1-1-1973

A comparative study of different instructional programs and their effects on student affect, cognition, and personality.

John Floyd Newby
University of Massachusetts Amherst

A COMPARATIVE STUDY OF DIFFERENT INSTRUCTIONAL PROGRAMS AND THEIR EFFECTS ON STUDENT AFFECT, COGNITION, AND PERSONALITY

A Dissertation Presented

By

John F. Newby

Submitted to the Graduate School of the University of Massachusetts in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 1973

Psychology

A COMPARATIVE STUDY OF DIFFERENT INSTRUCTIONAL PROGRAMS AND THEIR EFFECTS ON STUDENT AFFECT, COGNITION, AND PERSONALITY

A Dissertation

by

JOHN F. NEWBY

Approved as to style and content by:

Ronald K. Hambleton

Dr. Ronald K. Hambleton (Chairman of Committee)

James M. Royer

Dr. James M. Royer (Member)

Harry Schimmer

Dr. Harry Schumer (Member)

H. Swaminathan

Dr. Hariharan Swaminathan (Member)

for R.T.L.

Dr. Richard T. Louttit (Department Head)

Psychology

December, 1973

TABLE OF CONTENTS

ABSTRACT

In view of the importance of comprehensive evaluations of school programs, this investigation, was designed to accomplish two goals. One was to compare the effects of five instructional programs which differed on some very basic dimensions. The second objective was to determine whether the instructional programs interacted with student aptitudes to produce higher performance on selected criteria. This phase of the study is widely known as aptitude-treatment interaction research.

The total sample consisted of 1366 students in five instructional programs and the sixth, seventh and eighth grades. All schools were located in fairly affluent communities and were well supported. Comparability among students was confirmed by comparing group means on the Lorge-Throndike Intelligence Test, the Otis-Lennon Mental Ability Test and the STEA, a short intelligence scale included in the SRA achievement battery.

The failure to carefully define instructional programs under study, has led to considerable confusion in the literature over the importance as well as the possible implications of research studies. The problem is that treatments are too often vaguely discussed in very global terms. The differences between the programs in this study were highlighted by positioning each instructional treatment on the "Ten Dimensions of Schooling." The major

distinctions were in the setting of instructional objectives, selection of materials and activities, arrangement of the physical environment, scheduling, and the individualization of instruction.

The major concern in the comparison studies was with overall differences in group scores. Students in the more conventional programs had higher scores on achievement measures. This finding, however, has to be assessed in light of the fact that the innovative programs were still in the developmental stages and were constantly being modified. The analysis of study habits and attitudes scores revealed significant differences between grades. Sixth graders had higher scores than seventh grade students, who in turn, had significantly better attitudes and study habits than eighth grade students. Even though the study was not longitudinal, this result is consistent with previous research demonstrating declining attitudes towards school and school related activities as students progress from lower to higher grades.

Scores on the Learning Environment Inventory (LEI) and on tests of fluency, flexibility and originality were analyzed by multivariate analyses of variance. The results of the analysis of the LEI scores showed significant main effects for both grade and treatment. To identify the source of the differences discriminant functions associated with the significant latent roots were computed. These analyses

substantiated expectations that the learning environment in the most traditional program was significantly different from that of the other instructional programs. One discriminant function appeared to be related to instruction. The other seemed to be concerned with intergroup/interpersonal accord. Students in the most conventional program perceived their assignments as being more difficult, their programs as having fewer provisions for individual differences and the pace of instruction as being too rapid. Students in this treatment also tended to be apathetic about school though they perceived their classes as being congruous and as having less tension and disagreement. The analyses of the fluency, flexibility, and originality scores resulted in fairly consistent findings. The most conventional treatments differed from the others at each grade level. In general, these differences occurred on measures of spontaneous flexibility, ideational fluency and originality.

The ATI studies were conducted with composite achievement and study habits and attitudes scores as the criteria and a number of carefully chosen personality variables as predictors. Shy, dependent, submissive individuals appeared to excel in highly structured environments where a great deal of direction was provided. Dominant, enterprising students seemed to achieve best in the more open programs.

ACKNOWLEDGEMENTS

I would like to express appreciation to many individuals who, in various ways, contributed to the successful completion of this dissertation. I am grateful to my doctoral committee, first, for the patience and practical assistance which they each gave so freely and unselfishly. In addition, I extend special thanks to Dr. James M. Royer and Dr. Harry Schumer for their very many penetrating questions and insightful observations. Likewise, I am grateful to Dr. Harihan Swaminathan for his very able guidance in conducting the statistical analyses reported in this dissertation.

I have always held the very deepest respect and admiration for the chairman of my doctoral committee, Dr. Ronald K. Hambleton. He has contributed, in many ways to my maturation, as a student and researcher. His advice and guidance, not only in the conducting of this dissertation, but throughout my years of study at the University of Massachusetts have been invaluable. I am eternally grateful.

I also extend thanks to Mrs. Delores Fremown for her expert typing, Mr. Richard Rovinelli and Dr. Daniel Sheehan, who also worked on the project from which this dissertation evolved. Without their co-operation, perhaps none of the work which came out of that project would have seen fruition.

LIST OF TABLES

LIST OF FIGURES

Chapter I

Introduction

1.1 Background

Over the last ten years we have seen a plethora of educational programs of all forms and descriptions designed for implementation in our schools (for examples, see Gibbons, 1970; Heather, 1972; Hambleton, 1973). Usually, such innovations are tried for several years at most, with few, if any real positive consequences and as a result they are gradually phased out. Obviously, there are exceptions to this trend. The most prominent examples include innovations based on the concepts of Individually Prescribed Instruction (Glaser, 1968), Mastery Learning (Bloom, 1968; Carroll, 1970), and Open Education (Plowden, 1966). The singularity of such programs can be ascribed to several features. Foremost among these is the fact that instruction is individualized. Students are encouraged to assume a more active role in the learning process, teachers are permitted to give more individual attention, and instruction is individually paced to allow for differences in learning rate.

A consideration of programs which were not successful is also revealing. While these are quite varied, their goals are also typically expressed in terms of individualization and the maximization of learning. The fact that a large

proportion of innovative programs have experienced little or no success has led many educational decision-makers to the conclusion that the impact of such interventions are insignificant and that that the effects of innovative instructional programs, in general, are trivial and, therefore, do not represent any significant improvement over traditional methods of instruction. Supporting this view are, perhaps two of the most controversial but important investigations completed in the last decade (Coleman, 1966; and Jencks, Smith, Acland, Bane, Cohen, Gintis, Heyns, and Michelson, 1972).

The reasons for the apparent paucity of successful examples are easily identified since many so-called innovations in education are plagued by obvious, but hopefully, resolvable problems. The following would be foremost among explanations for program failures: Poor planning, poor leadership, insufficiently prepared personnel, inadequate support, and lack of relevant feedback and evaluation.

Hence, most available evaluative studies cannot be cited as evidence for or against the effectiveness of innovations in education. Evaluators must, therefore, share the blame for the failures of many instructional programs. Typically, evaluators do not adequately appraise the programs and the problems confronting them prior to conducting a study. In addition, the information given program decision-makers often is not the kind needed for formative evaluation or the kind that contributes to the ultimate improvement of the programs. Clearly, a number of steps need to be taken if we are to come

to grips with the problems confronting our schools. For one, we must facilitate greater cooperation and communication between school personnel and evaluators. More comprehensive evaluations must be conducted covering not only cognitive but affective and environmental variables as well. The instruments used must be carefully selected or developed and they must be valid for the purposes they are to serve. It is essential that information be gathered about different instructional programs, and that comparisons among them be made in order that the relative advantages of each is identified. Much more information about the effects of programs across grades must be collected in case differential programmatic effects are related to developmental differences. Finally, the possibility that certain programs may be best suited for certain kinds of individuals should be investigated.

1.2 Purposes

In view of the limited scope and superficiality of most evaluations of instructional programs, this study was designed to collect data on a variety of variables for the purpose of comparing five different instructional approaches. It was conceived upon realization that there is a need for studies which investigate the effects of programs that are adequately planned, staffed, and supported to determine if differences really exist among them. One general objective, then, was to assess the differences among the instructional programs or treatments on certain carefully chosen variables.

Student responses on selected achievement, attitude, and creativity tests represented the major variates on which assessment was based. These, as well as various personality tests, were administered to a large number of students. The attitudinal variables investigated in this study were those held toward studying and school-related activities. Since many instructional programs aspire to facilitate and encourage creativity, it was felt that this variable was also relevant and should be explored. Scores on these tests were augmented by I. Q.'s and scores on the Iowa Tests of Basic Skills (ITBS) and the Science Research Associates Achievement Test (SRA). All tests were selected on the basis of theoretical relevance to the learning criterion and the instructional setting.

The second general objective of the study was to determine whether the instructional programs interacted with student aptitudes to produce higher performance on selected criteria. The issue was not which environment was best, but rather which environment was best for a particular person to produce a specific effect. This part of the study, which has been termed, aptitude-treatment interaction research, was highly exploratory in nature, conducted with the idea of collecting data that might be used to generate hypotheses for more controlled studies in the future.

The aptitude-treatment interaction (ATI) investigations involved a generation of regression lines of several criterion variables on selected aptitude variables (one criterion and

predictor at a time) for students within grades and instructional treatments. Where regression lines derived for a fixed criterion, predictor, and grade, across instructional treatments were disordinal, aptitude-treatment interactions occurred and tentative explanations were offered.

Since the studies reported here were exploratory in nature, the general emphasis was as much on hypothesis generation as on hypothesis testing. Thus the design did not exhibit carefully constructed treatments with all extraneous factors clearly accounted for. Of course this kind of research also has its place and it is felt that the contribution made in designating fertile areas for further study justifies the effort.

1.3 Educational Importance

There appears to be widespread dissatisfaction with schools as they are typically conceived and thus the movement to identify more effective instructional programs and techniques is gaining momentum. The establishment of diverse instructional programs is accompanied by the need for more studies like the one described here which document the effectiveness of such programs on a number of relevant dimensions. More informative evaluations must be conducted, if for no other reason than that there is widespread interest in those instructional programs which hold the most promise of achieving the less estoeric goals of contemporary schools. The number of schools which now include creative productivity,

positive affect and the psychological well-being of students among their goals is substantial. Most previous research on creativity, attitudes and personality have been directed toward classes taught in conventional schools. Very few investigations of these variables have been conducted in so-called innovative schools or programs. Even fewer investigations have attempted to compare programs on these dimensions. Since the present study was concerned with exactly these questions it is felt to be of special significance.

The results of the study also have implications for evaluation methodology. Specifically, the method of quantifying apparent differences in instructional treatments, the way in which confidentiality was handled, and the preparation of students and teachers for testing may well have wider applicability.

It was strongly suspected that certain learner characteristics would interact with instructional treatments to produce greater student performance. The identification of several aptitudes and instructional treatments which are likely to result in ATI's would greatly facilitate the development of individualized instructional programs. Since the data was collected on a number of variables, the research has the value of comprehensiveness. Subsequent studies conducted under more controlled conditions might well demonstrate significant interactions based on the work presented here. Certainly, the specification of relevant aptitudes and treatments eventually will enable educators to predict programs

in which a student is likely to achieve best as well as design treatments which meet the specific needs of students.

1.4 Outline of the Study

The remainder of the study was organized into five chapters. Chapter II presents a description of ten dimensions of schooling along with information on the position of each instructional treatment on the dimensions. Chapter III is divided into three parts. In the first two parts are presented a selected review of research on school achievement, attitudes and creativity. In the third part of Chapter III, a number of studies which fit into an aptitude-treatment interaction framework are discussed. Chapter IV outlines the methodology used in the study along with complete descriptions on all of the variables. The results along with a discussion are included in Chapter V. In the final chapter the overall findings of the study are summarized, implications are discussed, and further areas of research are suggested.

Chapter II

Description of Instructional Programs

2.1 Introduction

The literature is full of studies comparing "conventional" vs. "innovative" or "structured" vs. "unstructured" programs. However, seldom do the researchers present a clear definition of just what they mean by these terms. This failure to carefully describe instructional programs under study has led to considerable confusion in the literature over the importance of the results and the possible implications. Perhaps one reason that research in this area is so muddled is that investigators have, in fact, been contrasting programs with very similar characteristics but different labels.

What is needed are a set of dimensions that describe school programs along with appropriate guidelines for positioning school programs on these dimensions to lead us from the dilemma. The first really important contribution along these lines was the outstanding work of Traub, Weiss, Fisher, and Masella (1972). The Dimensions of Schooling (DISC) questionnaire was developed initially by Traub et al. (1972) for describing open education in terms of observable characteristics. The ten dimensions selected are presented in Table 2.1.1. They are thought to have wide applicability

Table 2.1.1

Dimensions of School Life Included in DISC (II)[1]

Setting Instructional Objectives	Materials and Activities	Physical Environment	Structure for Decision Making	Time Scheduling	Individualization of Instruction	Composition of Classes	Role of Teacher	Student Evaluation	Student Control
Defining Objectives	Availability of Curriculum	Learning Environment	Assignment of Teachers	Horizontal Organization	Student Methods	Range of Age	Selection of Materials	Criteria for Promotion	Rule Making
Use of Objectives	Materials Usage		Sub-grouping Criterion	Time Structure	Pace of Materials	Vertical Organization	Decision Making	Promotion Timing	
	Media Usage			Unstructured	Number of Student Groups		Teacher Focus	Evaluation Audience	
				Attendance			Teacher Role	Timing of Evaluation	
				Independent Study Time				Student Role in Evaluation	
								Evaluation Procedures	

[1]Taken from a paper by Traub, Weiss, Fisher, and Masella (1973)

and include concepts that are central to the development of most recent innovative programs. They were, therefore, felt to be most appropriate for describing the programs in this study.

In the remainder of this chapter we have used the ten dimensions developed by Traub and his colleagues to structure our comments on the five instructional treatments.

2.2 Setting Instructional Objectives

Traub et al. (1972) hold that the more open a school's program, the greater the opportunity students, as individuals, will have to participate in setting objectives and the greater the degree to which objectives will be individualized.

A review of the various programs showed that teachers, in all cases, defined the overall objectives for the courses. However, within several of the programs, enabling objectives were individually set. For example, the math programs for instructional treatments 1 and 2 were individualized. While secondary objectives were stated for each student, providing different paths to the goals, all students were expected to achieve the same overall objectives established for the program. In instructional treatment 3, both the language and the math programs were individualized and required individually set enabling objectives. The science program was based on a contractual system. Again, the major objectives were predetermined.

2.3 Materials and Activities

In the most open programs, any available book, film, record, toy, tool, or other object or collection, would be acceptable for use in the school. The number of activities considered appropriate for students would almost be unlimited. Clearly, the emphasis here would be on diversity. The only criteria for selection being that the activity or resource interest someone, is available, or can be undertaken in the school or in the surrounding community.

Instructional treatment 3 obviously ranked high on this dimension. Not only was extensive use made of the community, but there also was a wide variety of materials and activities available at the school. An example would be the extensive media program. Teachers also made use of micro-filming on occasion.

Instructional treatments 1, 2, 3, and 4 had resource centers, and extensive collections of books, games, etc. The availability of such resources might well be expected since all schools in this study were located in upper middle class neighborhoods and were well supported. The differences among them was that students in the first three were free to visit the resource center any time of day. Students in instructional treatment 4 were permitted to utilize the resource center periodically at specified times only.

The real contrast, however, was between instructional treatments 1, 2, 3, 4, and instructional treatment 5. The principal of the school housing instructional treatment 5 felt that his school enjoyed as much parental interest and

community support as the other schools in the study. However, he agreed that there were very few extra resources or activities available to his students. Space for a resource center had been designated. At the time this study was conducted, there were a few tape recorders and a small collection of books located there. It was obvious, though, that the center was still in the developmental stage and little or no use was being made of it.

2.4 Physical Environment

As far as open education is concerned, the preference would be for flexible environments that can be readily modified to suit the situational requirements of an activity. The architectural features of the school was not felt to be a particularly important factor in this study. However, the use or arrangement of available space within the school was seen to be of some importance.

In instructional treatment 5, very little effort was directed toward organizing the available space to meet the special needs of a particular group. All instruction was conducted in self-contained classes. On only one occasion were desks ever observed to be arranged in any order other than straight rows. The other groups were known to make use of space in very different ways.

The philosophy permeating instructional treatment 3 evolved around getting the student out into the community, involving him in community affairs. Instead of spending most

of his time in school, the student was engaged in projects conducted in nearby cities or he spent time at a local bank or city hall. The objective, obviously, was to broaden his range of experiences.

2.5 Structure of Decision Making

One of the areas in which the decision-making process can be evaluated is the assignment of students to teachers. As far as formal course work is concerned, none of the students in this study were permitted to choose the teachers who taught their courses. Traub et al. (1972) state that in a relatively open decision-making structure, the students would be allowed to group themselves according to interest and to move from one group or activity to another without seeking permission. This was not generally practiced by the teachers in instructional treatments 4 and 5. In most cases, the formation of class subgroups was done by the teacher on the basis of achievement and expediency.

In instructional treatments 1, 2, and 3, students did have a voice in the selection of outside activities in which they participated. However, as far as the formation of subgroups or the selection of classes in required courses is concerned, they had no more control over assignments than students in instructional treatments 4 and 5.

2.6 Time Scheduling

One way in which a program uses time structuring is denoted by the absence or presence of a rigidly followed

timetable. Another is the amount of unstructured time in the school day; that is, time during which students are completely free to pursue their own interests. A third is the amount of independent study time provided in each different subject. A fourth indication of the extent of fixed scheduling in a school is attendance requirements. If students are required to attend scheduled activities then the school is much less flexible than if students are free to opt out of scheduled activities if they so choose.

In instructional treatments 1, 2, and 3, seventh and eighth grade students spent only 60% of the school day in formal classes. The other 40% was open to the student to use in any manner he chose. The schools provided what might be called learning opportunities or mini-courses. The greatest degree of freedom was enjoyed by seventh and eighth grade students in instructional treatment 3, where the entire community became their laboratory.

Very little freedom of choice was given to students in instructional treatments 4 and 5.

2.7 Individualization of Instruction

Individualization, according to Traub et al. (1972) means allowing the student the freedom to work at his own pace and to learn in a way that he finds satisfying and rewarding. The availability of different subgroups might be an indication that individualizing is occurring.

Instructional treatments 1, 2, and 3 were considered to

have flexible approaches to learning. They encouraged a great deal of problem solving activities and independent study. In this way, students were permitted to explore their own interests and develop their creative talents. As already stated, students in the seventh and eighth grades spent approximately 60% of their time in formal instruction. The mini-courses provided to fill the remaining time were modules of instruction on a number of pertinent topics such as ecology, economics, sociology or special topics in mathematics--subjects not typically covered in conventional programs. In some instances, interdisciplinary teams were formed to provide a more comprehensive coverage of subjects which crossed disciplines. One such topic was pollution jointly conducted by the chemistry, science, and social studies teachers. The attempt being made, obviously, was to draw upon the creative talents and interests of the teachers. Other features of the program included flexible scheduling, individually paced instruction in math and science courses and an innovative reading program. The environment at these schools was designed to promote the social, intellectual, emotional, physical and aesthetic growth of the student at his own rate of development. The underlying assumption was that aided by subtle teacher guidance, students would accept the responsibility to synthesize their own learning experiences--some of which would be fleeting, others encompassing several weeks of selected activity.

In instructional treatment 4, basic concepts and subject matter content were still conveyed to students mainly through formal coursework in regular classes. At the seventh and eighth grade levels, however, a number of additional learning opportunities were offered which were designed to meet individual needs, as well as to provide students with the opportunity to pursue their own interests. Each department at the school had developed some unique approaches to accomplish the goals of the school.

In English, where the goal was to help each student develop a communicable oral and written style, students attended five periods of English a week. Rigid adherence to a single text was not required. However, the English book used had several different reading levels for varying abilities. The English, along with the social studies department also offered a seminar program which consisted of elective courses, presented once a week by members of the departments. Topics for seminars were selected by faculty on the basis of student interest.

The school also afforded the opportunity to participate in individual study projects. Students chose their own topics and were assigned advisors who had expertise in the area of interest. The addition of a resource center where audio-visual equipment, research texts and tools and other materials were available greatly increased the opportunities for doing independent study and participating in new programs like student tutorials. Periods of unassigned time were used for

these purposes. The atmosphere in the resource center was informal and relaxed. Students were permitted to conduct their own studies and investigate questions that interested them.

The science program was 90% lab oriented where the students learned by actually conducting research. Key concepts were introduced through experiments. The program stressed the application rather than the memorization of knowledge. Experiments were presented in a structured sequence which could be condensed or synthesized. Mini-course electives were offered by members of the department during non-teaching hours. These were designed for students who were able and who wanted to do more work.

The math program at this school utilized computation and drills to perfect the basis skills. Achievement of mathematical concepts was aided by manipulative materials like math games which were available in the Resource Center. Other unique features of the program included a math team and computer club.

2.8 Composition of Classes

This dimension refers to the manner in which students are moved through the program. The central difference is between programs which group students according to age and those that allow students to group themselves without regard for age or past accomplishments.

None of the treatments in this study had eliminated the

grade distinction. Furthermore, there was very little peer teaching or interaction between students in different grades.

2.9 Role of the Teacher

The major distinction along this dimension is between the teacher as lecturer or purveyor of information and the teacher as facilitator of learning acting primarily as a resource person or counselor. Teachers, at all of the schools lectured to some extent at least. However, the most obviously didactic and authoritarian approach was observed in instructional treatment 5.

