discussion situation would almost have to bias the study in favor of those coming to class. The instructor would obviously, even if unintentionally, keep discussion more in the areas he thought more important and test likewise. Even though study guides were made and followed quite closely, the traditional class condition would inevitably get a little greater edge than those with study guides only. The extra time most probably spent and the possible bias on the material discussed could very well account for the fact that those in class at both ability levels performed significantly better.

Changes in Ss might have occurred which were not measured in the present study. Bigelow and Egbert (1968) found achievement in traditional environments to be equal to independent study achievement; however, more intellectual efficiency and responsibility result from independent study. While these latter two factors were not measured by either Webb and Baird (1968) or the present experimenter, retention results of the

present study might yield tangential support for this statement. The point is made in Travers (1973) that independent study is viewed favorably by most authors, many of whom do not evaluate the effects of independent study on students' achievement or attitudes. The study findings vary considerably, he goes on to say, according to the criteria used to evaluate.

Retention

Moore et al. (1973) measured retention of material from a college physics course one year after completion of the course by continuous progress (mastery) versus traditional instruction and found greater retention by the continuous progress students. Corley et al. (1971) found similar results in an introductory psychology course (they tested retention ten months after course completion). The retention examination in the present study is strongly supportive of the results of both of these studies even though there was only about a three month delay between the initial measure and the follow-up retention examination in the present study. The lapse from the mid-term examination to the follow-up retention examination on it was about three months while the final to the follow-up was slightly less than two months. The mid-term and follow-up show a significant interaction between the two variables, test and class/self-paced condition. Even though the traditional class group who returned to take the follow-up had scored better at the mid-term, they retained less than had the self-paced group who learned less originally as measured by the tests. This finding indicates that either the Ss in the traditional class knew the material less well; i.e.,

perhaps they simply "crammed" for the test upon which they were to be graded, or the self-paced group knew what they learned more deeply and lastingly.

The results of the final and follow-up on the material on the final had less dramatic results, possibly due to the shorter length of time the material had been "incubated." Of those taking follow-up examinations, the ones who had had the self-paced condition the last half of the course scored higher under this condition than those having class the last half (overall, including the Ss who failed to return for the follow-up, the class condition had done better). A disordinal interaction, therefore, as is found with the mid-term and follow-up should not be expected--based upon the results of the mid-term. The lines, however, do diverge with the self-paced group remembering more than the class group.

The disordinal interaction observed on the mid-term and follow-up shows, quite convincingly, that students who are self-paced do, indeed, retain more of what they learn than those who have been coming to class. Perhaps they structure their learning differently, or perhaps the fact that they structured it themselves made the difference. For whatever reason, the fact that after three months, in this case, and even longer in other research, students know more material when it was learned independently, brings into serious question the traditional belief that classes and teaching are necessary for the results which have been striven for. If scores on examinations and recollection of material from those examinations are what education is all about, perhaps the traditional college lecture and class periods should be abolished.

Though the interactions were of major concern in the present study, as an educator, one cannot help but be appalled at the great drop from the initial examinations (mid-term and final) to the follow-up examination. The two alternatives--teach students more and have them forget more and/or change methods and teach them less in order that they retain more of it--are both undesirable. The goal of the educator should be to get the most retention regardless of teaching (though most educators would probably tailor teaching to a particular method if it were proven that that method were better). In view of the findings of the study the additional question is posed: "How do we get more retention of what is taught?" Though studies have dealt with this issue in the past, little definitive evidence exists to put one method of teaching very much above another.

The examination of the possibility of complex interactions due to the type of material being learned and remembered by the two achievement levels under the two treatment conditions proved nonsignificant. Overall, of course, Ss scored better on problem-situation (applied) material than the association (recall) material. No differential effect was found, however, across achievement levels or treatment conditions. The same pattern was found in both the analysis of the mid-term and the portions of the retention examination based upon it and the final and portions of the retention examination based upon it.

The results of this analysis are to be interpreted cautiously. The within group variability was quite high and made the possibility of obtaining a significant F quite remote. A study designed with this as the major research question and with much more care taken in the

construction (validation, etc.) of the examination items would probably give a more dependable test of this research question. The results obtained from these analyses showed, however, that differential retention of the two types of material was not responsible for the overall difference found in retention of the self-paced group and the traditional class group.