Instructional treatment 5 followed what was felt to be a very traditional approach to instruction . . .being characterized by a self-contained classroom with a single teacher instructing anywhere from 25 to 35 pupils at one time. Typically, instruction in such systems are structured and expected to be well-organized. That is, material is presented sequentially in such a way that when new stimuli are introduced, the learner is able to recognize familiar elements and attack each problem on the basis of prior learning of fundamental skills, facts or principles. Teachers characteristically assume primary responsibility for specifying what is to be presented in classes and how much time is to be devoted to specific activities. Subject content is imparted to students almost exclusively through readings and lectures. Classes thus acquire a degree of regularity and predictability. Students know what is expected of them and many are able to discipline themselves accordingly and acquire the necessary

skills and behaviors which allow them to excel.

2.10 Student Evaluation

A program is considered most open when it:

a. employs observations, work samples and anecdotal reports but no formal tests;

b. adopts evaluation procedures to suit individual students;

c. collects evaluation data more or less continuously and involves students directly in planning and implementing evaluation procedures;

d. makes decisions whenever there is apparent need for changing the activities a student engages in or the materials he works with.

Students in instructional treatments 1, 2, and 3, were not evaluated in the usual sense since grades were not assigned. The procedure was to send parents periodic reports indicating the level of a student's performance on specific objectives. What evaluation there was in these schools was made in terms of mastery and non-mastery or satisfactory and unsatisfactory. Thus the evaluation procedures were being adopted to individual needs.

Standard methods of evaluating students were practiced in instructional treatments 4 and 5.

2.11 Student Control

The essential difference on this dimension is between a school which sets few constraints on students and one which

exercises a custodial philosophy of student control; a philosophy characterized by "stress on maintenance of order," impersonality and even a punishment-centered orientation toward students.

The extent to which students are involved in the setting of rules would also be of concern here.

While all of the schools fell in the middle ranges, they can be ranked on this dimension. In general, instructional treatment 3 exercised the least amount of control while instructional treatments 4 and 5 exhibited the kind of environment most associated with a traditional school.

Differences within treatments were also observed. For example, more constraints were imposed on the sixth graders in instructional treatment 3 than on students in the seventh and eighth grade. Teachers maintained greater control over the sixth graders and took more responsibility for rule setting.

A team teaching approach was used at the sixth grade level of instructional treatment 4. Each classroom was L-shaped and had a partition which could be left enclosed when it was desirable to combine the two groups of students. This was often done as the attempt was made to allow each teacher the opportunity to instruct in the areas of his greatest strengths. Each group consisted of about thirty students. Two teachers shared responsibilities for instruction and planning. They also collaborated in determining the composition of the groupings that were formed. Here again, greater control was being exercised over sixth graders than over seventh and eighth grade students in the same treatment.

2.12 Summary

In summary, treatments 1 and 2 had very flexible approaches. They encouraged a great deal of problem solving activities and independent study. Instructional treatment 3 differed from 1 and 2 in one major respect. It utilized the community at large in fostering student learning. A healthy respect for structure and traditional values influenced the philosophy of education and the organization of learning experiences available for students in Group 4. The leadership, however, was considered to be quite progressive having taken significant steps toward individualization. After a careful analysis of the five programs, it was concluded that instructional treatment 5 was our best example of what has typically been called, in studies of this type, "the traditional program."

A graphical representation of the relationships among the five instructional treatments on Traub's ten dimensions of schooling is shown in Table 2.12.1. The major differences between groups were on dimensions, 1. setting instructional objectives, 2. materials and activities, 3. physical environment, 5. time scheduling, and 6. individualization of instruction. In all cases treatment 5 was ranked lowest and treatment 1 was highest. The least amount of difference among the programs was thought to be on dimensions, 7. composition of classes, 8. role of the teacher, and 9. student evaluation. None of the programs had eliminated the grade distinction, most teachers in all schools lectured and

Table 2.12.1

Description of the Five Instructional Treatments on the Ten Dimensions of Schooling[1]

Dimension Name	Dimension Description		Treatments				
			1	2	3	4	5
Setting Instructional Objectives	Set for Individuals	Set for Class	3	3	2	4	5
Materials and Activities	Diverse	Limited	3	3	2	4	5
Physical Environment	Flexible	Restrictive	4	4	2	5	5
Structure for Decision Making	Decentralized	Centralized	4	4	3	5	5
Time Scheduling	Unstructured	Structured	3	3	2	4	5
Individualization of Instruction	Individually Determined	Group Paced Instruction	3	3	2	3	5
Composition of Classes	Ungraded	Graded	5	5	5	5	5
Role of Teacher	Facilitator	Lecturer	4	4	4	5	5
Student Evaluation	Suited to Individuals	Based on Formal Tests	4	4	4	5	5
Student Control	Decentralized	Centralized	3	3	2	4	4

[1]To rate the instructional treatments on each of the ten dimensions a five-point rating scale was used. A "1" indicated a judgment that the instructional treatment was best represented by the adjective describing the left end of the dimension. A "5" indicated a judgment that the instructional treatment was best represented by the adjective describing the right end of the dimension. Intermediate values "2", "3" and "4" were used to describe a program which fell somewhere between the two extreme values.

tests of one form or another were the primary means of assessing student achievement.

Chapter III

Review of the Literature

3.1 Introduction

By way of background for our investigation it was felt that three areas needed special attention. The first involved an assessment of the current status of our knowledge of the effects of various instructional methods on student achievement and attitudes. The second area of concern was the problem of measuring creativity. Finally, it seemed essential to review the aptitude-treatment interaction literature to generate possible bases for stating hypotheses in the study.

3.2 Effect of Instructional Methods on Student Achievement Attitudes

The empirical study of selected learner characteristics and their relation to academic success, for the most part, has dealth with cognitive and personality variables with increased concern in attitudes or interests. Quite recently there has also been an interest in projective techniques. Khan (1969) noted that the bulk of previous research on prediction has been concerned with achievement as a function of cognitive variables, such as verbal and numerical aptitude, general mental ability, etc. He concluded that the study of cognitive factors as determinants of achievement have "been

thoroughly exploited" and that the average relationship between aptitude variables and achievement criteria ranges between .50 and .75. Consequently, one-half to three-quarters of the variability in achievement thus remains unexplained.

The results of numerous studies comparing educational programs in terms of cognitive development can be best summarized as equivocal. Some of the problems characterizing this research is evident from a review of studies on team teaching.

Team teaching as an approach to instruction has shown rapid increase over a very few years (Borg and Brite, 1967). While reports describing and assessing various team-teaching projects throughout the country are plentiful, the number of evaluations employing control groups and providing adequate statistical treatment of data is quite small. Typically, such studies have compared team with traditional teaching at a single grade level. The investigations of Knox (1966) at Grade 1, Burningham (1968) at Grade 4, and Georgiades and Bjelke (1966) at Grade 9 are illustrative. Composition and functioning of instructional teams differ from study to study and there is little consistency in terms of dependent variables measured. In spite of these differences the majority of controlled evaluations have shown several common findings:

1. Achievement of students under team instruction is no better than achievement in traditional

self-contained classes;

2. Student attitude is the same under team and regular instruction; and
3. Teacher attitude is more positive under team instruction.

Attempts to predict individual differences in intellectual achievement by means of non-intellective variables have also produced inconsistent results. Intuitively, one would expect that favorable attitudes towards school should be positively related to school achievement. This however, has not been substantiated in studies conducted by Jackson and Lahaderne (1967), Jackson and Getzels (1959), and Diedrich (1966) using global measures of attitude.

Other research indicates that school satisfaction can be more fruitfully studied if it is considered as a multidimensional variable. Cullen and Katzenmeyer (1970) defined such factors as teachers, peers, subject matter difficulty and subject matter interest in a factor analytic study which did result in a relationship between attitudes and achievement. Support for this stance is also found in reports by Auria and Frankiewicz (1967) with the Student Opinion Poll and Khan (1969) and Khan and Roberts (1969) with the Survey of Study Habits and Attitudes (SSHA).

Khan (1969) administered the SSHA to students in the ninth grade. The purposes of the study were to (a) obtain relatively pure measures of affective variables suitable for use with junior high school students, (b) determine how well

3.3 Problems in the Study of Creativity

In view of the recent critiques of tests of creativity (Crockenberg, 1972) the major concern in planning this study was whether such tests could, defensibly, be used in an evaluation study. There is considerable disagreement in the educational and psychological literature on whether the creativity tests actually measure what they are supposed to. Creativity, as a construct, has defied definition. Since there is no concensus on the meaning of creativity, it is difficult to identify adequate criteria against which to validate the tests. As an example of the problems encountered one might consider the Torrance Tests.

Torrance offered as evidence for the construct validity of his tests, correlations between scores of high school seniors and follow-up information on "creative activities" the students had engaged in some six to seven years later (Erikson, 1966). On the basis of returns from 44 of the 66 subjects tested, the following rather disappointing correlations were derived: fluency, .27 ($p < .05$); flexibility, .24 ($p < .10$); originality, .17 ($p > .10$); elaboration, .16 ($p > .10$). Crockenberg (1972) noted that this should not be surprising given the criteria used. The activities labeled "creative" included such items as: Subscribed to professional magazines or journals, changed religious affiliation, elected or appointed to a student office, and learned a new language. Crockenberg further noted that all of these criteria could be considered typical of well-educated high I. Q. people.

they predict criteria of achievement both separately and in conjunction with aptitude variables and (c) to assess the contribution of the affective variables to the accuracy of prediction over that which is realized by intellective variables alone. The intellective predictors were scores on the verbal and mathematical parts of the SCAT. The achievement criteria were scores on six subtests of the Metropolital Achievement Test Series.

The affective predictors significantly increased ($p < .01$) the multiple correlations for males when they were used in conjunction with aptitude scores on the SCAT. For females, five multiple correlations attained significance when affective predictors were added to the aptitude battery. For males, one significant canonical correlation of .687 ($p < .01$) was obtained, with attitude toward teachers and achievement anxiety contributing significantly to the relationship from the predictor set and reading and social studies contributing significantly from the criterion set. One significant canonical correlation of .758 ($p < .01$) was observed for females, with Achievement Anxiety as the only heavily weighted variable from the predictor set, and reading and social studies constituting the composite criterion.

In short, it would appear that attitudes and achievement are related and that attitudes can best be studied using an instrument such as the SSHA which defines attitudes in terms of several dimensions.

Additional data on construct validity is equally unconvincing. In general, the procedure for determining construct validity takes one of two forms. One way is to ask teachers to nominate children highest and lowest on the four dimensions measured by the test (Torrance, 1962, 1963). The other way compares children nominated by teachers as exhibiting the most of each characteristic with all those not nominated. Studies by Yamamoto (1962) and Torrance and Myers (1962) showed that the test scores of those nominated as highest by their teachers were significantly higher than test scores of the comparison group on fluency, flexibility and originality, but not on elaboration. Williams (1965), however, found that pupils nominated as most original by their sixth grade teachers did not differ from those not nominated in mean originality scores on the Ask-and-Guess task. Only one teacher out of six successfully differentiated between high and low scoring children. The evidence is not entirely consistent as well as being subject to competing explanations. Similar conclusions have been drawn by other investigators, not only about Torrance's tests, but with regards to other available measures of creativity as well (Cronbach, 1968; Davis and Belcher, 1971).

It might be observed that so-called measures of creativity seem to have derived their respectability on the basis of face validity given the evidence presented thus far. The principles upon which the tests were based seem to be reasonable. For example, it would be logical to

expect creative people to be flexible thinkers who would readily desert old ways of thinking to strike out in new directions. Therefore, a factor of flexibility is often included among creativity tests. One would also expect to find a trait of originality defined as cleverness, infrequency and remoteness of response. People suspected of being creative are also thought to be able to produce ready alternatives when presented with a problem. Therefore a factor of fluency is considered in the development of most tests of creativity.

Over the years a large body of literature has developed in the area of creativity assessment. J. P. Guilford factor analyzed a large number of cognitive tests and developed the Structure of Intellect Model. The category, Divergent Production, includes most of the traits we now identify with the domain of creativity. Based on this, we might conclude that Guilford has, at least, presented evidence for the factorial validity of his tests.

Studies of the relationship between creativity and intelligence should also be considered here. The results of several major investigations have indicated that measures currently being used to assess creativity are strikingly independent of the conventional realm of general intelligence. This finding has been documented by Torrance (1962), Getzels and Jackson (1962), Clark, Veldman, and Thorpe (1955); Guilford and Hoepfner (1966), and Wallach and Kogan (1965).

Wallach and Kogan (1965) conducted the most defensible and comprehensive investigation of the relation between creativity and I. Q. In comparison to previous work, their study was the strongest, methodologically, as well as the best controlled. Ten creativity and ten standard intelligence measures were administered to the students. The average correlation between intelligence and creativity measures was .10. This was much lower than the average intercorrelations between creativity measures (.40) as well as the correlation between all intelligence measures (.50). On the basis of their findings, Wallach and Kogan concluded that a dimension of individual differences had been defined, which on the one hand, possessed generality and pervasiveness, but which on the other hand, nevertheless was quite independent of the traditional notion of general intelligence.

Getzels and Jackson (1962) found that teachers, when asked to rate students on the degree to which they would like to have them in class, clearly preferred the high-I.Q. over the highly creative pupil. This was true even though the high-I.Q. students and the highly creative students were equally high in school achievement. The study also showed that the high-I.Q. child tends to hold a self-image consistent with what he feels the teacher would approve, seeking to conform to the projected values of the teacher; the creative pupil on the other hand, tends to hold to a self-image consistent with his own projected values, often not conforming to the teacher's values. He considers high marks and

goals that projectively lead to adult success in life less important than does a member of the high-I.Q. group. He has a much greater interest in unconventional careers than his less creative peers. Getzels and Jackson (1962) and Torrance (1962) have shown that the ability to think creatively affects the acquisition of information and educational skills.

Let us look more carefully at the methods by which validity of creativity tests is assessed. Guilford (1971) noted that ratings were especially suspect as criteria, unless they are obtained from experienced observers who make observations under controlled conditions. Under normal classroom conditions, teachers have been found to be poor evaluators of creative qualities (Merrifield, Gardner and Cox, 1964). Their ratings of such traits are likely to correlate strongly with I.Q.'s of students. Guilford further holds that creative talent is not a single, broad ability, but that it draws upon a large number of the abilities associated with intelligence on different occasions, and more uniquely upon abilities associated with the divergent-thinking production and transformation categories. Since creative talent, from the standpoint of aptitudes, is composed of numerous special abilities, and since criteria of creative performance in everyday life are also complex, no one test of a creative ability can be expected to correlate highly with those criteria. Multiple predictions and multivariate procedures of validation seem called for.

There seems to be agreement that so called "tests of

creativity" measure something. The controversy pertains to the conceptual definition of this "something." One suggestion to the problem is that the tests should be referenced more exactly, that is, they should be called measures of "fluency" or "flexibility," etc., rather than measures of creativity. This procedure has the advantage of being more descriptive as well as obviating the need for making what may turn out to be incorrect assumptions.

3.4 A Survey of Selected Aptitude-Treatment Interaction Research Studies

A number of studies contrasting what were thought to be different instructional approaches have been conducted within the framework of studies of interaction among instructional treatments and aptitudes [known as aptitude-treatment interaction (ATI) research]. A major portion of Cronbach and Snow's (1969) critique of the literature in this area was concerned with methodological issues in constructing ATI studies and analyzing the results. The authors conclude that most previous studies are inconclusive due to the way the problems were posed, the methods by which the data were analyzed, and their contradictory results.

A major part of the problem with previous studies is the fact that they have weakly conceptualized both the aptitude and the treatment dimensions. Cronbach and Snow hypothesize, for example, that to simply characterize aptitudes in such terms as "spatial" is unlikely to identify combinations of variables worth investigating. They also

assert that treatments used in the past have suffered from brevity and artificiality.

The above criticisms were thought to be obviated in the present study. Here the differences between on-going treatments were documented and it was felt that the instructional treatments were divergent enough to produce significant interactions. In addition, the aptitude variables were carefully chosen on the basis of their relevance within the context studied.

Numerous studies using personality measures and what has been variously called "structured" and "unstructured" or "teacher-centered" and "student-centered" approaches as variables have been conducted. While most of these studies were not defined as attempts to establish aptitude treatment interactions, they can be considered within this framework. Grimes and Allinsmith (1961), for example, investigated the relationship among compulsivity, anxiety, and performance in structured and unstructured settings. They found that anxiety and compulsivity interacted with one another and with teaching method. High anxious students did poorly in unstructured treatments while they did relatively better in structured treatments. Results also indicated that students who are both high compulsive and high anxious perform better in a structured treatment.

Smith et al. (1956) reported that anxious individuals of permeable (flexible) structure made optimum gains in reading efficiency when exposed to a maximum of course

structure and direction, and made minimum progress when exposed to non-directive teaching procedures. The explanation of this finding was that anxious, permeable individuals gain security from their dependence upon structure provided by an authority figure. A second hypothesis that anxious individuals of impermeable, rigid structure will make optimum progress when exposed to non-directive teaching procedures was not substantiated.

Dowaliby and Schumer (1973) recently reported evidence of an ATI where anxiety was the aptitude and two methods of teaching were considered as treatments. Subjects for the study were students in two college classes. While a teacher-centered mode optimized learning for high-anxious students, a student-centered approach resulted in superior performance for low-anxious students. The results of the above studies suggest that anxiety is a variable that should be considered in ATI investigations.

Several other personality variables have been found to interact with structure to produce differential student performance. These include independence, sociability, and authoritarianism. Amidon and Flanders (1961) found that while independent children were unaffected by teaching method, dependent children performed better under indirect than direct teaching. Beach (1960) reported that students low on sociability (as measured in the Social Intraversion - Extraversion scale of the Guilford Inventory of Factors) perform better in lecture sessions whereas more sociable

students perform better in leaderless discussion groups.

Domino (1971), in a study that confirmed previous findings (Domino, 1968), reported an interaction between the personality traits of Achievement-via-Conformance and Achievement-via-Independence and the treatment of teaching in a conforming manner (lectures, high structure) vs. teaching in an independent manner (active student participation, less structure). The author found a significant interaction between type of achievement orientation and style of teaching on five of seven variables. Satisfaction and achievement were maximized when the teaching style and the achievement orientation were compatible.

Finally, Doty (1967) included personality and creativity measures in a study of student characteristics and achievement in two structured methods (conventional lecture and tape lecture) and an unstructured method (small group discussion). Correlations between personality variables and achievement were computed. One analysis revealed that the relationship between social needs and achievement in conventional lecture and small group discussion were .40 and .65 respectively. The reported correlation was -.53 when instruction was by audio-taped methods. This study is particularly relevant since a criterion other than achievement was analyzed. When the criterion was attitude toward the teaching method the correlations between social needs and the criterion in the lecture and small group treatments were .51 and .39; for the audio-taped method the correlation was -.12. Doty also

computed the correlations between creativity and achievement in conventional lectures and audio-taped lectures. These correlations were low and negative, -.21 and -.16. For the small group instructional treatment the correlation between these two variables was .37.

The above studies dealt with structure or the lack of structure as manifested by individual teachers. It might be questioned then how well these results generalize to a whole school where the program is, in a broad sense, "structured" or "unstructured." It might be argued that in the latter case, the teachers will have natural proclivities toward teacher-centeredness and student-centeredness and that this will supercede whatever influence there might be from the formal structure of the school. While the present study was designed to determine whether the organization, structure and climate of a school affected student outcomes, there is some reason to believe this to be the case. For example, Patton (1955) found that students who rejected traditional sources of authority and are highly motivated toward personal achievement were most favorably disposed toward experimental classes run by students themselves and most able to handle the responsibilities involved in such classes. The same author also found that authoritarians, a distinct minority at the University of Chicago, tended to have the highest dropout rate. They complained of looseness in the pedagogical approach that tolerated smoking in classrooms, did not require attendance and expected students to answer their own questions (see Stern, 1962, p. 694).

Chapter IV

Method

4.1 Subjects

The total sample consisted of 1367 students in five instructional treatments and the sixth, seventh and eighth grades. All schools were located in fairly affluent communities and were well supported--both financially and otherwise.

Comparability among students was difficult to assess since there were no common aptitude or intelligence test data available. Unfortunately the design of the study did not permit any pretesting. Hence strict comparison of the students was not possible. The schools, as part of their regular testing program, routinely administered intelligence tests for their own purposes. Scores on the STEA, a short intelligence scale included in the SRA achievement battery, were available for students in instructional treatments 1, 2, and 3. The Otis-Lennon Mental Ability Test was administered to sixth and seventh grade students in instructional treatment 4 and to all students in instructional treatment 5. Deviation I.Q.'s on the Lorge-Thorndike Intelligence Test were available for eighth grade students in instructional treatment 4.

Comparison of the means for each of the groups with the relevant norms for the respective tests clearly showed that all groups scored at about the same place in the distributions (i.e., around the 60th percentile). Thus, the limited data on the background ability of students across instructional treatments does support the hypothesis concerning the comparability of the students. While not conclusive data it is supportive of the hypothesis that students in the different instructional programs were similar in general ability.

4.2 Description of Variables

A description of the various measures administered to students in the investigation is presented in this section. The measures are organized under four subheadings: (a) Cognitive Measures, (b) Personality Measures, (c) Learning Environment and Attitude Measures, and (d) Creativity Measures.

(a) Cognitive Measures

Iowa Tests of Basic Skills

The Iowa Tests of Basic Skills (ITBS) was devised to test functional skills of children in grades 3 to 9 in the areas of vocabulary, reading, comprehension, language skills, work-study skills, and arithmetic. The focus of the tests are not considered to be that of typical achievement tests which are concerned with the common content areas. Rather the tests focus on the evaluation of generalized intellectual

skills and abilities involved in and required for achievement in the aforementioned subject areas. The test makers hold that the measurement of these basic intellectual skills is far more valuable for use in the improvement and individualization of instruction and educational guidance than is the assessment of the acquisition of specific information in school subjects. While this point of view is debatable, it does provide some insight in the logic underlying the development of the tests.

The battery consists of eleven separate tests for grades three through nine. All tests are contained in a single booklet. However, each student takes only items appropriate in content and difficulty to his own grade level.

Test-retest reliability coefficients for the test tend to be high. They range from .84 to .96 for the major tests and from .70 to .93 for the subtests. These correlations are felt to be sufficiently high for individual diagnosis and prediction.

A major strength of the test is its curricular validation. A careful analysis of the skill processes being tested was conducted before test items were devised. Besides this, the usual procedures for establishing validity were followed. Sample test items were administered to a number of different samples and discrimination and difficulty indexes established.

SRA Achievement Series

Science Research Associates' Achievement Series measure

the educational development of pupils in grades four through nine in the following broad curricular areas: Reading, language arts, mathematics, social studies, science and use of sources. The test profiles also provide a composite achievement score. Primary interest in the present investigation was with the composite, reading, language arts, and mathematics scores.

Kuder-Richardson (KR-20) reliabilities for the various tests range from the low .80's to the low .90's. These coefficients are indicative of generally high level structural quality and an acceptable level of consistency in test performance.

The product-moment intercorrelations among the various subtests generally run in the 0.50's and 0.60's. This seems to indicate that, while the separate tests are measuring several areas in common, each score is providing some unique information regarding educational achievement.

While no standardized achievement test could measure all the objectives of an instructional program the broad objectives of the SRA series and the Iowa Tests of Basic Skills probably provide the best basis for comparing students in different instructional programs. Both tests place emphasis on measuring broad understanding and general skills which reveal the ability of the student to apply what he has learned rather than to recall facts in isolation.

(b) Personality Measures

California Psychological Inventory

A massive amount of data attests to the usefulness of the California Psychological Inventory (CPI). In general, the consistency of measurement is high enough to permit use of the scales in both group and individual testing (Gough, 1957). The reliability of the scales have also been investigated using the test-retest method. In one study, two high school junior classes took the CPI in the fall and again a year later. The modest test-retest correlations among the high school students (a range of .49 to .77 across the 18 scales) may reflect, in part, the differing rates of maturation among those adolescents during the year between testings.

Studies of the validity of the various scales are reported in the manual. The results of these studies are reassuring. They indicate that the CPI has wide applicability.

The scales are concerned with characteristics of personality which are related to the favorable and positive aspects of human as opposed to the pathological. Thus, its scales are addressed principally to personality characteristics that are important for social living and social interaction.

The CPI consists of eighteen scales. The items are stated in the true-false format with nearly all items scored on more than one scale. Each scale is intended to cover one important personality trait, and the total set of eighteen is intended to provide a comprehensive survey of an individual from a social interaction perspective.

Since this investigation was to include persons from the lowest age group for whom the test was appropriate, it was advantageous to delete about 200 items which were thought to be objectionable to students, uninteresting or unimportant. Great care was exercised in removing items from the inventory in order that the scales would not be invalidated in the process.

Of the eighteen scales, eleven were selected for use in this study. A description of each is presented below:

1. Dominance - This scale assesses factors of leadership ability, dominance, persistence, and social initiative. High scorers tend to be seen as aggressive, confident, persistent, and planful; and as being persuasive and verbally fluent. Low scorers tend to be seen as retiring, inhibited, common-place, indifferent, silent, and unassuming; and as being slow in thought and action.
2. Capacity for Status - This scale serves as an index of an individual's capacity for status (not his actual or achieved status). The scale attempts to measure the personal qualities and attributes which underlie and lead to status. High scorers tend to be seen as ambitious, active, forceful, insightful, resourceful, and versatile. Low scorers tend to be seen as apathetic, shy, conventional, dull, mild,

simple, and slow.

3. Sociability - This scale identifies persons of outgoing, sociable, participative, temperament. High scorers tend to be seen as being outgoing, competitive, and forward. Low scorers tend to be seen as awkward, conventional, quiet, submissive, and unassuming.

4. Social Presence - This scale measures factors such as poise, spontaneity, and self-confidence in personal and social interaction. High scorers tend to be seen as clever, enthusiastic, imaginative, quick, informal, spontaneous, and talkative. Low scorers tend to be seen as deliberate, moderate, patient, self-restrained, and simple.

5. Self-Acceptance - This scale assesses factors such as sense of personal worth, self-acceptance, and capacity for independent thinking and action. High scorers tend to be seen as intelligent, outspoken, sharp-witted, demanding, aggressive, and talkative. Low scorers tend to be seen as methodical, conservative, dependable, conventional, easy-going, and quiet.

6. Sense of Well-Being - This scale identifies persons who minimize their worries and complaints, and who are relatively free from self-doubt and disillusionment. High scorers tend to be seen as energetic, enterprising, alert, ambitious, and versatile.

Low scorers tend to be seen as unambitious, leisurely, awkward, cautious, apathetic, and conventional.

7. Achievement Via Conformance - This scale identifies those factors of interest and motivation which facilitate achievement in any setting where conformance is a positive behavior. High scorers tend to be seen as capable, co-operative, efficient, organized, responsible, stable, and sincere. Low scorers tend to be seen as coarse, stubborn, aloof, awkward, insecure, and opinionated.
8. Achievement Via Independence - This scale identifies those factors of interest and motivation which facilitate achievement in any setting where autonomy and independence are positive behaviors. High scorers tend to be seen as mature, forceful, strong, dominant, demanding, and foresighted. Low scorers tend to be seen as inhibited, anxious, cautious, dissatisfied, dull, and wary.
9. Intellectual Efficiency - This scale indicates the degree of personal and intellectual efficiency which the individual has attained. High scorers tend to be seen as efficient, clear-thinking, capable, and intelligent. Low scorers tend to be seen as confused, defensive, shallow, and unambitious.
10. Psychological-Mindedness - This scale measures the degree to which the individual is interested in,

and responsive to, the inner needs, motives, and experiences of others. High scorers tend to be seen as observant, perceptive, talkative, resourceful, and changeable. Low scorers tend to be seen as apathetic, peaceable, serious, cautious, and unassuming.

11. Flexibility - This scale indicates the degree of flexibility and adaptability of a person's thinking and social behavior. High scorers tend to be seen as insightful, informal, adventurous, confident, humorous, rebellious, idealistic, assertive, and egotistic. Low scorers tend to be seen as deliberate, worrying, industrious, cautious, guarded, mannerly, rigid, and methodical.

Children's Manifest Anxiety Scale

The Children's Manifest Anxiety Scale (CMAS) was developed by Castenada, McCandless, and Palermo (1956). It is an adaptation of the Taylor Manifest Anxiety Scale appropriate for use with elementary school children. The scale consists of 42 anxiety items and 11 items which provide an index of the subject's tendency to falsify his responses. The anxiety items can be grouped into roughly the same five categories as those in the Manifest Anxiety Scale. Those categories are (1) physiological disorders, (2) general emotionality, (3) the direct admission of worry or nervousness, (4) physiological stress, and (5) self-

consciousness and self-confidence. The CMAS is regarded as a measure of generalized anxiety. It is one of the most popular measures of general anxiety in children.

One-week test-retest reliabilities are about .90 for the anxiety scale and about .70 for the lie scale. Sex differences are typically found on both scales. Girls score significantly higher than boys. Grade differences are frequently observed.

School Anxiety Scale

The School Anxiety Scale developed by Phillips (1966), makes use of items from the Test Anxiety Scale for Children, the Achievement Anxiety Scale, the Audience Anxiety Scale, and other personality instruments. It was designed to assess anxiety associated with a broader range of stressful school situations than is encompassed by the Sarason scales (e.g., the Test Anxiety Scale for Children). A factor analysis revealed four factors which roughly parallel those found for the Test Anxiety Scale for Children: (a) fear of taking tests; (b) physiological reactivity associated with a low tolerance for stress; (c) lack of confidence in meeting the expectations of others, particularly teachers; and (d) fear of negative evaluation by others, particularly in public performances. Fewer items in the School Anxiety Scale load on the "test anxiety factor" than is the case for the Test Anxiety Scale for children. Phillips reported that the School Anxiety Scale correlates positively with the Prone-

ness toward Neuroticism subscale of the Children's Personality Questionnaire.

The Intellectual Achievement Responsibility Scale

Crandall, Katkovsky, and Crandall (1964) provide the rationale for the Intellectual Achievement Responsibility Scale. They state that individuals have been found to differ in the degree to which they believe that their actions produce the reinforcements which follow their efforts, or they feel that the rewards and punishments meted out to them are at the discretion of powerful others or are in the hands of luck or fate.

The Intellectual Achievement Responsibility Scale (IRA) attempts to measure beliefs in internal versus external reinforcement responsibility. It is aimed at assessing children's beliefs in intellectual-academic achievement situations and limits the source of external control to those persons who most often come in face-to-face contact with a child, his parents, teachers, and peers (Crandall, Katkovsky, and Crandall, 1964). The IAR scale is composed of 34 forced-choice items. Each item stem describes either a positive or a negative achievement experience which routinely occurs in children's daily lives. This stem is followed by one alternative stating that the event occurred because of the behavior of someone else in the child's immediate environment. A child's I+ score (indicating belief in internal responsibility for successes) is obtained by summing all positive

events for which he assumes credit. A child's I- score (indicating belief in internal responsibility for failures) is obtained by summing all negative events for which he assumes credit. His total I score is the sum of his I+ and his I- subscores.

The developmental sample consisted of 923 elementary and high school students and was drawn from five different schools so that it would be representative of children in diverse kinds of communities. Test-retest reliabilities for the young children were .69 for total I, .66 for I+, and .74 for I-. For the ninth grade students, test-retest reliabilities were .65 for total I, .47 for I+, and .69 for I-. Measures of internal consistency, split-half reliabilities were computed for the separate subscales. For a random sample of 130 of the younger children, the correlations were .54 for I+ and .57 for I-. For a similiar sample of older children, the correlations were .60 for both the I+ and I- subscales.

The authors cite several other statistics to lend some additional support to the construct validity of children's beliefs in their control of reinforcements. Among these are the low correlations between the I+ and I- subscales. As for sex and age differences, I+, I-, and total I scores tend to increase only slightly with age and girls' scores tend to be somewhat higher than boys', especially from grade 6 upward. First-born children in the upper grades tend to give higher total I scores. Children's Social Desirability

Scale scores correlate only slightly with IAR scores. IAR scores predict various achievement measures, especially course grades. The authors discuss all of these findings in the context of the theory developed around internal and external belief systems.

Junior Index Motivation

The Junior Index of Motivation Scale (JIM) was designed to assess the desire of junior high school students to learn in school. It is based on the assumption that whatever causes one to try to achieve in school comes primarily from within rather than from without. It also assumes that this motivation is rooted in the individual's personality structure, his value system and his curiosity (Frymier, 1970).

The JIM scale consists of 80 statements. The student has to choose one of four alternatives to indicate varying levels of agreement.

In one of the studies cited in support of the validity of the scale, students who were seen by teachers as being highly motivated made significantly higher scores than students who were seen by their teachers as being low in motivation. Further evidence of validity was obtained by correlating JIM scale scores with scores from another measure of motivation (Farquhar's M-scale). For grade nine students, the correlation was .44, while for grade eleven students it was .57.

Split-half reliabilities of .83 and test-retest reliabilities of .70 seem to indicate that the JIM scale is internally consistent and dependable over time.

Gill Self Concept Scale

The Gill Self Concept Scale (Gill and D'Oyley, 1968) was designed to produce a measure of the self concepts of students in terms of their perceptions of themselves and their experiences in school. Two dimensions of the self are measured: The self as it is now perceived, and the ideal self. A basic assumption in the development of the Scale was that the individual is capable of making a subjective judgment of himself by arranging his self percepts along a subjective continuum from "never like me" to "always like me." These judgments are the data for the perceived self.

To attain some measure of the values attached to these judgments, ratings of the ideal self are also made by the respondent. The ideal self is defined as the organized conceptual pattern of qualities that an individual considers desirable or undesirable for himself. The task for the respondent, here, is to order his self-perceptions along a four point "value" continuum from "what I would always want to be like" to "what I would never want to be like." An attempt was made to reduce the effects of acquiescence by including some negative statements. The scoring system for positive items is 1, 2, 3, 4. For negative statements, the order is reversed and alternatives are scored 4, 3, 2, and 1.

To obtain a measure of test-retest reliability, the self concept scale was readministered to 67 students about eight weeks after the first administration. For boys, the reliability coefficient for the perceived self scale (.69) was higher than that obtained for the ideal self scale (.60). For girls, however, the stability coefficient for the perceived self scale (.60) was lower than the corresponding value for the ideal self scale (.67).

The factorial structures of the perceived self and the ideal self scales were investigated by means of a principal components analysis. The psychological interpretations that were attached to the interpretable factors appeared to substantiate the validity of the scales.

To assess the usefulness of the scale as a predictor of academic achievement, validity coefficients were computed using final average grades as the criterion. The coefficients for the perceived self scale (.42 for boys and .35 for girls) were higher than those for the ideal self scale (.25 for boys and .19 for girls). Both the perceived self as well as the ideal self scale seem to be more efficient in predicting the academic performance of boys than of girls.

(c) Learning Environment and Attitude Measures

Learning Environment Inventory (LEI)

The Learning Environment Inventory (LEI) was designed to measure the social climate of learning of a class as it is perceived by the students within it. In this study two subject areas were selected so that students could focus

their perceptions on specific referents. These were the mathematics and English classes. (Test booklets alternately referring to a mathematics class or to an English class were randomly distributed to the students.) Choosing specific classes allows students to relate to a restricted range of experiences, thus easing the decision-making process. Mathematics and English classes have such widespread attendance that all students will have some experiences on which to focus their perceptions. Since testing time was short it was not possible to have each student relate to both.

The LEI has two distinct uses according to its author (Anderson, 1971): To assess the perceptions of an individual student of his class, and to gauge the learning environment of the class as a group. The class mean provides the best estimate of collective student perceptions of the class and the class mean should be used when one is examining different conditions or treatments across classes.

The LEI contains 105 statements descriptive of typical school classes and the respondent expresses his agreement or disagreement with each statement on a four-point scale. The 105 items are divided into 15 scales. (There is no overall LEI score.) In selecting the 15 climate dimensions, the author only considered concepts previously identified as good predictors of learning or concepts considered relevant in terms of social psychological theory and research. A description of each scale is presented below:

1. Cohesiveness - The degree of intimacy or the

feeling of cohesiveness that develops within a group.

2. Diversity - The extent to which the class provides for a diversity of student interests and activities.
3. Formality - The extent to which behavior within the class is guided by formal rules.
4. Speed - The extent to which the rate of progress of the class is matched to the characteristics of individual students within it.
5. Environment - The amount of space available and the type of recreational equipment included in the physical environment.
6. Friction - The degree to which the class is characterized by disagreement, tension, and antagonism within the class.
7. Goal Direction - The degree to which the goals of a class are expressed in objective terms and accepted by the class members.
8. Favoritism - The degree to which low academic self concepts characterize the members of the group.
9. Difficulty - The level of difficulty of the work or assignments typically given students.
10. Apathy - The extent to which class members evidence an affinity with class activities.
11. Democratic - The extent to which decision making

is shared by the members of the class.

12. Cliqueness - The extent to which students tend to stick together in small groups.
13. Satisfaction - The degree to which students like or enjoy their classes.
14. Disorganization - The degree to which students consider the class disorganized.
15. Competitiveness - The extent to which competition prevails within the class.

Survey of Study Habits and Attitudes

The Survey of Study Habits and Attitudes (SSHA) instrument was developed to measure study methods, motivation for studying, and certain attitudes toward scholastic activities which are important in the classroom. The purposes of the SSHA are: (a) to identify students whose study habits and attitudes are different from those of students who earn high grades; (b) to aid in understanding students with academic difficulties; and (c) to provide a basis for helping such students improve their study habits and attitudes and thus more fully realize their best potentialities.

One form of the SSHA can be used with grades 7-12. It consists of 100 statements concerning study activities and attitudes. The student replies to each statement with one of the following answers: rarely, sometimes, frequently, generally, or almost always. The 100 statements were originally categorized by psychologists into four basic

subscales. The particular subscales and subscores of the SSHA are as follows: Work Methods (use of effective study procedures, skill and efficiency in doing academic assignments) plus Delay Avoidance (promptness in completing assignments and ability to resist distractions) combine to yield a Study Habits score (a measure of academic behavior). Teacher Approval (feelings and opinions about teachers, their classroom behavior, and their methods) plus Education Acceptance (approval of educational objectives, practices and requirements) combine to yield a Study Habits score. The Study Habits score plus the Study Attitudes score combine to give a total Study Orientation score (an overall measure of study habits and attitudes).

Subscale intercorrelations ranged from .44 to .84 for men and from .27 to .76 for women, with medians of .53 and .39 respectively. Kuder-Richardson formula 8 estimates of internal consistency yielded coefficients for the four basic subscales ranging from .87 to .89. Test-retest correlations after a fourteen-week interval ranged from .83 to .88. The authors concluded that the four subscale scores are sufficiently stable through time to justify their use in predicting future behavior or in assessing the degree of change in study habits and attitudes after counseling (Brown and Holtzman, 1964).

Extensive validity evidence is presented by the authors. Validity coefficients reported on SSHA total scores with grade point averages ranged from .25 to .45.

with a time limit of two minutes for each of five problems. Responses are scored for the following factors:

Ideational Fluency - Number of different responses,

Spontaneous Flexibility - Number of different categories of responses,

Originality - Sum of responses that did not fall into any of the following categories:

1) Things that are actually possible and have been or are being done,
2) Irrelevant responses,
3) Commonplace consequences, frequently talked about consequences in folklore, mythology, etc.

A number of attempts to determine the validity and reliability of the Torrance Tests of Creative Thinking are reported. A discussion of the problems involved in establishing validity for creativity tests was presented in Chapter 3. Test-retest reliability estimates obtained from studies of mentally retarded youngsters and average fifth grade students range from .61 to .93. Torrance holds that the tests developed through the Minnesota Studies of Creative Behavior are applicable to students from kindergarten through graduate school.

Controlled Associations Test

This test is adopted from Thurstone and copyrighted by ETS. In reference to Guilford's Structure of Intellect

Additional validity evidence showed that the partial correlation between SSHA total scores and grade point averages with scholastic aptitude held constant was highly significant, ranging from .41 to .47. According to Brown and Holtzman (1964) these results combined with others reported in the test manual clearly indicate the importance of the SSHA in providing measures of personal traits that are relevant to academic success but are not covered by scholastic aptitude tests.

(d) Creativity Measures

The Consequences Test

The Consequences Test was originally developed by Guilford and his associates (1951) to provide measures of ideational fluency (divergent production of semantic units) and originality (divergent production of semantic transformations). The test was scored for fluency by simply counting the numbers of obvious consequences. The number of remote consequences produced an originality score. The originality factor, requiring the ability to produce clever, or uncommon responses, appears to be a relatively stable grouping, having been found six times in Guilford's project. Torrance's (1962) modification of the Consequence Test was chosen for use in this study. As in Guilford's test, students are presented with improbable situations and asked to list as many outcomes of these hypothetical circumstances as they can. The Consequences Test is administered as a group test

Model, it is a measure of associational fluency or a factor identified with the divergent projection of semantic relations. It requires the ability to produce words from a restricted area of meaning. Respondents are told to write as many snynonyms as possible for each of four words. The score is the number of words written that are related to the stimulus word. Performance on this test involves an awareness of similiarity in the meanings of words amid the differences. The more associations that the examinee has that are tied to a word and the more he is willing or flexible enough to work at a crude level of analogy or similarity, the higher will be his score. The test has been found suitable for sixth graders through college.

The Word Beginnings Test

Both the Word Beginnings and the Word Beginnings and Endings Test are measures of the factor of divergent production of symbolic units or word fluency. They require the ability to produce many words that conform to simple specifications not involving meanings. In the Word Beginnings Test, the examinee is asked to write words beginning with a specified prefix. This task is very similar to Thurstone's Suffix Test where the score is the number of words written. The Word Beginnings Test consists of two parts, each presenting a different prefix. The respondents are given three minutes in each part to record their answers. The test is suitable for grades 6-16 and can be obtained from ETS.

The Word Beginnings and Endings Test

Most of the discussion of the Word Beginnings Test is also appropriate for the Word Beginnings and Endings Test with the exception of the description of the task required of the examinee. The Word Beginnings and Endings Test is very similar to Thurstone's First and Last Letters Test and involves writing as many words as possible beginning with one given letter and ending with another. Reliability, validity and norming information, usually presented in test manuals, are not provided for this test, the Word Beginnings Test nor the Controlled Associations Test. The reason for this is that in each case the test was designed mainly for factorial research purposes.

4.3 Problems Investigated

This section provides a further elaboration of the purposes of the study stated in Chapter One. It was delayed until this point so that the variables under investigation would be clear. It is divided into two parts. The first part provides a statement of questions relating to achievement, attitudes, and creativity. The second part includes a set of questions relating to aptitude-treatment interactions.

(a) Achievement, Attitudes, and Creativity

One of the basic assumptions of an individualized instructional program is that learning is ultimately personal

and individual. In the present study, it was felt that instructional treatments 1, 2, and 3 offered students the greater number of alternatives. Thus they should be better able to accommodate a wide range of individual differences. The prediction was that students in these programs would be more favorable toward school and school-related activities than students in the other programs as a consequence.

It was also expected that students in instructional treatments 1, 2, and 3 would have greater confidence in themselves, be more productive when working in an unstructured situation and more adept at coming up with alternative solutions to problems. Thus, they were also expected to have higher mean scores on tests of fluency, flexibility and originality.

In view of other research in the area, it was also expected that students in the various instructional treatments would differ in terms of achievement.