Even though the above caution is given, the strong results significantly in favor of the retention of applied material are noteworthy. The goals of education, when stated, almost always cover the goal of generalizability of material learned. The teaching of concepts and principles which have general applicability and the bases of which will transfer from situation to situation is strongly advocated. The inclusion of the applied, problem-situation items on tests in this study and the orientation of the teaching in the course (including the learning packets) in this direction were aimed at this goal. Though the low overall retention in the present study was cause for dismay, the fact that applied material was retained significantly better than association or recall material was encouraging. It becomes obvious, once again, that focusing on details, facts, and recall (memorization) in teaching might well result in teacher satisfaction with "learning" at the end of a course or when a test is given but the low retention of such material brings into question the efficacy of promoting learning of such material and including it on tests. Perhaps the focus in all courses of the sort this study dealt with should be on the applied and generalizable kind of information which is of more practical benefit and retained better.

The above ideas are pointing toward the evaluation of instruction based upon the delayed consequences (what is retained) of instruction. The research on the effectiveness of instruction too often focuses on what happens in the classroom and immediate consequences thereof. Perhaps an alternative approach would be to look at the nature of the material retained sometime after a course is completed and then go back and try to reconstruct what kinds of things (experiences of the students actions by the instructor, etc.) are responsible for the retention such that other concepts could be presented differently and improve the retention thereof. This is an area open to research.

Press of Class

Not of major importance as a dependent measure but of interest and a prerequisite to the present study according to Siegel and Siegel (1966) was the establishment of an intrinsic instructional press. The results indicate that this was done. By establishing an intrinsic instructional press where the student is regarded as mature and capable, the low ability student, according to Siegel and Siegel, should have had her achievement facilitated and higher ability students should not have been inhibited. Gruber and Weitman (1962) tend to agree.

Though Siegel and Siegel refer to the instructional press of classrooms, they do not discuss the press of the entire school environment to any degree. They do advocate a very strong independent study orientation especially in earlier, broad survey courses in a student's college career. One is left to guess how successful one course being different can be in relation to the rest of a student's college

environment. It is clear that the one experimental course represented by this study was but a small "drop in the bucket" of the students' total college experience. The effects or potential effects of the course might well be "washed out" by the remainder or overall school environment. Most traditional colleges and universities, including the one in which this study was done, probably have extrinsic instructional presses. To obtain meaningful results in changing the approach in a classroom, it might first be essential to change the entire environmental press. This point must be considered in interpreting any results of the present study. Gruber and Weitman (1962) make this point quite explicitly: "Experiments conducted with a single course suffer from the defect that the students are involved in a number of other courses not under the investigator's control, which may have unknown or unexpected effects on the experiment" (p. 6-1). The results of the "Academic Experience Questionnaire" must be considered for the effect the conditions reflected by them may have in looking at the results of the present study, also.

CHAPTER V

SUMMARY

This study investigated the possibility and nature of the occurrence of aptitude-treatment interactions at the college level and the degree of retention of material due to two different methods of instruction. The aptitude investigated was represented by prior achievement as measured by GPA, and the treatment was two types of instruction: traditional lecture/discussion and self-pacing.

Fifty-four females at The University of Tennessee, predominantly in their second and third year of college, who were enrolled in a basic educational psychology course required of all elementary education majors were divided into two groups based upon GPA. They were then randomly assigned either to class or self-pacing in a counter-balanced design wherein they would receive one treatment for one-half the quarter and the other treatment the second half. All Ss were given study guides containing pre-tests, objectives, activities, and post-tests for ten weeks (one unit per week). Students in the traditional class condition were given check-ups (quizzes) each week while those who were self-paced were encouraged to take these check-ups at their convenience but attended no formal class sessions.

Important to the study was creation of an intrinsic instructional press in which students saw themselves (and perceived the instructor as seeing them) as mature and capable of coping with appropriately difficult subject matter. Good rapport was necessary; the instructor had to be

seen as facilitative rather than punitive; examinations had to be seen as fair and just; the instructor had to be seen as enthusiastic. A questionnaire was designed to survey the degree to which these conditions were met by the present instructor relative to other instructors. On all items, the experimental course and instructor were seen as providing a more intrinsic press than was provided by other courses and instructors.