(b) Differential Effects as a Function of Treatment

This part of the study represented an exploratory search for ATI's where aptitude was variously defined as scores on personality measurements and tests of fluency, flexibility and originality. One of the goals that is generally expressed for instructional programs is that student satisfaction and attitudes toward school and school related activities be enhanced. Therefore, responses on the Survey of Study Habits and Attitudes test were considered

as one criterion. The other was scores on a standardized achievement measure.

Given the descriptions of the treatments in this study, it was possible to generate a number of hypotheses of interaction. These hypotheses were derived from previous research findings and logical reasoning. In instructional treatment 4 and 5, the teacher was primarily responsible for setting the tone of the class as well as defining acceptable student behaviors. This kind of approach would be expected to stimulate pupil conformity. Instructional treatments 1, 2, and 3 should facilitate independence and individual responsibility. If this were the case, then students who excelled in instructional treatments 4 and 5 should score high on the Ac (achievement via conformity) scale of the California Psychological Inventory (CPI) and low on tests of fluency, flexibility and originality as well as the Ai (achievement via independence) scale of the CPI. Students who excelled in instructional treatments 1, 2, and 3 were expected to score high on the AI scale and high on tests of fluency, flexibility and originality but low on the AC scale.

Beach's (1960) results indicating differential student performance based on sociability suggested that outgoing, enterprising students would probably perform better in instructional treatments 1, 2, and 3 than in an environment where continuous interactions with other people was discouraged. In instructional treatments 1, 2, and 3, students were freer to pursue activities they selected. Learning

could well be enhanced by students gathering in groups and discussing, challenging and stimulating one another. If this were true, then students scoring high on the Sy (sociability) scale of the CPI should fare better in instructional treatments 1, 2, and 3. In Beach's study the less sociable student performed best in lecture classes. Similar results were predicted for less sociable students in the present study. Less sociable students were expected to perform best in instructional treatments 4 and 5.

An interaction between flexibility and school attended was also expected with the Fx (flexibility) scale of the CPI. The flexible person is adventurous, confident, idealistic, assertive, and highly concerned with personal pleasure and diversion. Such an individual would be expected to excel in a highly unstructured school environment where he is freer to follow his own inclinations. The deliberate, cautious, worrying, industrious individual who is overly deferential to authority or custom would probably perform better in instructional treatments 4 and 5.

Therefore, it was hypothesized that students with high scores on the Fx scale of the CPI would have higher scores on the standardized achievement tests when they were in instructional treatments 1, 2, and 3; conversely, students with low scores on the Fx scale would have higher scores on the standardized tests when they were in instructional treatments 4 and 5.

The Ie (intellectual efficiency) scale of the CPI

indicates the degree of personal and intellectual efficiency which the individual has attained. Low scores tend to represent conventional and stereotyped thinking; such individuals are also seen as lacking in self-direction and self-discipline. In the present study students with low scores on this dimension were expected to perform best in instructional treatments 4 and 5, where the teacher provided guidance, support and direction. Individuals with high scores on the intellectual efficiency scale tend to be efficient, clear-thinking, planful, and resourceful. They are also alert and well-informed and tend to place a high value on cognitive and intellectual matters. Such students would be expected to achieve best in instructional treatments 1, 2, and 3.

The several studies suggesting that anxiety interacts with environment in determining student performance in structured and unstructured classes (Smith et al., 1956; Grimes and Allinsmith, 1961; McKeachie, 1951) and in teacher-centered approaches to instruction (Dowaliby and Schumer, 1973) suggest it may also be profitable to look for similar interactions with the treatments in this study. The prediction was that the achievement of students high in anxiety would be higher in instructional treatments 4 and 5 than in instructional treatments 1, 2, and 3. The rationale for this was that in instructional treatments 4 and 5, student responsibilities and course requirements would be more clearly delineated. The instructor would determine the requirements and the standards of performance.

The student high in anxiety would be expected to perform better in a school where he knew exactly what he was to do. The student low in anxiety, being freer to pursue his own inclinations, should achieve highest in instructional treatments 1, 2, and 3.

It was hypothesized that a disordinal interaction would be obtained with regards to the Do (dominance), Cs (capacity for status), and Gi (good impression) scales of the CPI. Students who succeeded in instructional treatments 4 and 5 should have higher scores on each of these variables than students who scored lowest on the criteria. The reverse was hypothesized in the case of instructional treatments 1, 2, and 3. It was felt that students with high scores on the Do, Cs, and Gi scales would learn best in instructional treatments 3 and 4 where competitiveness was a subtle consequence of the teacher's exercising greater control over the learning environment.

4.4 Procedure

The studies reported here grew out of a larger evaluation. The results of that evaluation have been disseminated as a final report to school systems in Massachusetts (Hambleton, Rovinelli, Sheehan, & Newby, 1972). A considerable amount of time went into the planning of these investigations. To begin with, the semester preceding the gathering of data was spent reviewing papers, tests, and procedures that were thought to be of potential use.

Preparation of Tests and Questionnaires

Since the intent was to machine score as many of the tests as possible, many of the test directions had to be rewritten. In some cases, where tests were used that had appeared in education and psychology journals, original test directions were written. To insure the appropriateness of the directions prepared for the tests, principals, guidance personnel and interested teachers in the schools were asked to comment on the drafts. Modifications were then made on the basis of their responses. Several of the instruments were pretested with sixth grade children in a school in Vermont to further substantiate the validity of the procedures followed. Finally, permission was obtained from Consulting Psychologists Press, Inc. to reproduce a subset of 291 items from the California Personality Inventory for inclusion in the study.

Preparation of Students and Teachers

All tests were administered by teachers and guidance counselors at each school. To help insure that a standard procedure was followed, an examiner's manual was prepared outlining every step in the testing process. Training sessions were held for the teachers and guidance counselors. Every step outlined in the manual was discussed with the teachers. The rationale for including the various tests was also repeated. That the success of the testing program depended, to a great extent, on the attitudes of the test

administrators towards the evaluation was stressed in all meetings with the teachers as well as in the examiner manual.

About one week prior to the first test day, a meeting was held with all participating students to inform them personally of the study and explain its importance.

The talk that was read by various members of the evaluation team to all participating students in their classrooms was as follows:

Good Morning:

My name is _____ and I am a member of an evaluation team from the University of Massachusetts who has been asked to find out how you feel about school and the many things you do while in school. Since there are many students and only four of us on the evaluation team, we cannot sit down and talk with you individually, therefore we have made up a number of questionnaires which we are going to ask you to take during the next couple of weeks so that we can find out about your feelings. These questionnaires are not like others you have taken in the past. In fact, we think you may even enjoy taking most of our questionnaires. The questionnaires will be used to determine things about students in this school that we hope will eventually be used to improve your school. For most of the questionnaires there are no correct answers. Therefore we encourage you to give honest answers, since if you do otherwise, that is, if you give dishonest answers, the results of our study will be meaningless. Since some of the questions are personal we are going

to ask you to put a special identification number on each answer sheet instead of your name so that no one else in the school will know how you answered the questionnaires. After you finish each questionnaire, a fellow student will collect the answer sheets and put them in an envelope which he will seal. Members of my team will then collect the envelopes at the end of the day. We will analyze the results and give a report to your school about how the whole student body feels about the various ideas in the questionnaires. Do you have any questions about anything I have said or perhaps not made clear?

A modified version of the above statement was read again on the morning of the first test as a reminder and also to catch any students who may have missed our earlier talk. Discussion with students included a brief question and answer session which appeared to go well.

The point most emphasized in that meeting was that all responses to questions were to be kept in strictest confidence. To insure the privacy of individual responses each student was assigned a number which was recorded on all answer sheets in place of names. At the end of a testing period, one of the students collected the answer sheets, placed them in an envelope, sealed it, and carried them to the office where they were picked up by members of the evaluation team. Testing time was spread out over at least four days. Students were in testing for only about two hours each day, so it is

unlikely that fatigue would be a confounding factor in this study.

4.5 Experimental Design

Comparison Studies

The first part of the investigation was concerned with overall differences among the instructional treatment groups. Four basic analyses were run. A univariate analysis of variance was conducted on composite scores derived from the ITBS in the case of treatment 5 and the SRA achievement series for the other treatment groups. A univariate analysis of variance was also computed on the Study Orientation scores. A multivariate analysis of variance was conducted on the fifteen LEI subscale scores and also on the fluency, flexibility, and originality test scores.

This part of the investigation could be characterized as a post-instructional analysis of the five instructional treatment groups across the three grades. The basic design, showing the total number of students per treatment x grade combination is given in Table 4.5.1.

Justification for applying multivariate analysis on the LEI and creativity scores is appropriate at this point. An alternative procedure would be to utilize univariate analyses of variance conducted on each of the variates separately. Bock and Haggard (1968), however, suggest that when univariate tests, such as F-tests, are performed on each variable separately a single probability statement applicable to all variables jointly cannot, in general, be obtained from the

Table 4.5.1

Representation of the Experimental Design for the Comparison Studies[1]

Grade	Treatment 1	2	3	4	5
Six	Group 61 (n = 167)	Group 62 (N = 77)	Group 63 (n = 99)	Group 64 (n = 160)	Group 65 (n = 52)
Seven	Group 71 (n = 130)	Group 72 (n = 56)	Group 73 (n = 44)	Group 74 (n = 149)	Group 75 (n = 54)
Eight	Group 81 (n = 101)	Group 82 (n = 36)	Group 83 (n = 34)	Group 84 (n = 149)	Group 85 (n = 54)

[1] $Group_{ij}$ corresponds to the j'th treatment in the i'th grade.

separate F ratios.

These dependent variables may be correlated in some arbitrary and unknown way, and the separate F-tests would not be statistically independent. No exact probability that at least one of them will exceed some critical level on the null hypothesis can be calculated. Multivariate tests, on the other hand, are based on sample statistics which take into account the correlations between variables and have known exact sampling distributions from which the required probabilities can be obtained.

One might then raise the possibility of throwing all of the variates under consideration together and conducting an overall multivariate analysis of variance. This approach was a feasible alternative. It was rejected because of missing scores. Since complete data profiles were not available for all subjects in the study, the choice of reducing the sample size for the entire study (resulting in a reduction in power) had to be weighed against running separate analyses on the subsets of variates. The latter alternative was selected with full recognition of the consequences on the overall type I error rate.

All students but those in instructional treatment 5 had scores on the SRA Achievement Series. Students in instructional treatment 5 had taken the Iowa Tests of Basic Skills (ITBS). In order to impose a common metric on the achievement test data, the Iowa Tests of Basic Skills scores were converted to stanines to match the SRA achievement series.

While there was general commonality among the subtests of each test, it seemed most appropriate to equate scores on three of the subtests: Total reading, total language, and total mathematics. Scores on these subtests were expressed as stanines.

The tests of fluency, flexibility, and originality required hand scoring. Since the process involved careful reading and judging of responses on subjective criteria, reliability estimates were obtained. Two separate checks resulted in Pearson product-moment correlations of .83 and .85. Interrater agreement, then, was high for this part of the study.

ATI Investigation

The aptitude-treatment interaction phase of the investigation focused on the regression of the study orientation (SSHA) and the SRA/ITBS composite achievement scores on the following aptitude variables:

Self Concept

Intellectual Achievement Responsibility

School Anxiety

Personality (9 scales)

Achievement Motivation

Study Habits and Attitudes

Creativity (6 scales)

The composite achievement score was selected as a criterion because it appeared to be the most reliable and

valid measure of general educational achievement available to us. The Study Orientation Score, by the same token, was the best overall measure of student affect available. Both criteria were widely accepted by the schools involved in the study.

The basic design for the ATI studies has been presented. By way of review we note the following features: Students in the different instructional treatments constituted our sample. There were no selection factors at work. All students at the designated schools were included in the pool of subjects considered for the study. It was felt that real differences existed between the treatments and that these differences could be utilized in planning programs which result in improved student achievement and satisfaction. A number of aptitudes thought to be relevant were identified. Appropriate measures of these aptitudes were then selected and administered to the students.

4.6 Method of Analysis

Comparison Studies

The data were analyzed within a multivariate general linear hypothesis framework. Standard tests of hypotheses of the form

$$A \; B \; C = 0 \qquad 4.6.1$$

were tested. In this equation, B denotes the matrix of unknown parameters specified by the design matrix, X, in the

matrix equation,

$$Y = X B + \varepsilon \qquad 4.6.2$$

where

(i) Y is an N x p matrix of p dependent measures on each of N experimental units. In the typical behavioral experiment, a general element of Y, $y_i^{(j)}$, would designate the response measure for the i'th subject on the j'th dependent variate.

(ii) X is an N x m matrix of m known predictor and/or design variables on each of N experimental units. For example, in the multiple linear prediction situation, X might contain the measurements on m predictor variables for each of N individuals. In an analysis of variance situation, X would be a matrix which describes the actual experimental design under which the data were obtained as elements. X has rank m with $r \leq m < N$. Usually X is reparameterized such that the number of columns in X equals its rank.

(iii) B is an r x p matrix (assuming X is reparameterized to be N x r) of unknown parameters specified by the hypotheses of interest. Depending upon the choice of X, the elements of B may represent any contrasts among population parameters, expected values of the dependent variates, population regression coefficients, etc.

(iv) ε is the random error component.

The matrices A and C in (4.6.1) are specified by the researcher and are used to select contrasts among the elements of B.

The linear model for the 3 x 5 analysis of variance was

$$y_{ijk}^{(\ell)} = \mu^{(\ell)} + \alpha_i^{(\ell)} + \beta_j^{(\ell)} + \alpha\beta_{ij}^{(\ell)} + e_{ijk}^{(\ell)} \qquad 4.6.3$$

where

(i) $y_i^{(\ell)}$ denotes the measure on the ℓ^{th} dependent variate for the i'th subject;

(ii) $\mu^{(\ell)}$ denotes the usual overall grand mean effect of the ℓ'th dependent variate;

(iii) $\alpha_i^{(\ell)}$ denotes the effect due to grades for the ℓ^{th} dependent variate;

(iv) $\beta_j^{(\ell)}$ denotes the effect due to school for the ℓ^{th} dependent variate;

(v) $\alpha\beta_{ij}^{(\ell)}$ denotes the grade x school interaction for the ℓ^{th} dependent variate;

(vi) $\varepsilon_{ijk}^{(\ell)}$ denotes the error component associated with the ℓ^{th} dependent variate.

The most general hypothesis tested is

$$A\ B\ C = D \qquad 4.6.4$$

where

(i) A is a g x r matrix of rank g whose elements, a_{ij} (i = 1, 2. . ., g; j = 1, 2. . ., r) are used to select particular combinations from the rows of B;

(ii) C is a p x u matrix of rank u whose elements, c_{ij} (i = 1, 2. . ., p; j = 1, 2. . ., u) are used to select linear combinations among the columns of B; and

(iii) D is any specified g x u matrix of constants.

Test Criteria

Multivariate test criteria are usually a function of the characteristic roots of HE^{-1}. Three popular test criteria can be cited. Wilks likelihood ratio criterion is the most widely applied. This criterion makes use of the statistic

$$\Lambda = \left[\prod_{i=1}^{u} (1 + \lambda_i) \right]^{-1}$$

where λ_i (i = 1, 2, . . ., u) are, again, the characteristic roots of HE^{-1}. An equivalent form of the above expression is

$$\Lambda = \frac{|E|}{|H + E|}.$$

If N is sufficiently large, the quantity

$$\chi^2 = - [N - r .5 (u - g + 1)] \ln\Lambda$$

is distributed as chi square with g x u degrees of freedom (Bartlett, 1951). A better approximation (Rao, 1965) is given by

$$F = \frac{1 - \Lambda^{1/s}_{u,g,\ N-r}}{\Lambda^{1/s}_{u\ ,\ g,\ N-R}} \cdot \frac{s[2(N-r) + g-u-1] - 15(gu-2)}{gu}$$

where

$$s = \sqrt{\frac{g^2 u^2 - 4}{g^2+u^2 - 5}}$$

Under the null hypothesis, F is approximately distributed as a F statistic with gu and {S[u(N-r) + g - u - 1]} - .5(gu-2) degrees of freedom.

Once significance has been determined, interest focuses on determining the nature of the effect and the source of the differences in terms of the variates which contribute to it. Following a rejection, Cramer and Bock (1966) recommended that univariate analysis of variance be run on each variate separately. This suggestion has been echoed by Hummel and Sligo (1971). The reason this procedure was not followed was that a large number of analyses conducted in this fashion would increase alpha substantially. When the variables are independent, the experiment-wise error rate is given by 1 -

$(1 - \alpha)^P$. (P is the number of variables in the study). Since the variables in the present study were not independent, we have no way of determining the overall type I error. It would, however, increase as a function of the number of variables. For this reason significant effects were followed by discriminant analysis to determine the nature of differences.

Computer Program for Tests of the Multivariate Linear Hypothesis

The statistical analyses in this phase of the investigation were conducted using Mulgen, a computer program written by Olson (1970) in Fortran IV. Designed to run on the CDC 6400 computer system, it provides the following features:

(i) double precision arithmetic;

(ii) printouts of the basic matrices utilized in the analyses;

(iii) significance tests of general linear hypotheses of the type given in (4.6.4) using Wilks maximum likelihood criterion and the corresponding F approximation; and

(iv) efficient methods of handling data input.

Background on the ATI Paradigm

Typically, we alter the learning treatments for individuals or groups of individuals in one of two ways: (1) by altering the rate of presenting materials to different students or (2) by presenting qualitatively different materials to different students in the learning groups. In

ATI we are interested in matching modes of instruction (loosely called treatments) to the learner.

Let us assume that some area of instruction has been identified in which it is important that all students be able to perform. This performance outcome or criterion might be called p. Typically, we design competing methods or treatments, say M_1 and M_2, and we then attempt to determine which of the two methods yields the greatest amount of average p. Large and significant differences favoring either M_1 or M_2 are rarely found. However, if one examines the variability of students on the criterion measures p around the treatment means, one frequently finds wide pupil variability.

The ATI strategy attempts to utilize this variance in seeking leads for developing and improving treatments that will interact with aptitudes in yielding high level performance on the criterion. The variability around the treatment means for individuals can be theoretically broken down into a source of random error or measurement and into one or more systematic sources of individual differences. Hopefully, these systematic sources of individual differences form constructs of ability that can be matched with appropriate treatments. ATI methodology then becomes a system to maximize the output of p as an interactive function of type of treatment and aptitude patterns of the learner.

The ATI procedure might also be presented as a form of moderated prediction system. Treatments can be conceptualized

as the moderator variables, i.e., they moderate the regression of criterion scores on aptitudes. As an example, we can consider some hypothetical ATI outcomes. The symbol A will be used to represent the aptitude dimension or measure, the symbol P the outcome dimension or measure, and the regression lines of P on A will be represented by M_1 and M_2 (methods 1 and 2). Assume there is variability around the regression lines.

An inspection of Figure 4.6.1 prompts us to draw three conclusions:

1. Method 1 is superior to Method 2 at all levels of A.
2. No interaction exists.
3. Subjects will perform better at all levels of A if they are assigned to the M_1 condition.

Early in the history of ATI research, a significant interaction effect was considered to be ordinal when the treatment lines did not cross (Figure 4.6.2) and disordinal when the treatment lines crossed (Figure 4.6.3). Bracht (1970) argued that this distinction does not provide adequate protection against a type I error. There seems to be consensus now that the crossing of the treatment lines is not a sufficient requirement for the existence of a stable ATI.

For research on ATI's that uses a treatments-by-levels factorial design, Bracht and Glass (1968) suggest that an

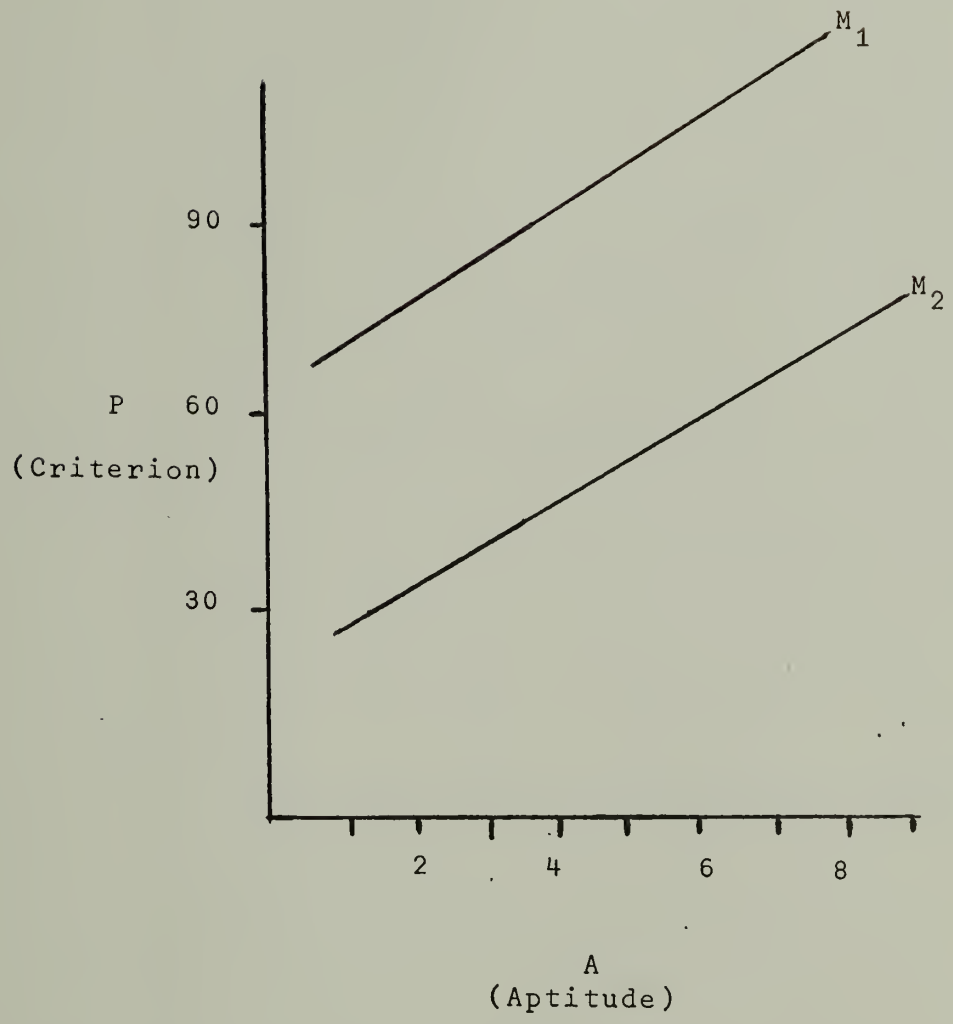


Figure 4.6.1 Illustration of no interaction between aptitude and treatment.

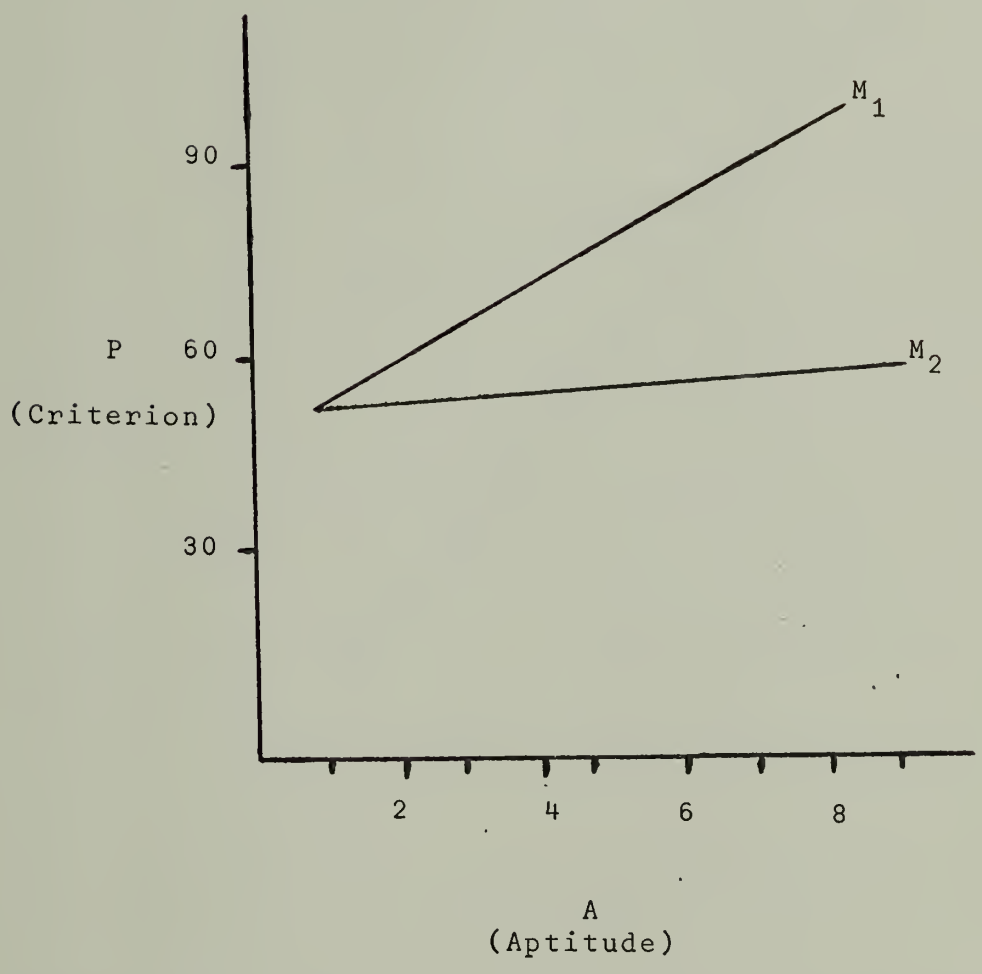


Figure 4.6.2 Illustration of an ordinal aptitude treatment interaction.

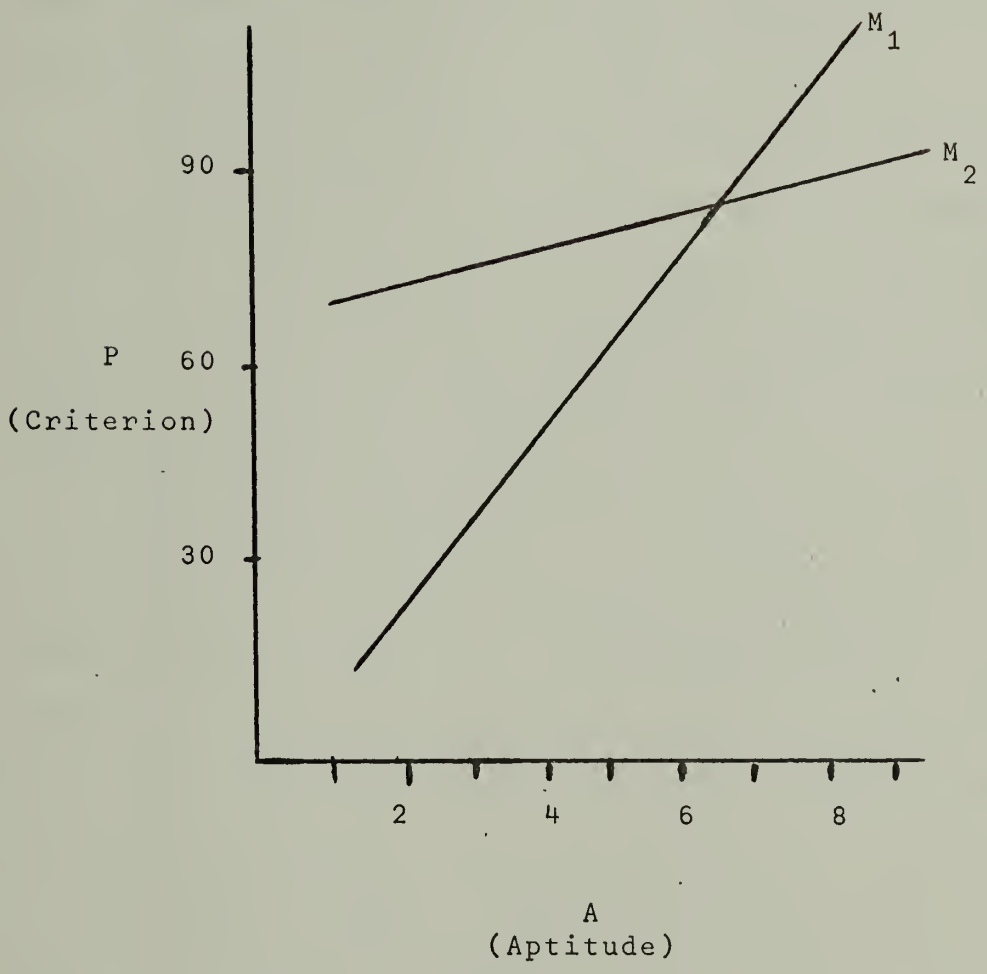


Figure 4.6.3 Illustration of an ordinal aptitude-treatment interaction. (Differences between the regression lines at the high end of the aptitude scale are not significant.)

interaction effect should be defined as disordinal only when the differences between alternative treatments at two levels of a personalogical variable are both significantly non-zero and different in algebraic sign. Thus the interaction in Figure 4.6.3 would be regarded as ordinal by Bracht and Glass because of the non-significant treatment difference in the high ability group.

A disordinal interaction is represented in Figure 4.6.4. The perpendicular lines intersecting A approximately at points 2 and 6.5 create three zones. The "D2" zone indicates a point on the A scale where subjects do "significantly" better on P if they are assigned to the M_2 condition than if assigned to the M_1 condition.

The other critical zone is labelled "D1." Subjects with scores of 6.5 or better on the A measure will perform significantly better on P if they are assigned to the M_1 rather than to the M_2 condition. Subjects with scores below 2.0 on the A measure will perform significantly better on P if they are assigned to the M_2 rather than to the M_1 condition.

The zones between 2 and 6.5 on the aptitude dimension can be interpreted as meaning that students show no significant differences on P as a function of being placed in M_1 vs. M_2. Whatever differences are observed between groups in the region 2 to 6.5 are attributed to sampling error.

The procedure for detecting interactions requires testing the parallelism of regression slopes. If the test of parallelism is rejected and the regression lines cross within

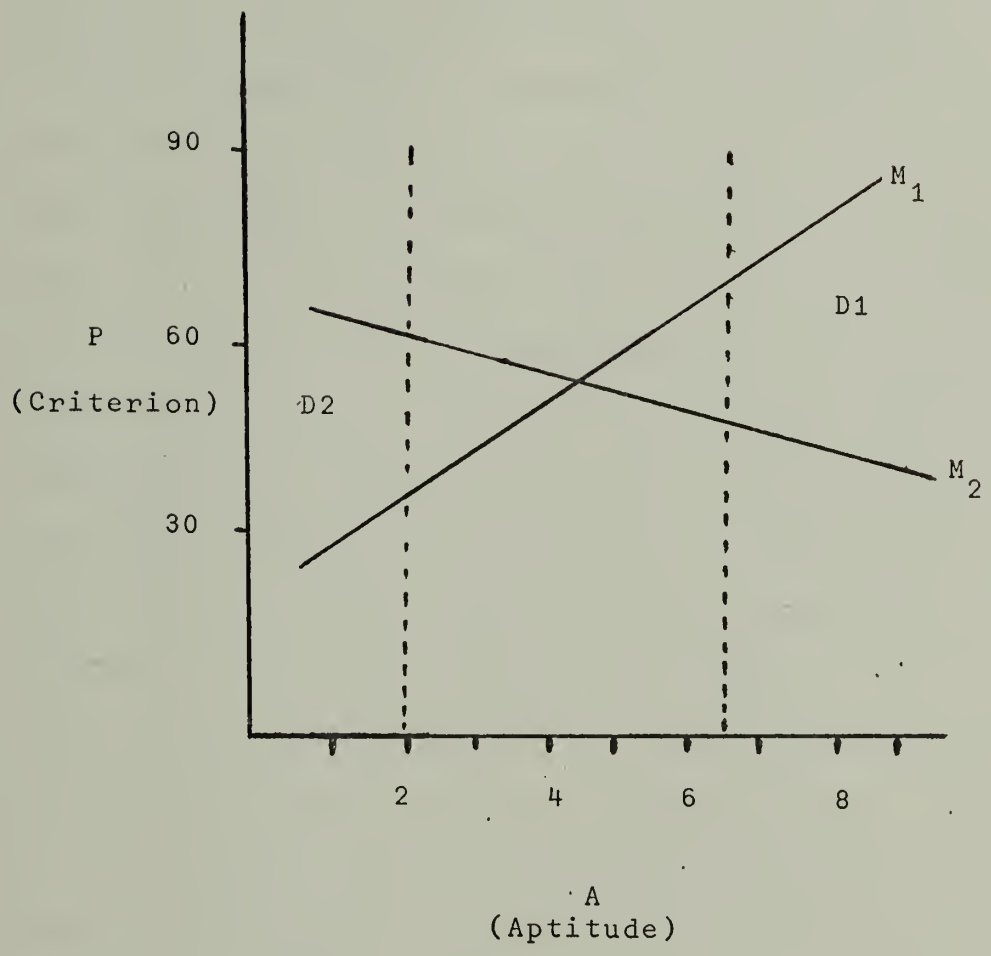


Figure 4.6.4 Illustration of a disordinal aptitude treatment interaction.

the range of the measured aptitude variable the interaction is disordinal. If one fails to reject the null-hypothesis test of parallelism, then no interaction exists.

The computer program used in this study follows the analysis of covariance method of Gulliksen and Wilks (1950). The general purpose of the method is to entertain three statistical hypotheses. Initially, the hypothesis of the homogeneity of variance of the criterion scores about the regression line of criterion score on aptitude score is tested for the two groups. The statistic used for the test is χ^2 (Chi square). If the hypotheses is rejected, it is not appropriate to go on with the tests of the second and third hypotheses. If, however, we do not reject the initial hypothesis, we test the hypothesis that the slopes of the regression lines are equal. Specificially, the second hypothesis is that the slope of the regression of criterion scores on aptitude scores is the same for each treatment group. The statistic that tests this hypothesis is F. From the point of view of ATI research, this second hypothesis must be rejected for significant interactions to exist. To complete the cycle, the third hypothesis is a test to see if the regression lines are identical. Specifically, the final hypothesis states that the intercept of the regression of criterion scores on aptitude scores is the same for the two treatment groups. From an analysis of covariance point of view, if one fails to reject all three hypotheses, the conclusion is that the groups are from the same basic

population, or at least from populations having regression lines with the same slope and the same intercepts, and having the same standard errors of estimate. As far as ATI's are concerned - once the homogeneity of variance assumption is substantiated, the second hypothesis becomes the important one. A significant interaction exists only if the second hypothesis is rejected.

Two approaches have been used in reaching decisions about the points along the aptitude continuum that yield significantly different amounts of outcome as a function of treatment assignment. One approach involves the setting of confidence bands around the regression lines. The other approach is the Johnson-Neyman technique (1936).

The Johnson-Neyman technique has special application to ATI research. When the experimenter rejects the hypothesis of equal slopes of regression, the Johnson-Neyman technique can be used to define the regions of the predictor space (personalogical variables) in which the treatments are significantly different on the criterion. Thus, the Johnson-Neyman technique is used to test for ordinal versus disordinal interactions between treatments and personalogical variables. Although the technique was originally developed for designs with 2 treatments and two personalogical variables, it has been extended to the case of more than two personalogical variables (potthoff, 1964); Abelson, 1953).

A dearth of evidence on just what student characteristics interact with what instructional programs or educational environments is the most immediate obstacle to matching students with instruction.

Much of the formal research on teaching centered about differences in classroom instructional methods, with academic achievement being studied as the dependent variable. However, comparative little attention has been given to the personality of the learner as he performs under the various forms of instructional methods and in varying learning situations. One of the major reasons for the lack of cumulative knowledge on the matching of students to learning experiences has been the failure to take seriously the implications of an interactive model that coordinates the effects of educational environments upon particular types of students to produce specific objectives. The need for considering individual differences in instructional planning has been recognized and is often expressed as a desirable goal. Yet educational planners and decision-makers continue to work from models for the student-in-general. The consideration of the importance of differential student characteristics leads to questions about the general effectiveness of educational procedures, such as whether a discovery approach is more effective than a structured approach. Little or no account, until recently, has been given the differential effectiveness of such approaches on different kinds of students.

Chapter V

Results and Discussion

5.1 Introduction

The chapter is organized around the five analyses that were conducted on the data. The first two were univariate analyses of the composite achievement scores and the study orientation scores. The third and fourth were multivariate analyses of the subscales of the learning environment inventory (15 scales) and the creativity scores (six scales). For each of the first four analyses a two factor design was considered. The factors were grade (3 levels) and instructional treatment (5 levels).

The final section is a report of the ATI investigation. The means and standard deviations of all aptitude and dependent variables for the students in the three grades and five instructional treatments are reported in Tables 5.1.1 to 5.1.3.

5.2 Analysis of Achievement Test Data

A two way analysis of variance was conducted on the composite achievement test scores to test the hypothesis of equal effects of experimental treatments. To accommodate the unequal cell frequencies, the data was appropriately analyzed by the method of unweighted means (Anderson and Bancroft, 1952). Analysis is carried out directly on the

cell means, but the error mean square is obtained by dividing the within-cell mean square by the harmonic mean of the subclass numbers. The results are presented in Table 5.2.1.

The F ratio was significant at the .01 level. Hence, the data contradicted the hypothesis that the main effect of schools was zero. The grade by school interaction was also significant ($p < .01$). A graphical representation of the means is given in Figure 5.2.1. This figure represents the profiles corresponding to the simple effects of the grades for each of the schools.

It is noted here that since scores on the achievement tests had been rescaled by conversion to stanines within grades the possibility of a main effect for grades was eliminated. However, rather than analyze each grade separately, for convenience the data was analyzed in a two-factor design.

When subjected to simple effects analysis, the data indicated differences between the schools at each grade level. A further breakdown of the interaction for the eighth grade groups using the Newman-Keuls procedure showed the following results. Students in instructional treatment 5 scored significantly higher on the criterion than students in treatment 2 and treatments 1, 3, and 4. Students in treatment 4 also scored significantly higher than the students in treatment 2. Individual comparisons for the seventh grade groups indicated that the effects of instructional treatment 5 was different from the effects of the other schools. The students

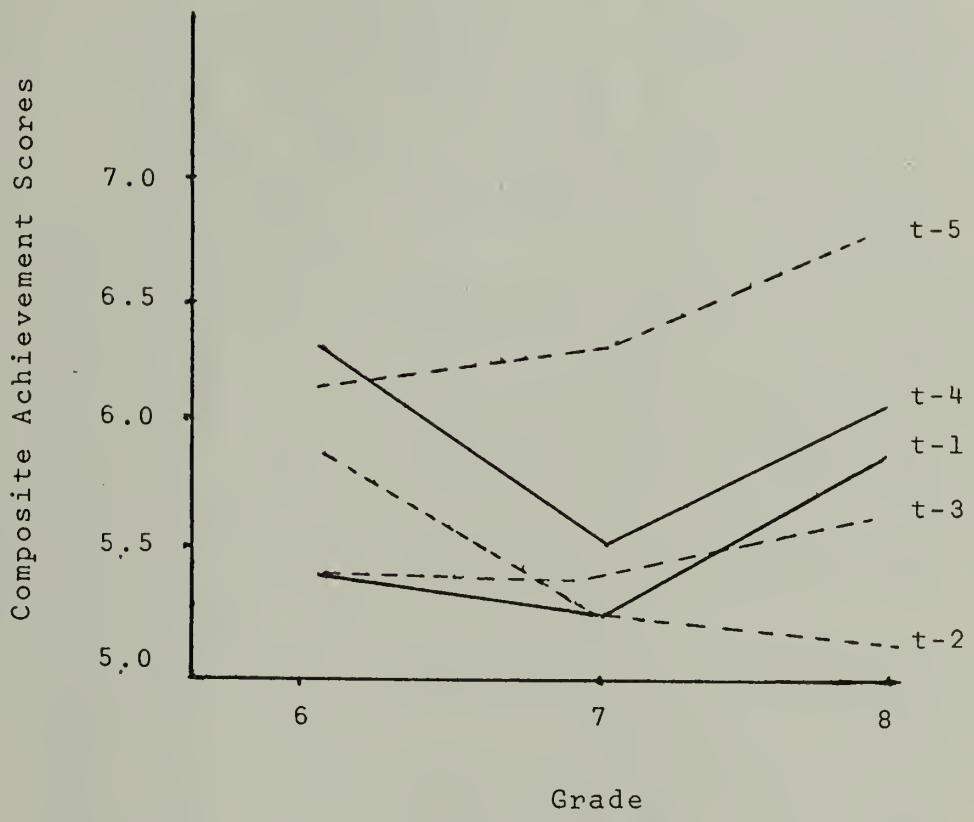


Figure 5.2.1 Profiles of Simple Effects for Grades as a Function of Composite Achievement Scores.