A modified Lindquist Type I design was used with the repeated measure being the treatment rather than the trial. The dependent measures were scores on examinations. Results showed the traditional class condition to produce higher examination scores (using all Ss and the extreme 20 Ss at each end of the GPA spectrum; this was not the case using the two extreme 10 Ss; no significant difference was found in the latter case), and, as would be anticipated, higher GPA Ss scored significantly better than low GPA Ss. No interaction was found.

Forty Ss returned completed retention examinations about three months after the course was completed. A Lindquist Type I analysis of covariance was used to check on differential retantion from the mid-term to the follow-up (both experimental conditions) and from the final to the follow-up. In the first case, the Ss who had had class the first half of the course scored higher on the mid-term but scored lower on the follow-up than those who had had self-pacing, giving a significant interaction. Those having self-pacing last scored better on the final and on the follow-up. The interaction on this latter comparison was not significant though it approached significance (calculated $F = 4.00$; critical $F = 4.17$). This retention part of the study did show that

people retain better information they have learned under self-pacing.

A further analysis was done to investigate the type of material retained best, by whom, and learned under what condition. A Lindquist Type III analysis of covariance, with original scores being the covariates was done on the two types of material tested--applied, problem-situation and association items or recall information. No statistically significant interactions were found from this analysis. Across treatments and achievement levels, however, Ss scored significantly better on applied material than on association material.

Overall, the findings suggest that initial support for traditional class might not hold up when retention is examined three months later. What is learned under self-paced conditions is retained better. Under all conditions, applied material seems to be learned better than recall (or association) material.

Recommendations for Further Research

A number of areas for further research suggest themselves:

1. The same or similar research to the present study needs to be done with different ages, with males, and in different ecological settings to test the generalizability.

2. Much research needs to be done in defining personological variables which show promise and can be reliably measured.

3. The same research efforts need to be done in defining treatments which make a difference and, of course, the most effort needs to be expended to see what aptitudes fit best with what treatments.

4. A great deal more work needs to be done on retention. The question of how we increase overall retention is posed. If, as a number of studies suggest, students learn as much and retain more with less class time and less, or different, efforts by instructors, perhaps new broad experimental programs should be undertaken to see if more change in the desired directions can be made with greater economy than present educational programs permit.

LIST OF REFERENCES

LIST OF REFERENCES

Astin, A. W. The college environment. Washington, D. C.: American Council on Education, 1968.

Atkinson, R. C., and Paulson, J. A. An approach to the psychology of instruction. Psychological Bulletin, 1972, 78, 49-61.

Barton, K., Dielman, T. E., and Cattell, R. B. The prediction of school grades from personality and I. Q. measures. Personality, 1971, 2, 325-333.

Becker, W. C., Engelman, S., and Thomas, D. R. Teaching: A course in applied psychology. Chicago: Science Research Associates, 1971.

Bigelow, G. S., and Egbert, R. L. Personality factors and independent study. Journal of Educational Research, 1968, 62, 37-39.

Bracht, G. H. Experimental factors related to aptitude-treatment interactions. Review of Educational Research, 1970, 40, 627-645.

Bracht, G. H., and Glass, G. V. The external validity of experiments. American Educational Research Journal, 1968, 5, 437-474.

Cofer, C. N. The specificity of individual differences in learning. In R. M. Gagné (Ed.), Learning and Individual Differences. Columbus, Ohio: Merrill, 1967. Pp. 136-140.

Corley, J. R., Valente, R. G., and Shamow, N. K. The retention of material learned in a personalized introductory psychology course. Paper presented as part of a symposium (Personalized Instruction: A Symposium in Honor of Fred Keller) at the APA 1971 meeting in Washington, D. C.

Cronbach, L. J. How can instruction be adapted to individual differences? In R. M. Gagné (Ed.), Learning and Individual Differences. Columbus, Ohio: Merrill, 1967. Pp. 23-39.

Dayton, C. M. The design of educational experiments. New York: McGraw-Hill, 1970.