TABLE 5.1.1

Means and Standard Deviations for the Sixth Grade Groups on the Variables Considered in the Study

Variable	Treatment 1		Treatment 2		Treatment 3		Treatment 4		Treatment 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Study Habits and Attitudes										
Delay Avoidance	30.9	4.3	30.0	3.7	30.6	3.6	31.7	4.6	30.5	4.0
Work Methods	32.3	4.6	32.0	3.8	30.8	4.0	32.4	4.9	30.9	4.0
Study Habits	63.2	8.2	60.1	6.9	62.1	6.7	64.0	8.8	61.3	7.3
Teacher Approval	33.1	5.7	32.6	5.6	34.1	5.5	34.1	5.9	31.3	6.8
Educational Acceptance	33.3	4.6	33.0	3.0	32.3	6.1	33.8	4.8	31.4	3.4
Study Attitudes	66.3	9.8	66.4	9.1	66.9	9.0	67.9	9.9	62.7	6.5
Study Orientation	229.5	16.9	228.5	14.2	224.7	14.4	231.9	16.8	224.0	12.6
Self Concept Scale	175.4	22.2	179.1	17.0	183.4	19.1	185.6	19.4	177.5	21.3
Intellectual Achievement Responsibility Scale	22.8	4.5	23.0	4.0	25.0	4.3	24.3	4.0	21.3	4.8
General Anxiety	14.9	6.7	14.0	7.9	13.0	6.8	14.6	7.2	17.1	7.4
School Anxiety	19.0	4.9	17.6	3.8	19.2	4.4	18.3	4.5	20.0	5.5
Achievement Motivation	106.1	21.8	120.2	20.6	120.4	20.6	121.6	19.3	117.0	21.6
Learning Environment										
Cohesiveness	19.3	2.5	18.4	2.6	19.3	3.2	19.2	2.6	18.2	2.7
Diversity	19.5	2.8	18.4	2.0	21.3	3.2	19.4	2.7	17.8	2.6

TABLE 5.1.1 (Cont'd)

Means and Standard Deviations for the Sixth Grade Groups on the Variables Considered in the Study

Variable	Treatment 1		Treatment 2		Treatment 3		Treatment 4		Treatment 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Formality	18.3	2.6	19.6	1.6	16.4	2.9	18.7	2.3	17.4	3.3
Speed	15.9	2.8	17.3	2.3	16.9	2.0	17.3	2.9	18.4	3.0
Environment	18.0	2.7	19.0	1.2	18.8	2.8	18.0	2.9	16.9	2.1
Friction	18.7	2.9	19.0	2.9	18.3	2.6	19.1	3.2	17.7	2.2
Goal Direction	18.2	2.8	15.4	3.0	16.8	3.0	18.1	2.6	18.0	2.1
Favoritism	15.7	3.4	15.4	2.4	19.8	2.6	17.1	2.7	16.3	2.8
Difficulty	18.9	3.0	17.9	2.2	19.3	2.2	19.3	3.2	17.6	2.5
Apathy	17.2	2.4	16.9	2.6	17.0	2.9	16.4	2.9	15.4	2.9
Democratic	16.4	3.2	15.9	3.0	14.4	2.9	16.9	3.3	16.2	2.9
Cliqueness	17.9	2.5	16.7	2.6	19.8	2.0	18.1	2.3	17.7	3.3
Satisfaction	17.1	3.3	16.8	3.0	19.2	3.1	17.1	3.0	17.8	2.1
Disorganization	16.9	2.4	19.1	3.5	15.3	2.5	16.0	2.5	15.8	3.0
Competitiveness	17.5	2.5	18.3	3.3	16.9	3.1	18.7	2.8	18.2	2.4
Personality										
Dominance	20.2	5.1	18.6	5.3	18.3	5.4	20.3	5.3	20.2	3.9
Capacity for Status	11.6	3.8	11.8	4.0	12.0	3.2	11.7	3.5	12.3	2.7
Sociability	18.9	4.4	17.0	4.9	19.0	4.8	20.0	4.9	16.8	3.5
Social Presence	28.4	4.2	28.2	4.6	26.6	4.9	29.7	5.4	26.4	4.9
Self Acceptance	14.8	3.3	15.6	3.9	13.2	2.8	15.3	3.7	14.1	3.6
Sense of Well-Being	25.8	6.8	23.0	8.4	25.8	6.2	27.4	8.1	21.9	4.5

TABLE 5.1.1 (Cont'd)

Means and Standard Deviations for the Sixth Grade Groups on the Variables Considered in the Study

Variable	Treatment 1		Treatment 2		Treatment 3		Treatment 4		Treatment 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Achievement Via Conformity	20.3	4.9	19.6	4.9	23.0	5.6	20.2	5.1	18.1	3.6
Achievement Via Independence	14.8	4.2	16.8	3.6	14.4	4.5	16.1	4.2	15.0	4.0
Intellectual Efficiency	25.6	5.2	25.2	6.5	22.6	6.3	27.0	6.2	22.7	4.0
Psychological Mindedness	6.9	2.1	7.8	2.2	6.9	2.8	7.6	2.5	7.5	2.3
Flexibility	8.4	3.7	9.2	3.9	9.1	3.8	9.6	3.5	10.0	2.6
Creativity										
Fluency - Consequences	18.6	7.2	16.8	6.1	15.5	8.0	20.4	8.3	16.9	7.1
Flexibility - Consequences	17.0	6.1	15.2	5.3	16.9	6.9	18.7	7.5	16.0	6.4
Originality - Consequences	1.6	3.1	1.4	1.7	1.2	1.4	1.6	3.5	.6	.9
Fluency - Controlled Associations	22.5	9.3	20.6	8.5	25.9	12.8	23.5	10.1	25.8	10.9
Fluency - Word Beginnings	16.2	5.2	15.6	5.4	16.2	5.3	13.3	4.5	16.2	4.9
Fluency - Word Beginnings and Endings	16.3	5.2	16.6	6.4	15.7	5.5	15.1	4.5	20.0	6.1
Achievement Measures										
Composite	5.5	1.6	5.8	1.4	5.9	1.6	6.2	1.6	6.1	1.5
Reading	5.9	1.6	6.6	1.5	5.4	1.8	6.5	1.7	6.4	1.8
Language	5.4	1.7	5.9	1.6	4.3	1.9	6.1	1.6	5.9	1.7
Math	5.1	1.7	5.6	1.7	4.4	1.4	6.1	1.9	5.8	1.8

TABLE 5.1.1 (Cont'd)

Means and Standard Deviations for the Sixth Grade Groups on the Variables Considered in the Study

Variable	Treatment 1		Treatment 2		Treatment 3		Treatment 4		Treatment 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Intelligence										
STEA	5.7	1.6	5.8	1.9	5.9	1.4				
Otis Lennon							112.0	12.4	112.6	10.5

TABLE 5.1.2

Means and Standard Deviations for the Seventh Grade Groups on the Variables Considered in the Study

Variable	Treatment 1		Treatment 2		Treatment 3		Treatment 4		Treatment 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Study Habits and Attitudes										
Delay Avoidance	30.3	4.4	30.8	4.7	31.3	5.3	29.2	4.2	30.7	4.6
Work Methods	30.7	4.7	31.2	4.4	32.0	5.1	30.3	4.7	31.2	4.3
Study Habits	60.7	7.8	62.0	8.5	63.2	9.7	59.6	8.3	61.9	8.3
Teacher Approval	31.2	4.8	32.1	5.4	34.3	6.0	30.2	5.0	30.9	4.4
Educational Acceptance	31.6	4.4	32.0	4.7	34.4	5.4	31.0	4.3	32.3	4.5
Study Attitudes	62.8	8.6	64.1	9.8	69.0	11.1	61.2	8.8	63.2	8.3
Study Orientation	223.8	16.0	226.2	17.4	232.0	19.3	220.7	15.9	225.0	15.5
Self Concept Scale	171.6	21.9	173.7	25.2	180.2	20.7	175.0	18.2	176.3	17.7
Intellectual Achievement Responsibility Scale	22.1	4.8	22.5	5.5	24.2	4.7	23.3	4.3	21.8	5.3
General Anxiety	14.4	7.4	13.9	7.6	13.9	6.9	16.3	7.4	17.4	6.9
School Anxiety	19.2	4.7	18.0	4.6	17.3	3.7	19.6	4.5	18.4	4.9
Achievement Motivation	113.9	20.2	114.1	20.1	115.2	20.3	108.5	20.9	120.8	19.1
Learning Environment										
Cohesiveness	18.1	2.7	18.2	3.2	18.9	1.2	18.1	2.9	20.0	3.2
Diversity	19.4	3.1	18.6	3.0	19.1	1.4	18.9	2.4	18.2	3.3

TABLE 5.1.2 (Cont'd)

Means and Standard Deviations for the Seventh Grade Groups on the Variables Considered in the Study

Variable	Treatment 1		Treatment 2		Treatment 3		Treatment 4		Treatment 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Formality	17.3	2.8	17.0	3.1	18.6	2.3	17.5	2.8	19.0	4.1
Speed	16.2	2.6	17.1	3.6	16.5	2.3	16.6	2.6	17.9	2.4
Environment	17.3	3.0	17.7	3.4	17.6	2.1	17.1	2.8	16.7	4.6
Friction	17.7	3.0	17.3	4.0	19.1	2.2	19.0	3.0	17.2	3.8
Goal Direction	17.5	3.4	18.1	3.5	19.3	2.6	18.1	3.3	18.6	4.3
Favoritism	15.3	3.6	16.1	3.6	15.8	2.3	14.6	3.2	16.0	3.8
Difficulty	19.3	2.7	18.3	4.2	18.6	2.7	19.5	2.9	18.4	5.1
Apathy	16.7	3.5	16.0	3.1	16.7	2.2	16.8	2.9	15.9	3.5
Democratic	16.2	4.3	16.4	3.3	16.8	3.1	16.9	3.6	15.8	4.6
Cliqueness	16.8	2.6	17.0	3.2	17.0	1.6	17.2	2.9	18.3	2.7
Satisfaction	17.5	3.1	16.8	3.3	17.4	2.5	17.9	3.1	16.7	4.4
Disorganization	17.3	2.8	17.0	2.0	17.3	1.9	16.4	3.0	17.3	3.1
Competitiveness	17.1	2.7	17.6	2.5	17.2	1.9	17.8	3.4	16.4	2.5
Personality										
Dominance	20.1	6.8	20.0	5.7	22.2	5.2	20.9	5.2	20.0	5.9
Capacity for Status	11.9	3.9	11.9	3.1	12.8	3.4	12.1	2.3	12.1	3.5
Sociability	19.9	5.4	19.1	5.4	22.5	4.3	20.5	4.4	19.9	4.7
Social Presence	28.9	5.5	28.3	3.5	31.8	4.5	29.9	4.7	28.8	4.2
Self Acceptance	15.3	3.6	15.1	4.2	18.1	2.8	16.0	3.4	15.5	3.8
Sense of Well-Being	26.6	7.6	25.3	10.2	28.7	7.4	26.2	6.0	25.6	6.8

TABLE 5.1.2 (Cont'd)

Means and Standard Deviations for the Seventh Grade Groups on the Variables Considered in the Study

Variable	Treatment 1		Treatment 2		Treatment 3		Treatment 4		Treatment 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Achievement Via Conformity	20.7	5.4	19.1	5.1	20.1	4.8	20.9	4.2	20.2	4.6
Achievement Via Independence	16.2	3.3	15.4	3.4	16.6	3.7	16.0	3.5	15.9	4.1
Intellectual Efficiency	26.5	6.2	27.4	5.4	29.6	6.2	26.2	5.1	25.5	6.8
Psychological Mindedness	7.8	2.3	7.3	3.2	7.3	1.8	7.4	2.5	7.8	2.2
Flexibility	9.7	3.7	8.6	3.6	11.1	3.5	11.3	3.4	9.0	4.2
Creativity										
Fluency - Consequences	19.2	7.0	16.0	7.9	20.8	7.9	18.4	8.4	17.9	10.1
Flexibility - Consequences	16.6	5.7	13.9	6.6	18.9	6.8	17.5	8.7	16.6	8.8
Originality - Consequences	1.5	2.2	1.7	3.4	2.4	1.9	.7	1.1	.8	1.6
Fluency - Controlled Associations	24.7	10.8	31.0	15.0	28.9	13.1	24.0	11.4	23.2	12.9
Fluency - Word Beginnings	14.8	4.7	18.3	5.7	20.4	6.2	15.8	5.6	15.9	4.6
Fluency - Word Beginnings and Endings	15.8	5.4	18.6	5.5	19.7	5.6	18.4	6.0	16.1	5.7
Achievement Measures										
Composite	5.4	1.5	5.4	1.5	5.5	1.6	5.6	1.6	6.2	1.7
Reading	5.9	1.7	5.9	1.4	6.0	1.9	6.0	1.5	6.2	1.6
Language	5.4	1.5	5.2	1.8	5.5	1.6	5.5	1.7	5.6	1.5
Math	5.0	1.6	4.9	1.6	4.8	1.8	5.4	1.7	5.8	1.7

TABLE 5.1.2 (Cont'd)

Means and Standard Deviations for the Seventh Grade Groups on the Variables Considered in the Study

Variable	Treatment 1		Treatment 2		Treatment 3		Treatment 4		Treatment 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Intelligence										
STEA	5.4	1.6	5.3	1.5	5.6	1.7				
Otis Lennon							110.6	13.4	113.9	11.1

TABLE 5.1.3

Means and Standard Deviations for the Eighth Grade Groups on the Variables Considered in the Study

Variable	Treatment 1		Treatment 2		Treatment 3		Treatment 4		Treatment 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Study Habits and Attitudes										
Delay Avoidance	29.7	3.9	29.3	4.1	27.7	4.2	29.4	3.7	29.2	3.7
Work Methods	30.9	4.2	30.5	4.1	29.2	4.6	29.9	3.8	30.0	4.0
Study Habits	60.6	7.2	59.7	7.4	56.9	8.3	59.3	6.8	59.2	6.8
Teacher Approval	32.0	5.1	30.1	4.1	30.9	5.7	30.6	4.2	30.8	3.9
Educational Acceptance	32.2	4.7	31.0	4.2	30.2	5.0	31.1	4.1	30.4	3.5
Study Attitudes	64.2	9.3	61.5	7.7	61.1	10.2	61.6	7.9	61.2	6.8
Study Orientation	224.7	15.2	221.2	13.8	218.0	17.5	220.9	13.4	220.4	12.6
Self Concept Scale	177.9	19.7	176.8	25.5	173.1	20.6	175.9	19.8	167.1	25.9
Intellectual Achievement Responsibility Scale	21.7	4.7	21.8	5.1	23.9	4.7	22.8	4.5	19.0	6.0
General Anxiety	13.8	7.3	17.2	7.6	14.2	6.1	15.3	7.4	15.2	7.6
School Anxiety	18.3	4.1	19.8	4.3	16.7	4.1	18.8	4.5	18.6	5.9
Achievement Motivation	115.4	20.2	113.3	19.1	116.5	18.4	116.5	18.3	116.8	19.3
Learning Environment										
Cohesiveness	18.1	2.7	18.3	4.2	18.9	2.6	18.1	2.9	20.0	3.2
Diversity	19.4	3.1	17.9	2.9	20.1	2.2	18.9	2.4	18.2	3.3

TABLE 5.1.3 (Cont'd)

Means and Standard Deviations for the Eighth Grade Groups on the Variables Considered in the Study

Variable	Treatment 1		Treatment 2		Treatment 3		Treatment 4		Treatment 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Formality	17.3	2.8	17.3	3.5	18.3	1.9	17.5	2.8	19.0	4.1
Speed	16.2	2.6	17.3	4.2	16.0	2.0	16.6	2.6	17.9	2.4
Environment	17.3	3.0	17.8	3.0	18.0	2.3	17.1	2.8	16.7	4.6
Friction	17.7	3.0	16.7	3.8	18.4	2.2	19.0	3.0	17.2	3.8
Goal Direction	17.5	3.4	18.2	3.8	18.3	2.1	18.1	3.3	18.6	4.3
Favoritism	15.3	3.6	15.5	3.2	16.0	2.6	14.6	3.2	16.0	3.8
Difficulty	19.3	2.7	16.9	2.3	18.9	2.8	19.5	2.9	18.4	5.1
Apathy	16.7	3.5	16.0	3.6	16.1	2.5	16.8	2.9	15.9	3.5
Democratic	16.2	4.3	16.3	3.8	17.5	2.7	16.9	3.6	15.8	4.6
Cliqueness	16.8	2.6	16.6	3.4	16.9	3.7	17.2	2.9	18.3	2.7
Satisfaction	17.5	3.1	17.5	2.3	18.7	2.4	17.9	3.1	16.7	4.4
Disorganization	17.3	2.8	17.1	3.1	17.0	2.4	16.4	3.0	17.3	3.1
Competitiveness	17.1	2.7	15.8	4.5	16.6	1.8	17.8	3.4	16.4	2.5
Personality										
Dominance	21.1	4.9	18.9	5.5	21.1	5.0	20.7	6.1	19.5	5.7
Capacity for Status	11.9	3.6	12.5	2.8	13.5	3.8	11.5	3.7	11.4	3.5
Sociability	21.1	4.8	18.8	3.6	20.0	5.5	20.0	5.1	18.4	4.6
Social Presence	30.1	5.1	27.9	4.9	29.1	4.6	29.3	4.8	26.4	5.0
Self Acceptance	16.7	3.3	14.7	3.2	16.0	4.0	16.7	3.9	14.8	3.6
Sense of Well-Being	28.3	6.1	21.8	7.6	28.9	4.5	26.9	5.9	23.4	8.0

TABLE 5.1.3 (Cont'd)

Means and Standard Deviations for the Eighth Grade Groups on the Variables Considered in the Study

Variable	Treatment 1		Treatment 2		Treatment 3		Treatment 4		Treatment 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Achievement Via Conformity	21.3	5.0	16.5	4.3	20.5	4.4	19.2	4.3	18.5	5.2
Achievement Via Independence	17.2	4.3	15.5	4.3	17.7	3.5	15.9	3.8	15.0	3.3
Intellectual Efficiency	28.1	5.4	23.2	5.9	28.1	4.5	25.8	5.5	24.0	6.0
Psychological Mindedness	7.2	2.3	8.5	2.0	7.9	2.6	7.2	2.3	7.0	2.4
Flexibility	10.5	3.7	9.9	3.6	12.4	3.3	10.5	3.2	9.6	3.8
Creativity										
Fluency - Consequences	19.2	7.0	16.0	7.9	20.8	7.9	18.4	8.4	17.9	10.1
Flexibility - Consequences	16.6	5.7	13.9	6.6	18.9	6.8	17.5	8.7	16.6	8.8
Originality - Consequences	1.5	2.2	1.7	3.4	2.4	1.9	.7	1.1	.8	1.6
Fluency - Controlled Associations	24.7	10.8	31.0	15.0	28.9	13.1	24.0	11.4	23.2	12.9
Fluency - Word Beginnings	14.8	4.7	18.3	5.7	20.4	6.2	15.8	5.6	15.9	4.6
Fluency - Word Beginnings and Endings	15.8	5.4	18.6	5.5	19.6	5.6	18.4	6.0	16.1	5.7
Achievement Measures										
Composite	5.9	1.7	5.3	1.5	5.7	1.7	6.0	1.9	6.5	1.8
Reading	6.0	1.7	5.7	1.7	6.0	1.5	5.8	1.8	6.3	1.9
Language	5.9	1.8	5.3	1.8	5.7	1.7	5.9	2.0	5.9	2.0
Math	5.6	1.8	4.8	1.5	5.1	2.0	6.0	1.9	6.2	1.9

TABLE 5.1.3 (Cont'd)

Means and Standard Deviations for the Eighth Grade Groups on the Variables Considered in the Study

Variable	Treatment 1		Treatment 2		Treatment 3		Treatment 4		Treatment 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Intelligence										
STEA	5.7	1.7	5.4	1.5	6.0	1.6				
Otis Lennon									114.1	9.3
Lorge Thorndike							116.0	15.0		

TABLE 5.2.1

Analysis of Variance of
Achievement Test Scores

Source	df	SS	F
Grades	2	3.06	.57
Treatment	4	118.41	10.92*
Grade x Treatment	8	154.34	7.12*
Error	1187	3217.80	

* $p < .01$

in treatment 5 had substantially higher scores on the criterion than did students in treatments 1 and 3.

Since students in treatment 5 had been given a different standardized achievement test than those in the other four schools, it was felt that the differences observed here might have been caused by differences in the nature of the respective tests taken by the students. Therefore, separate analyses were run with treatment 5 excluded. The result was that the treatment main effect vanished. There were no differences between the treatment groups at the seventh grade level. Students in instructional treatment 4 scored significantly better than treatment 2 students at the eighth grade level and treatment 4 students performed better than students in treatment 1 and treatment 3 at the sixth grade level. Further discussion of these results is provided in Chapter Six.

5.3 Analysis of the Study Orientation Test Scores

A two-way analysis of variance, similar to that reported in the last section, was conducted on the scores from the Study Orientation (SO) scale - one of seven included in the Survey of Study Habits and Attitudes. The SO score represents an overall measure of study habits and attitudes. It was thought to be the best measure of those constructs presently available.

The results of the analysis are presented in Table 5.3.1. The main effect due to grades and the grade x instructional

TABLE 5.3.1

Analysis of Variance of
Study Orientation Scores

Source	df	SS	F
Grades	2	4969.82	7.53*
Treatment	4	954.03	.72
Grades x Treatment	8	8779.17	3.33*
Error	1187	391536.33	

* $p < .01$

treatment interaction effect were significant. Profiles of the simple effects of the schools for each of the grades are presented in Figure 5.3.1. Simple effects analysis revealed differences between the schools at the sixth and seventh grade levels. Comparisons between means within grades, using the Newman-Keuls procedure showed the following results: For the sixth grades, treatment 4 students had better study habits and attitudes toward learning than did students in treatments 3 and 5. For the seventh grade the trend was somewhat reversed with the students in treatment 3 having higher scores on the criterion than students in treatment 4 and treatment 1. Speculation on why differences were not observed in the eighth grade are presented in the next chapter.

5.4 Analysis of the Learning Environment Inventory Scores

If the programs in this study differ in the effects they have on students, then responses on an instrument designed for assessing the climate of schools should manifest those differences. To investigate this possibility, a multivariate analysis of variance was conducted on the students' responses to the fifteen scales of the Learning Environment Inventory.

There was little reason to expect differences between the seventh and eighth graders on this instrument. Few distinctions could be made in the learning environments fostered by the various treatments for these grades. The sixth grade programs were organized somewhat differently in a few of the schools. One example was the team teaching

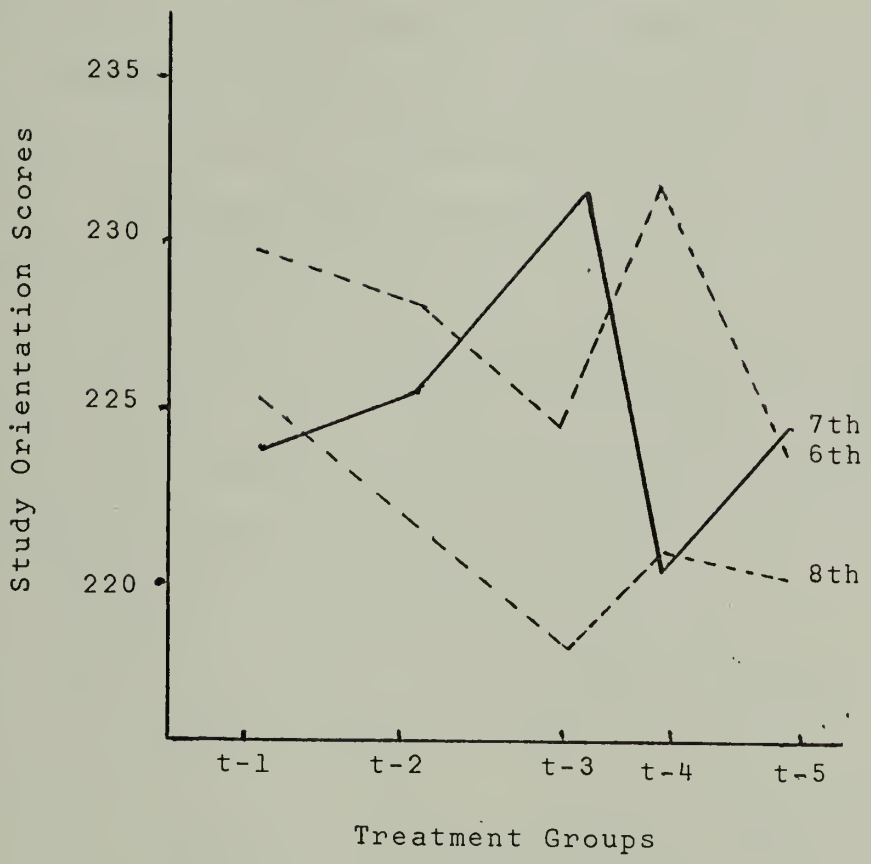


Figure 5.3.1 Profiles of Simple Effects for Schools as a Function of Scores on the Study Orientation Scale of the SSHA.

approach in instructional treatment 4. The responses of the seventh and eighth graders in each school were, therefore, combined for this analysis to produce a 2 x 5 design in which one level of the first factor represented the sixth grade and the second level represented the seventh and eighth grades.

Parameters of the models (4.6.3) were tested against zero by constructing standard hypotheses of the type given in 4.6.4. The significance tests of those hypotheses are summarized in Table 5.4.1.

Both the main effects for grade and treatment were significant at the .05 level. To identify the source of the differences, discriminant functions (V) associated with each significant latent root were computed. The function for the main effect due to grades was as follows:

$$V_G = .289X_1 - .202 + X_2 + .223X_3 - .201X_4 + .318X_5 + .270X_6 + .019X_7 + .411X_8 - .473X_9 - .092X_{10} - .137X_{11} + .481X_{12} - .188X_{13} - .610X_{14} + .337X_{15}.$$

The relative magnitudes of the standardized discriminant function coefficients can be compared to determine which variables contribute most to the definition of the composite function. In general, the differences in grades can be attributed to the contrast between the scores on X_8 (favoritism) and X_{12} (cliqueness) and the scores on X_9 (difficulty) and X_{14} (disorganization). The numerical value of the function tended to be positive for Grade 6 and negative for Grades 7 and 8. Sixth grade students were likely to have

TABLE 5.4.1

Multivariate Analysis of Variance of the Fifteen Scales of the Learning Environment Inventory

Source	df	F
Grade	15/460	2.84*
Treatment	60/1798	2.52*
Grade x Treatment	60/1798	1.35

* $p < .05$

lower self concepts and were more inclined to seek group support. Seventh and eighth graders perceived their assignments as being more difficult and their classes less ordered or organized.

The analysis of the main effect for treatment resulted in the identification of two significant dimensions of differences among the five groups. The first was

$$V_t^{(1)} = .277X_1 - .409X_2 - .243X_3 + .777X_4 + .218X_5 - .175X_6 + .271X_7 + .171X_8 - .445X_9 - .365X_{10} - .018X_{11} + .152X_{12} - .119X_{13} - .156X_{14} + .058X_{15}.$$

Inspection of the first function indicates that differences between treatments are largely due to X_4 (speed). Results also suggest that X_2 (diversity), X_9 (difficulty), and X_{10} (apathy) contribute to discrimination between the schools, independently of the effect of X_4.

On close examination, it appears that the function relates to instruction - course requirements, classroom activities, assignments, etc. At one end of the continuum would be those treatments where students perceived the manner in which material was presented the class as being hurried and the rate of progress through the curriculum as rushed. That is, proceeding at a pace generally felt to be unfavorable by students. Such treatments were further

characterized as having few provisions for individual differences in interests and learning styles. Students in such treatments perceived their class assignments as being more difficult than students in other groups perceived theirs. On the whole, students in such treatments were also apathetic about class activities.

The second discriminant function,

$$V_t^{(2)} = .542X_1 + .079X_2 - .082X_3 - .069X_4 - .821X_5 - .462X_6 - .532X_7 + .532X_8 + .163X_9 + .199X_{10} - .502X_{11} + .563X_{12} - .120X_{13} + .346X_{14} - .049X_{15},$$

represents a contrast between scores on X_1 (cchesiveness), X_{12} (cliqueness), X_{14} (disorganization) and scores on X_5 (environment), X_6 (friction), X_7 (goal direction) and X_{11} (democratic). The coordinates as a group, seem to be concerned with intragroup/interpersonal accord. Tension and disagreement among students would be at a minimum in treatments high on this dimension. A general consensus on class objectives and goals would be evident. Cooperation among members of the group would also be a distinguishable feature. The negative loading for environment implied that students in such groups tend to perceive available recreational equipment as being inadequate. The prediction was that the more traditional schools would score highest on this dimension.

To further clarify the relationships between the five treatments, group means on the two discriminant functions

discussed above were computed by applying the two sets of discriminant function weights to the group means on each of the original variables (presented in Table 5.4.2). The results are given in Table 5.4.3.

The mean values for $V_t^{(1)}$ and $V_t^{(2)}$ taken as Cartesian coordinates can be used to plot the locations of the five treatment groups in a two-dimensional space. These results are presented in Figure 5.4.1. It is evident that treatment 5 differs substantially from all others and occupies an extreme position on both dimensions. Treatment 2 lies close to treatments 3 and 4 on the axis of the second discriminant function, but tends to separate from treatments 1, 3 and 4 along the axis of the first discriminant function. We might conclude that the primary dimension of separation among treatment groups represents the difference between treatments 2 and 5 and the remaining three groups. The secondary dimension of group difference discriminates between treatment 5 and treatments 1, 2, 3, and 4.

Summary

The results of these analyses substantiated the expectation that the learning environment created by the most traditional program, instructional treatment 5, was significantly different from that of the other instructional treatments. It is interesting to note that the responses of treatment 4 students did not correspond more with treatment 5 students. The fact that they did not is interpreted

TABLE 5.4.2

Summary of Mean Scores on the LEI Scales for Students in Each of the Five Instructional Programs Across the Sixth, Seventh, and Eighth Grades[1]

Variable	Treatment				
	1	2	3	4	5
Cohesiveness	18.7	18.9	18.3	18.4	19.0
Diversity	19.7	18.8	19.5	18.7	18.2
Formality	17.6	17.4	18.6	17.4	17.8
Speed	15.5	17.2	16.9	16.9	17.9
Environment	17.5	18.4	17.8	17.4	17.0
Friction	18.1	17.3	19.1	18.5	17.4
Goal Direction	17.9	18.7	18.2	17.7	17.9
Favoritism	15.3	15.8	16.3	15.6	15.9
Difficulty	19.4	17.9	19.3	19.2	18.0
Apathy	16.7	16.6	16.4	15.9	16.0
Democratic	16.3	16.2	16.7	17.0	15.7
Cliqueness	17.1	16.9	17.2	17.2	17.9
Satisfaction	17.3	16.6	18.2	17.6	16.8
Disorganization	17.0	16.8	16.6	16.0	16.5
Competitiveness	17.1	17.1	17.1	17.8	17.0

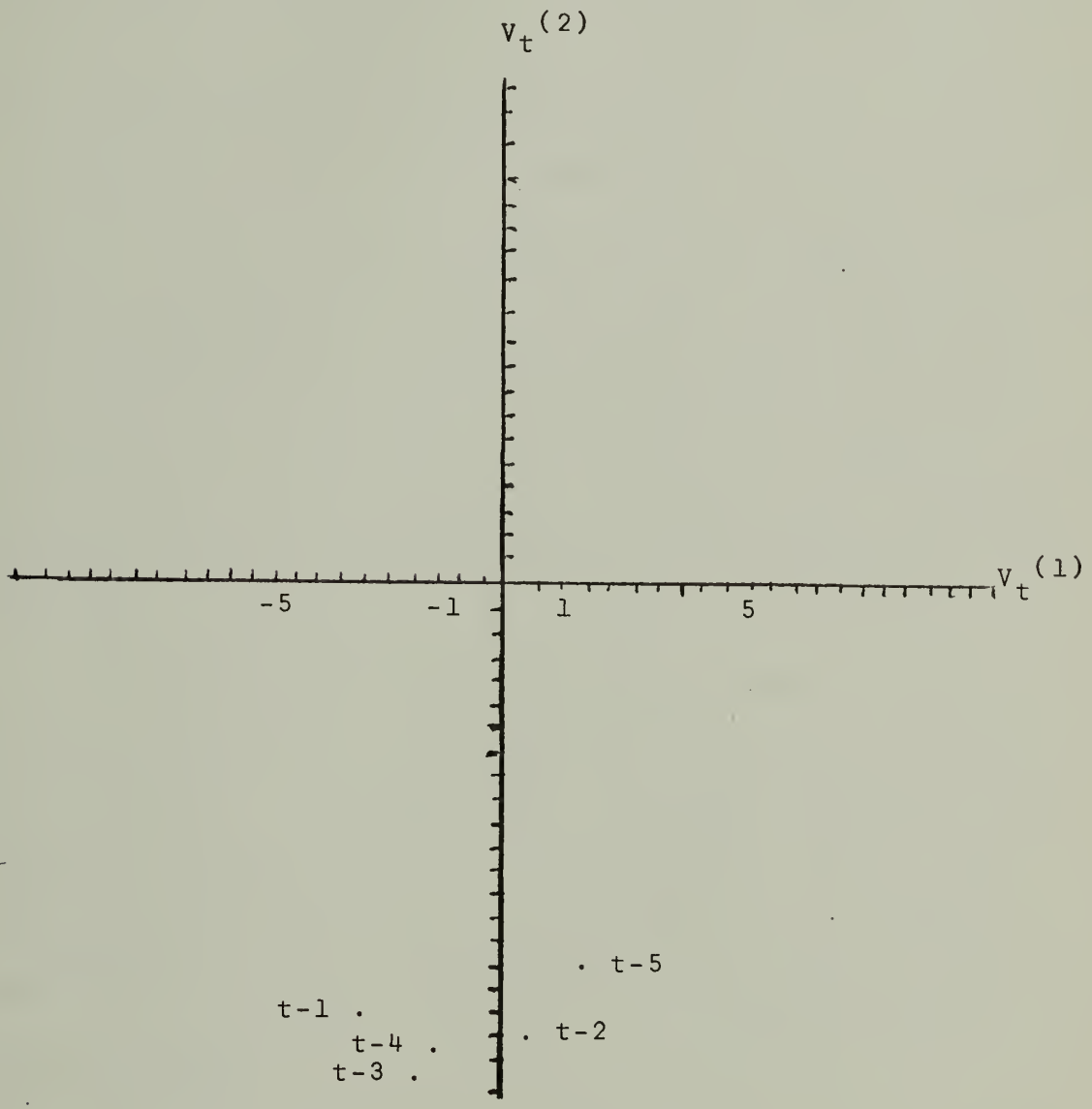
[1] $N_1=165$, $N_2=85$, $N_3=52$, $N_4=123$, $N_5=161$

TABLE 5.4.3

Summary of Mean LEI Discriminant Function Scores
for Students in the Five Instructional Programs
Across the Sixth, Seventh, and Eighth Grades

Discriminant Function	Treatment				
	1	2	3	4	5
$V_t^{(1)}$	-2.53	+ .08	- 2.08	1.23	1.63
$V_t^{(2)}$	-8.67	-9.79	-10.30	-9.80	-7.73

Figure 5.4.1 Graphic Representation of Instructional Treatment Groups Using Mean LEI Scores on the Discriminant Functions as Coordinate Values.



as evidence that a well-controlled and disciplined school with a broad and progressive program can be perceived as being just as open, unconfining, and satisfying as one which almost places no restrictions on students.

5.5 Analysis of the Creativity Scores

For the analysis of the creativity scores the dependent variates were the three Consequences Test subscores, the Controlled Associations Test score, the Word Beginnings and the Word Beginnings and Endings Test Scores. The linear model of concern, was given in (4.6.3). The summary of the significance tests is presented in Table 5.5.1. Both main effects and the interaction proved to be significant at the .01 level.

Obviously, the results were not as clear as would be desired for explicit interpretation. Subsequent analyses were, therefore, concerned with the simple effects of the treatments within each level of the grades factor. The source of the differences were investigated by computing discriminant functions associated with significant roots for each grade.

Multiple Discrimination Analyses for the Sixth Grade Groups

The majority of the total discriminable variance for the sixth grade can be accounted for by two discriminant functions. The first was identified as

TABLE 5.5.1

Multivariate Analysis of Variance of the Creativity Scores

Source	df	F
Grade	12/2694	12.86*
Treatment	24/4700	16.40*
Grade x Treatment	48/6632	9.39*

* $p < .001$

$$V_6^{(1)} = +.258X_1 + .514X_2 + .194X_3 - .014X_4 - .520X_5 - .542X_6.$$

It represents a contrast between the first two variables - ideational fluency (X_1) and spontaneous flexibility (X_2) and the last two variables (X_5 and X_6) - both of which were measures of word fluency. The additional discriminability provided by the second function was due to the contrast of ideational fluency (X_1), originality (X_3), and word fluency (X_5) with spontaneous flexibility (X_2), associational fluency (X_4) and word fluency (X_6).

Treatments with high scores on the first discriminant function are probably composed of students who are more productive in unstructured situations. They are fluent producers of ideas and are able to rapidly think of many approaches to a given problem. High scores on the word fluency tests are dependent upon the ability to rapidly produce words in response to some specification having to do with letter composition. Since the signs for these variables are negative, it is likely that high scorers on the first discriminant function are not particularly productive when the problem presented is limited by superficial restrictions.

The second discriminant function for the sixth grade groups was

$$V_6^{(2)} = -1.325X_1 + 1.238X_2 - .490X_3 + .693X_4 - .842X_5 + .623X_6.$$

Groups with high scores on this function included students who were able to identify different approaches to problems and to recall ideas or words related to some given thing in a specified way. It is probable, however, that they would not generate many responses and the responses produced would tend to lack uniqueness. The difference here may center on the ability to readily recall appropriate material or knowledge as opposed to the ability to freely generate new approaches or methods of relating to or dealing with a problem.

A comparison of group means for each treatment group on the two discriminant functions revealed the following ordering. Treatment 4 had the highest mean score (-.83) on the first function. Treatment 1 had the next highest score (-4.33) followed closely by treatment 3 (-4.30) and treatment 2 (-5.50). Treatment 5 had the lowest mean (-7.56). A ranking of group means on the second discriminant function resulted in treatment 5 having the highest score (13.8). Treatment 3 (9.94), treatment 4 (9.84), treatment 1 (9.52) and treatment 2 (7.33) followed in that order.

The conclusion drawn here is that the sixth grade groups can be differentiated and that students in treatment 5 stand apart from students in the other groups on both relevant dimensions. On the first function, treatment 5 students

scored significantly lower than students in the other groups. Treatment 4 students also tend to differ from students in the other groups on the variables associated with the second discriminant function. The difference is in the direction opposite to that of students in treatment 5. The results of ranking on the second discriminant function showed that treatment 5 students scored well above the other groups in the study.

Multiple Discriminant Analyses for the Seventh Grade Groups

The differences between seventh grade groups can be explained in terms of two discriminant functions. The first was largely characterized by the contrast between spontaneous flexibility (X_2) and ideational fluency (X_1). It was identified as

$$V_7^{(1)} = -1.8X_1 + 2.4X_2 - .54X_3 - .54X_3 - .54X_4 - .35X_5 + .52X_6.$$

The ability to produce a number of categories or classes of uses for a given object is thought to be essential to creative productivity. In cases where one approach to a problem proves unprofitable, the flexible individual will readily shift his focus and attack the problem from a different perspective. The less flexible individual would be inclined to persist with the more obvious approaches. The reason that such an ability would be associated with

high performance on the Word Beginnings test should be evident. The production of words having certain specifications would require many shifts in focus for high performance.

Ideational fluency, to a large extent, and originality and associational fluency, to a lesser extent, all are represented by negative coefficients. This implies that students in groups with high scores on this function are not likely to be the most fluent producers of ideas. Nor are their ideas likely to be noted for their originality.

A ranking of groups in terms of mean scores on the first discriminant function resulted in the following order. Treatment 4 had the highest score (-.34) followed by treatment 5 (-1.59), treatment 1 (-5.83), treatment 3 (-5.88) and treatment 2 (-9.75). It might be concluded, then, that treatments 4 and 5 had significantly different effects on seventh grade students as compared to that of the other treatments.

The proportion of the total discriminable variance accounted for by the second discriminant function is mainly attributable to the contrast between ideational fluency (X_1) and spontaneous flexibility (X_2) with word fluency (X_5) also represented with a moderately large coefficient. The function

$$V_7^{(2)} = 1.8X_1 - 1.4X_2 + .06X_3 - .29X_4 - .66X_5 - .29X_6$$

is almost characterized by the same association of variables as $V_7^{(1)}$ except that the signs are different. Students in groups scoring high on this function tend to perform well on tests measuring ideational fluency but low on tests of spontaneous flexibility and word fluency. A comparison of group means on the function revealed that treatment 1 students ($\bar{V}$ = -10.11) are significantly higher than the other treatment groups and that students in treatment 5 ($\bar{V}$ = -13.09) and treatment 4 ($\bar{V}$ = -13.96) tend to score higher than students in treatment 3 ($\bar{V}$ = -16.52) and treatment 2 ($\bar{V}$ = -17.16).

Multiple Discriminant Analysis for the Eighth Grade Groups

The discriminant functions obtained for the eighth grade groups are almost exact replications of the two reported for the seventh grade groups. The first was as follows:

$$V_8^{(1)} = -1.3X_1 + 1.4X_2 - .48X_3 - .44X_4 + .09X_5 + .96X_6.$$

The major contrast represented by the function is between spontaneous flexibility (X_2) and ideational fluency (X_1). Word fluency (X_6) with a positive coefficient is also a factor here. Originality (X_3) and associational fluency (X_4) are represented by negative coefficients as was the case with $V_7^{(1)}$.

The contrast between spontaneous flexibility and ideational fluency is the outstanding feature of the second discriminant function for the eighth grade groups. That function was

$$V_8^{(2)} = -1.9X_1 + 2.8X_2 - .80X_3 - .10X_4 - .45X_5 - .16X_6.$$

Originality (X_3) and word fluency (X_5) also contribute to the discriminability of the function with negative coefficients. Therefore, students in groups scoring high on $V_8^{(2)}$ are characterized by high scores on flexibility and a tendency to score low on measures of originality and word fluency.

An examination of group means on the first discriminant function revealed that treatment 5 students (20.09) were much higher on this dimension than students in the other treatment groups (treatment 4: 8.51; treatment 3: 7.77; treatment 2: 6.67; and treatment 1: 2.9). Students in treatments 4, 3, and 2 appear not to differ greatly. Students in treatment 1 tend to score considerably below students in the other four groups.

The major group differences on $V_8^{(2)}$ were between students in treatment 4 (-.1), treatment 1 (-1.3), treatment 5 (-2.9) and students in treatment 3 (-4.4) and treatment 2 (-5.6). At the eighth grade level, instructional treatments 1, 4, and 5 facilitate flexibility at the expense of fluency and originality.

Summary

In summarizing the results of this section, it appears that treatments 4 and 5 differ substantially from the other treatments at each grade level. In general, these differences occur on measures of spontaneous flexibility and ideational fluency.

Specifically, the analyses showed that the differences between treatments at the seventh and eighth grade levels were very similar. In the seventh grade, treatments 4 and 5 had higher means than treatments 1, 2, and 3 on a function mainly characterized by the contrast between spontaneous flexibility and ideational fluency. At the eighth grade level virtually the same function again appeared as the primary dimension of difference among treatment groups. Treatment 5 was substantially higher than the remaining treatments but was followed by treatment 4 in magnitude.

The primary discriminant function identified for the sixth grade treatment groups differed from those obtained for the seventh and eighth grade groups. This could be explained by the fact that the programs at the sixth grade level were considerably different from those offered seventh and eighth grade students. The sixth grade group with the highest mean score on the first function (treatment 4) was distinguished from the others by its team teaching approach to instruction.

The pervasiveness of the effects of the contrast between spontaneous flexibility and ideational fluency is demonstrated by the fact that the second dimension of difference between the sixth grade programs was virtually the same as those identified as having primacy for the other grades. Even though the programs differ between grades there appear to be enough carry-over from the seventh and eighth grades to influence the sixth grade programs in instructional treatments 1, 2, and 3.

5.6 ATI Analyses

The ATI studies were initially conducted with composite achievement scores as the criterion and twenty-two carefully chosen variables as predictors. The greatest expected differences involved contrasts between treatments 4 or 5 and treatments 1, 2, and 3. Since instructional programs 1 and 2 were also similar in most respects, treatment 1 was not considered in the ATI phase of the study. Too many investigators have selected weakly conceptualized or undifferentiated treatment conditions. The decision to focus in on the most obviously distinct experimental treatments was done to eliminate such superfluity.

The total number of potential aptitude variables available for investigation in this study were indeed large. Had ATI analyses been conducted with all of them the probability that chance results alone could have accounted for significant results would have been high. For this reason, the following procedures were adopted:

1) The criterion and aptitude variables selected for investigation were limited to those having theoretical importance. That is, only variables which seemed promising on the basis of prior research or some underlying theory were even considered for analyses.
2) Interactions which held up across grades or different scales measuring similar constructs were retained for discussion and given additional emphasis.

3) Greater weight was placed on variables which resulted in the rejection of hypotheses of no difference with alpha rates of .01 or better.