DiVesta, F. J. Theories and measures of individual differences in studies of trait by treatment interactions. Educational Psychologist, 1973, 10, 67-75.

Domino, G. Interactive effects of achievement orientation and teaching style on academic achievement. ACT Research Reports No. 39, December, 1970, Research Development Division of ACT Program.

Domino, G. Differential prediction of achievement in conforming and independent settings. Journal of Educational Psychology, 1968, 59, 256-260.

Entwistle, N. J. Personality and academic attainment. British Journal of Educational Psychology, 1972, 42, 137-151.

Feldman, K. A., and Newcomb, T. M. The impact of college on students. San Francisco: Jossey-Bass, 1969.

Gruber, H. E., and Weitman, M. Self-directed study: Experiments in higher education. University of Colorado, Behavior Research Laboratory Report No. 19, April, 1962.

Guilford, J. P. The structure of intellect. Psychological Bulletin, 1956, 52, 267-293.

Haskell, R. W. Effect of certain individual learner personality differences on instructional methods. AV Communication Review, 1971, 19, 287-297.

Haskell, R. W. Effect of personality characteristics upon learning via selected modes of instruction: An experimental investigation. Unpublished Docotral dissertation, Purdue University, 1970.

Jennings, W. N., IV. The assessment of instructor differences using the word-association examination. Unpublished Doctoral dissertation, The University of Tennessee, 1968.

Jensen, A. R. Varieties of individual differences in learning. In R. M. Gagné (Ed.), Learning and Individual Differences. Columbus, Ohio: Merrill, 1967. Pp. 117-135.

Keller, F. S. "Goodbye Teacher." Journal of Applied Behavior Analysis, 1968, 1, 79-89.

Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks/Cole, 1968.

Koran, M. L. Varying instructional methods to fit trainee characteristics. AV Communication Review, 1972, 20, 135-146.

Levin, J., Rohwer, W., and Cleary, T. Individual differences in the learning of verbally and pictorally presented paired associates. American Educational Research Journal, 1971, 8, 11-26.

Lindquist, E. F. Design and analysis of experiments. New York: Houghton-Mifflin, 1953.

McNally, L. The effect of instruction on learning. Paper presented at the AERA annual meeting in New Orleans, La., January, 1973.

Miller, D. M., and O'Conner, P. Achiever personality and academic success among disadvantaged college students. Journal of Social Issues, 1969, 25, 103-116.

Milton, O. Alternatives to the traditional. San Francisco: Jossey-Bass, 1972.

Mitchell, J. V. Education's challenge to psychology: The prediction of behavior from person-environment interactions. Review of Educational Research, 1969, 39, 695-721.

Moore, J. W., Hauck, W. E., and Gagné, E. D. Acquisition, retention, and transfer in an individualized college physics course. Journal of Educational Psychology, 1973, 64, 335-340.

Moore, J. W., Mahan, J. M., and Ritts, C. A. Continuous progress concept of instruction with university students. Psychological Reports, 1969, 25, 887-892.

Newson, R. S., Eischens, R., and Looft, W. R. Intrinsic individual differences: A basis for enhancing instructional programs. Journal of Educational Research, 1972, 65, 387-392, 396.

Nunney, D. N., and Hill, J. E. Personalized educational programs. Audiovisual Instruction, 1972, 17(2), 10-15.

Randhawa, B. S., and Fu, L. W. Assessment and effect of some classroom environment variables. Review of Educational Research, 1973, 43, 303-321.

Reynolds, M. C., and Balow, B. Categories and variables in special education. Exceptional Children, 1972, 38, 357-366.

Rosenshine, B., and Furst, N. The use of directed observation to study teaching. In R. M. W. Travers (Ed.), Second Handbook of Research on Teaching. Chicago: Rand-McNally and Co., 1973. Pp. 122-183.

Schultz, R. E. The control of "error" in experimental education. School Review, 1966, 74, 150-158.

Siegel, L. The contributions and implications of recent research related to improving teaching and learning. In O. Milton and E. J. Shoben, Jr. (Eds.), Learning and the Professors. Athens, Ohio: Ohio University Press, 1968. Pp. 136-157.

Siegel, L., and Siegel, L. C. The instructional gestalt in televised university courses. U. S. Department of Health, Education, and Welfare, Office of Education. Bureau of Research. No. 5-0852, Title VII Project No. 609, NDEA Act of 1958, Grant No. 7-45-0350-122, 1966.

Snow, R. E. Research on media and aptitudes. In G. Salomon and R. E. Snow (Eds.), Viewpoints, Bulletin of Indiana University School of Education, 1970, 46(5). Pp. 63-89.

Stern, G. G. People in context: Measuring person-environment congruence in education and industry. New York: Wiley and Sons, 1970.

Stroup, A. L. The prediction of academic performance from personality and aptitude variables. Journal of Experimental Education, 1970, 38(3), 83-86.

Stroup, A. L., and Eft, J. H. The CPI as a predictor of college academic success. Alberta Journal of Educational Research, 1969, 15, 191-194.

Travers, R. M. W. (Ed.) Second handbook of research on teaching. Chicago: Rand-McNally, 1973.

Wadsworth, B. J. Piaget's theory of cognitive development. New York: David McKay, 1971.

Walberg, H. J. A model for research on instruction. School Review, 1970, 78, 185-200.

Webb, C., and Baird, J. H. Learning differences resulting from teacher- and student-centered teaching methods. Journal of Higher Education, 1968, 39, 456-460.

Wessell, N. Y. The process of innovation. In C. B. T. Lee (Ed.), Improving College Teaching. Washington, D. C.: American Council on Education, 1967. Pp. 205-210.

Winer, B. J. Statistical principles in experimental design. Second edition. New York: McGraw-Hill, 1971.

APPENDIX

SAMPLE ITEMS, DIRECTIONS, AND FORMAT FOR

THE TWO TYPES OF TEST ITEMS

Association

DIRECTIONS: Give your associations to the following words or phrases; make them brief and to the point. A single word will usually be adequate to demonstrate your grasp of the material and that you could write more, given time. (Scoring 0, 1, or 2 for each response.)

Extinction		Assimilation		Punishment	
Extinction		Assimilation		Punishment	
Extinction		Assimilation		Punishment	
Extinction		Assimilation		Punishment	

Applied

DIRECTIONS: Respond to the following situations in a concise manner. Don't ramble; make your point(s) and move on!

A. Sarah is a typical "behavior problem" in your third grade class. She gets out of her seat too often, will not do her work for very long periods of time, bothers other children, talks at the wrong times, does not respond at the appropriate times, etc. What means would you use to get Sarah out of the category of "behavior problem"? (Three ways of dealing with her should be mentioned.)

B. Give an example of a concept to be taught at two levels which incorporates the invariance notion of concept acquisition cognitive development (i.e., fix a concept into the curriculum at two levels utilizing the appropriate levels of cognitive development).

VITA

Robert Allen Gregory was born in Knox County, Tennessee, on 13 October 1945. He attended Powell Elementary School and graduated from Powell High School in 1963. He graduated with honors from The University of Tennessee in Knoxville in June, 1967 with a Bachelor of Science degree in secondary education and, at the same time, was commissioned a Second Lieutenant in the United States Air Force. He continued in school as an NDEA Fellow to obtain a Master of Science degree in the Department of Educational Psychology and Guidance in 1968. He then went on active duty with the Air Force at Maxwell Air Force Base in Montgomery, Alabama, as an educational advisor in Air University's Academic Instructors School. Subsequently, he served as a test constructor for the Extention Course Institute and then held a position in Air University Institute for Professional Development's Professional Personnel Management Course as Chief, Behavioral Science in Management Division. During this time he was promoted to First Lieutenant and Captain and was granted a regular commission.

Under an Air Force program, he returned to The University of Tennessee in 1972 to be awarded a Doctor of Education degree with a major in Educational Psychology and Guidance in June of 1974. He is a member of Phi Kappa Phi, Phi Delta Kappa, and Kappa Delta Pi national honorary societies. He also holds membership in the American Psychological Association and the American Educational Research Association.

He is married to the former Ruth M. Way of Norris, Tennessee, and has a daughter, Rachel Oriana.