The goals of at least three of the programs in the study included enhanced student attitudes toward school and learning. To investigate the possibility of differential effects, analyses were repeated with the Study Orientation scores as the criterion. A subset of the original predictor variables, excluding the SSHA test score served as aptitudes.

The analysis of covariance method of Gulliksen and Wilks (1950), described in Chapter 4 was used to search for interactions. The data from different grades were analyzed separately. A summary of the results are reported in Tables 5.6.1 to 5.6.6. The F statistic in the second column of each table is used to test the parallelism of slopes from regressing the criterion variables on each of the predictor variables for the various instructional treatment groups. Three significant F's were observed for the sixth grade groups (see Tables 5.6.1 and 5.6.2). The test of the parallelism of slopes was rejected when the composite achievement test score was the criterion and Dominance, Capacity for Status, and Originality (from the Consequences Test) were the aptitudes. Dominance, significant at the .01 level was probably the most stable of the aptitudes for the sixth grade groups. The other two were significant at the .05 level. Obviously, none of the F statistics

TABLE 5.6.1

Summary of Analysis of Covariance Results for the Sixth Grade Students in the Four Instructional Treatments with Composite Achievement Test Scores as the Criterion

Aptitude	Hypothesis					
	Homogeneity of Variance (X^2 Statistic)	Degrees of Freedom	Equal Slopes of Regression (F Statistic)	Degrees of Freedom	Equal Intercepts of Regression (F Statistic)	Degrees of Freedom
Study Habits and Attitudes						
Teacher Approval	1.74	3	2.50	3/419	9.22**	3/424
Educational Acceptance	1.66	3	2.01	3/419	9.87**	3/424
Self Concept Scale	.32	3	.95	3/393	5.98**	3/398
Intellectual Achievement Responsibility Scale	.70	3	.83	3/425	8.88**	3/430
General Anxiety	.32	3	2.57	3/400	7.41**	3/405
School Anxiety	.35	3	.16	3/400	7.33**	3/405
Personality						
Dominance	3.64	3	3.79*	3/214	2.25	3/219
Capacity for Status	3.82	3	2.68*	3/214	2.21	3/219
Sociability	3.00	3	1.98	3/214	2.73*	3/219
Self Acceptance	3.17	3	2.00	3/214	2.22	3/219
Sense of Well-Being	2.78	3	.56	3/214	3.42*	3/219
Achievement Via Conformity	2.93	3	.11	3/214	3.63*	3/219

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

TABLE 5.6.1 (Cont'd)

Summary of Analysis of Covariance Results for the Sixth Grade Students in the Four Instructional Treatments with Composite Achievement Test Scores as the Criterion

Aptitude	Hypothesis					
	Homogeneity of Variance (X^2 Statistic)	Degrees of Freedom	Equal Slopes of Regression (F Statistic)	Degrees of Freedom	Equal Intercepts of Regression (F Statistic)	Degrees of Freedom
Achievement Via Independence	3.01	3	.87	3/214	2.11	3/219
Intellectual Efficiency	1.67	3	1.13	3/214	3.75*	3/219
Flexibility	3.12	3	.37	3/214	2.94*	3/219
Achievement Motivation	3.66	3	2.10	3/419	3.50*	3/424
Creativity						
Fluency - Word Beginnings	1.54	3	.79	3/166	.19	3/171
Fluency - Word Beginnings and Endings	.58	3	1.72	3/418	6.53**	3/423
Fluency - Consequences	3.79	3	1.71	3/439	6.43**	3/444
Flexibility - Consequences	3.27	3	1.25	3/438	6.06**	3/443
Originality - Consequences	3.44	3	2.69*	3/265	3.71*	3/270
Fluency - Controlled Associations	.95	3	.61	3/438	5.27**	3/443

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

TABLE 5.6.2

Summary of Analysis of Covariance Results for the Seventh Grade Students in the Four Instructional Treatments with Composite Achievement Test Scores as the Criterion

Aptitude	Hypothesis					
	Homogeneity of Variance (X^2 Statistic)	Degrees of Freedom	Equal Slopes of Regression (F Statistic)	Degrees of Freedom	Equal Intercepts of Regression (F Statistic)	Degrees of Freedom
Study Habits and Attitudes						
Teacher Approval	2.04	3	.87	3/310	4.48**	3/315
Educational Acceptance	2.81	3	.05	3/310	3.94**	3/315
Self Concept Scale	.59	3	.18	3/342	5.17**	3/347
Intellectual Achievement Responsibility Scale	1.32	3	3.04*	3/337	4.86**	3/342
General Anxiety	.07	3	.24	3/313	4.58**	3/318
School Anxiety	.05	3	.10	3/316	4.77**	3/321
Personality						
Dominance	7.44	3	.64	3/165	2.73*	3/170
Capacity for Status	5.86	3	.25	3/165	2.68*	3/170
Sociability	5.73	3	1.05	3/165	3.24*	3/170
Self Acceptance	7.34	3	.57	3/165	3.28*	3/170
Sense of Well-Being	5.85	3	2.29	3/165	3.14*	3/170
Achievement Via Conformity	5.57	3	1.17	3/165	2.32	3/170

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

TABLE 5.6.2 (Cont'd)

Summary of Analysis of Covariance Results for the Seventh Grade Students in the Four Instructional Treatments with Composite Achievement Test Scores as the Criterion

Aptitude	Hypothesis					
	Homogeneity of Variance (X^2 Statistic)	Degrees of Freedom	Equal Slopes of Regression (F Statistic)	Degrees of Freedom	Equal Intercepts of Regression (F Statistic)	Degrees of Freedom
Achievement Via Independence	4.30	3	2.29	3/165	3.17*	3/170
Intellectual Efficiency	6.09	3	2.38	3/165	5.55**	3/170
Flexibility	7.20	3	3.15*	3/165	3.55*	3/170
Achievement Motivation	7.04	3	.60	3/310	.95	3/315
Creativity						
Fluency - Word Beginnings	.68	3	.56	3/118	2.55	3/123
Fluency - Word Beginnings and Endings	.82	3	2.24	3/324	4.19**	3/329
Fluency - Consequences	2.64	3	1.76	3/322	4.34**	3/327
Flexibility - Consequences	2.87	3	1.91	3/322	4.15**	3/327
Originality - Consequences	.62	3	.49	3/151	5.93**	3/156
Fluency - Controlled Associations	.86	3	.20	3/312	5.32**	3/317

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

TABLE 5.6.3

Summary of Analysis of Covariance Results for the Eighth Grade Students in the Four Instructional Treatments with Composite Achievement Test Scores as the Criterion

Aptitude	Hypothesis					
	Homogeneity of Variance (X^2 Statistic)	Degrees of Freedom	Equal Slopes of Regression (F Statistic)	Degrees of Freedom	Equal Intercepts of Regression (F Statistic)	Degrees of Freedom
Study Habits and Attitudes						
Teacher Approval	3.12	3	.75	3/290	5.55**	3/295
Educational Acceptance	2.87	3	.57	3/290	5.95**	3/295
Self Concept Scale	7.55	3	.61	3/334	6.59**	3/339
Intellectual Achievement Responsibility Scale	3.51	3	.05	3/311	6.92**	3/316
General Anxiety	2.68	3	.43	3/297	4.29**	3/302
School Anxiety	2.78	3	1.51	3/300	4.76**	3/305
Personality						
Dominance	6.20	3	2.41	3/154	2.59	3/159
Capacity for Status	5.71	3	2.30	3/154	3.13*	3/159
Sociability	5.96	3	.03	3/154	2.76*	3/159
Self Acceptance	5.76	3	.12	3/154	2.50	3/159
Sense of Well-Being	5.26	3	.48	3/154	2.30	3/159
Achievement Via Conformity	6.44	3	.53	3/154	2.40	3/159

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

TABLE 5.6.3 (Cont'd)

Summary of Analysis of Covariance Results for the Eighth Grade Students in the Four Instructional Treatments with Composite Achievement Test Scores as the Criterion

Aptitude	Hypothesis					
	Homogeneity of Variance (X^2 Statistic)	Degrees of Freedom	Equal Slopes of Regression (F Statistic)	Degrees of Freedom	Equal Intercepts of Regression (F Statistic)	Degrees of Freedom
Achievement Via Independence	5.53	3	.18	3/154	3.13*	3/159
Intellectual Efficiency	4.72	3	.39	3/154	2.79*	3/159
Flexibility	6.20	3	.64	3/152	2.38	3/157
Achievement Motivation	6.40	3	.40	3/290	2.35	3/295
Creativity						
Fluency - Word Beginnings	14.24**	3	.26	3/125	4.30**	3/130
Fluency - Word Beginnings and Endings	8.19*	3	.44	3/307	5.08**	3/312
Fluency - Consequences	4.32	3	1.10	3/276	2.48	3/281
Flexibility - Consequences	4.65	3	.95	3/276	2.23	3/281
Originality - Consequences	4.12	3	.46	3/135	8.58**	3/140
Fluency - Controlled Associations	6.10	3	.23	3/269	5.67**	3/274

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

TABLE 5.6.4

Summary of Analysis of Covariance Results for the Sixth Grade Students in the Four Instructional Treatments with Study Orientation Subscale Scores as the Criterion

Aptitude	Hypothesis					
	Homogeneity of Variance (X^2 Statistic)	Degrees of Freedom	Equal Slopes of Regression (F Statistic)	Degrees of Freedom	Equal Intercepts of Regression (F Statistic)	Degrees of Freedom
Self Concept Scale	4.00	3	2.54	3/493	4.17**	3/498
Intellectual Achievement Responsibility Scale	6.66	3	.60	3/534	2.02	3/539
General Anxiety	5.58	3	.65	3/491	1.59	3/496
School Anxiety	5.72	3	.74	3/492	1.76	3/497
Personality						
Dominance	11.85**	3	.74	3/258	3.61*	3/263
Capacity for Status	8.94*	3	.86	3/258	4.48**	3/263
Sociability	8.88*	3	.90	3/258	2.09	3/263
Self Acceptance	10.17*	3	.03	3/258	3.74*	3/263
Sense of Well-Being	6.89	3	1.38	3/258	1.32	3/263
Achievement Via Conformity	7.34	3	.75	3/258	2.32	3/263
Achievement Via Independence	12.13	3	.43	3/258	2.49	3/263
Intellectual Efficiency	8.49*	3	1.15	3/258	1.37	3/263

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

TABLE 5.6.4 (Cont'd)

Summary of Analysis of Covariance Results for the Sixth Grade Students in the Four Instructional Treatments with Study Orientation Subscale Scores as the Criterion

Aptitude	Hypothesis					
	Homogeneity of Variance (X^2 Statistic)	Degrees of Freedom	Equal Slopes of Regression (F Statistic)	Degrees of Freedom	Equal Intercepts of Regression (F Statistic)	Degrees of Freedom
Flexibility	11.38**	3	.49	3/258	3.73*	3/263
Achievement Motivation	1.74	3	2.50	3/419	9.22**	3/424
Creativity						
Fluency - Word Beginnings	1.86	3	1.64	3/419	2.10	3/424
Fluency - Word Beginnings and Endings	5.97	3	.80	3/166	.60	3/171
Fluency - Consequences	5.88	3	1.52	3/418	1.92	3/423
Flexibility - Consequences	9.54*	3	1.43	3/439	1.47	3/444
Originality - Consequences	10.12*	3	.76	3/438	1.26	3/443
Fluency - Controlled Associations	4.24	3	1.32	3/265	.69	3/270

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

TABLE 5.6.5

Summary of Analysis of Covariance Results for the Seventh Grade Students in the Four Instructional Treatments with Study Orientation Subscale Scores as the Criterion

Aptitude	Hypothesis					
	Homogeneity of Variance (X^2 Statistic)	Degrees of Freedom	Equal Slopes of Regression (F Statistic)	Degrees of Freedom	Equals Intercepts of Regression (F Statistic)	Degrees of Freedom
Self Concept Scale	1.70	3	1.45	3/299	3.95**	3/304
Intellectual Achievement Responsibility Scale	3.24	3	1.38	3/308	4.53**	3/313
General Anxiety	2.33	3	2.20	3/272	5.05**	3/277
School Anxiety	3.34	3	.45	3/273	5.31**	3/280
Personality						
Dominance	2.23	3	.10	3/143	4.57**	3/148
Capacity for Status	3.23	3	.56	3/143	4.44**	3/148
Sociability	2.35	3	1.51	3/143	4.57**	3/148
Self Acceptance	2.19	3	.04	3/143	4.65**	3/148
Sense of Well-Being	.98	3	.97	3/143	3.82*	3/148
Achievement Via Conformity	3.90	3	1.01	3/143	5.22**	3/148
Achievement Via Independence	2.37	3	1.32	3/143	4.19**	3/148
Intellectual Efficiency	2.58	3	2.93*	3/143	3.24*	3/148

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

TABLE 5.6.5 (Cont'd)

Summary of Analysis of Covariance Results for the Seventh Grade Students in the Four Instructional Treatments with Study Orientation Subscale Scores as the Criterion

Aptitude	Hypothesis					
	Homogeneity of Variance (X^2 Statistic)	Degrees of Freedom	Equal Slopes of Regression (F Statistic)	Degrees of Freedom	Equal Intercepts of Regression (F Statistic)	Degrees of Freedom
Flexibility	1.14	3	.88	3/143	4.96**	3/148
Achievement Motivation	6.10	3	.04	3/310	.76	3/315
Creativity						
Fluency - Word Beginnings	5.67	3	.61	3/324	.75	3/329
Fluency - Word Beginnings and Endings	5.45	3	1.63	3/322	1.11	3/327
Fluency - Consequences	5.48	3	1.77	3/322	1.23	3/327
Flexibility - Consequences	4.81	3	.04	3/151	2.11	3/156
Originality - Consequences	7.96*	3	.63	3/312	.74	3/317
Fluency - Controlled Associations	2.24	3	.22	3/102	2.68*	3/102

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

TABLE 5.6.6

Summary of Analysis of Covariance Results for the Eighth Grade Students in the Four Instructional Treatments with Study Orientation Subscale Scores as the Criterion

Aptitude	Hypothesis					
	Homogeneity of Variance (X^2 Statistic)	Degrees of Freedom	Equal Slopes of Regression (F Statistic)	Degrees of Freedom	Equal Intercepts of Regression (F Statistic)	Degrees of Freedom
Self Concept Scale	9.24*	3	6.05**	3/290	.95	3/295
Intellectual Achievement Responsibility Scale	8.05*	3	2.46	3/297	.53	3/302
General Anxiety	2.22	3	.31	3/264	.25	3/269
School Anxiety	3.30	3	.25	3/267	.14	3/272
Personality						
Dominance	17.77**	3	1.95	3/132	.71	3/137
Capacity for Status	18.41**	3	1.49	3/132	.85	3/137
Sociability	17.84**	3	.62	3/132	.64	3/137
Self Acceptance	20.12**	3	.78	3/132	.74	3/137
Sense of Well-Being	18.90**	3	1.40	3/132	1.38	3/137
Achievement Via Conformity	21.64**	3	.67	3/132	1.27	3/137
Achievement Via Independence	18.76**	3	.98	3/132	.98	3/137
Intellectual Efficiency	19.08**	3	.76	3/132	1.28	3/137

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

TABLE 5.6.6 (Cont'd)

Summary of Analysis of Covariance Results for the Eighth Grade Students in the Four Instructional Treatments with Study Orientation Subscale Scores as the Criterion

Aptitude	Hypothesis					
	Homogeneity of Variance (X^2 Statistic)	Degrees of Freedom	Equal Slopes of Regression (F Statistic)	Degrees of Freedom	Equal Intercepts of Regression (F Statistic)	Degrees of Freedom
Flexibility	11.58**	3	1.49	3/130	.74	3/135
Achievement Motivation	1.23	3	1.66	3/280	1.50	3/287
Creativity						
Fluency - Word Beginnings	5.30	3	.27	3/98	6.01**	3/103
Fluency - Word Beginnings and Endings	5.65	3	.43	3/290	3.43*	3/290
Fluency - Consequences	5.06	3	.84	3/125	1.96	3/130
Flexibility - Consequences	4.58	3	.82	3/307	1.42	3/312
Originality - Consequences	4.64	3	.64	3/276	.43	3/281
Fluency - Controlled Associations	4.52	3	.66	3/276	.43	3/281

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

reached significance when the Study Orientation scores were the criterion.

The plots of the significant interactions appear in Figures 5.6.1, 5.6.2, and 5.6.3. When Dominance was the aptitude, treatment 5 students scored higher than students in treatments 2 and 3 at all levels. The regression line for treatment 4 crossed the regression lines for each of the other groups. Thus students in treatment 4 with the highest score in dominance tend to score higher on the composite achievement test than students in all other groups. Students in treatment 4, who scored lowest in dominance, conversely, were the lowest achievers.

Since this study was exploratory, no further analyses were conducted following a rejection of the equal slopes of regression hypothesis. The next step, in most cases, would be the determination of the point along the aptitude dimension at which students high on dominance would profit most from instruction in treatment 4 and students low on dominance from another instructional program. As discussed in a previous section, the Johnson-Neyman technique defines a region of homogeneity about the crossover point of two non-parallel regression lines and could be invoked for this purpose. To identify these regions of homogeneity, however, would imply greater generality of the ATI results than is felt to be warranted. Further work, in this direction, obviously, is strongly suggested.

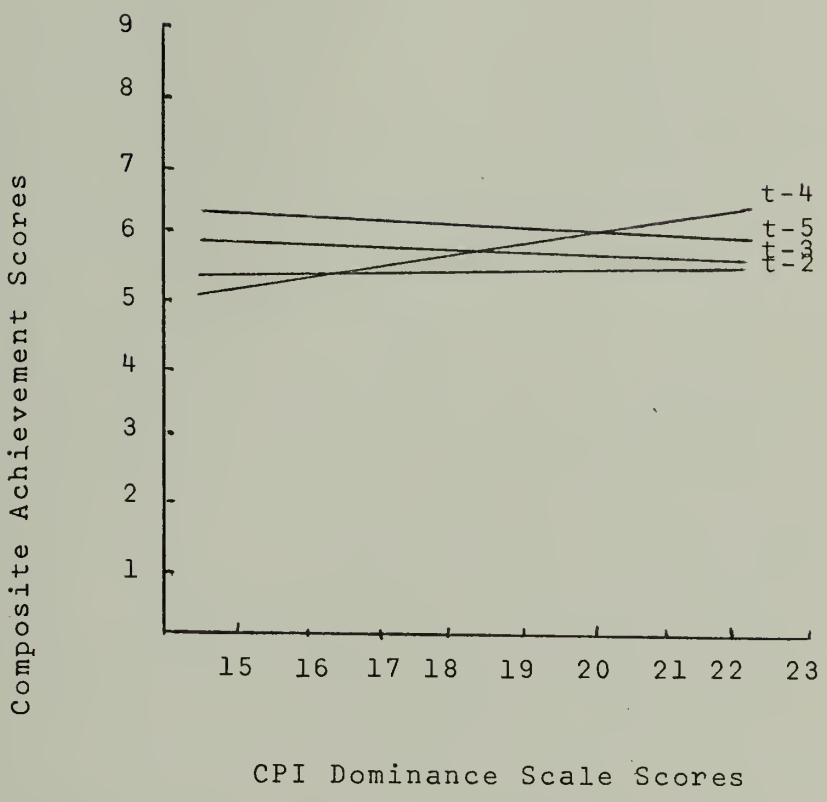


Figure 5.6.1 The regression of composite achievement scores on CPI dominance scores for all instructional treatment groups at the sixth grade level.

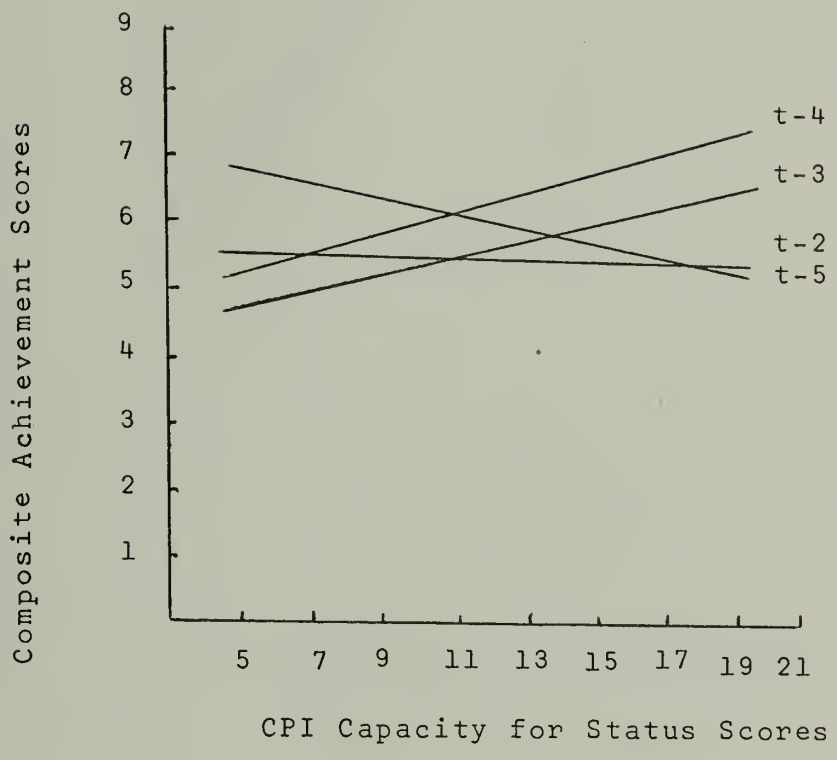


Figure 5.6.2 The regression of composite achievement scores on CPI capacity for status scores for all instructional treatment groups at the sixth grade level.

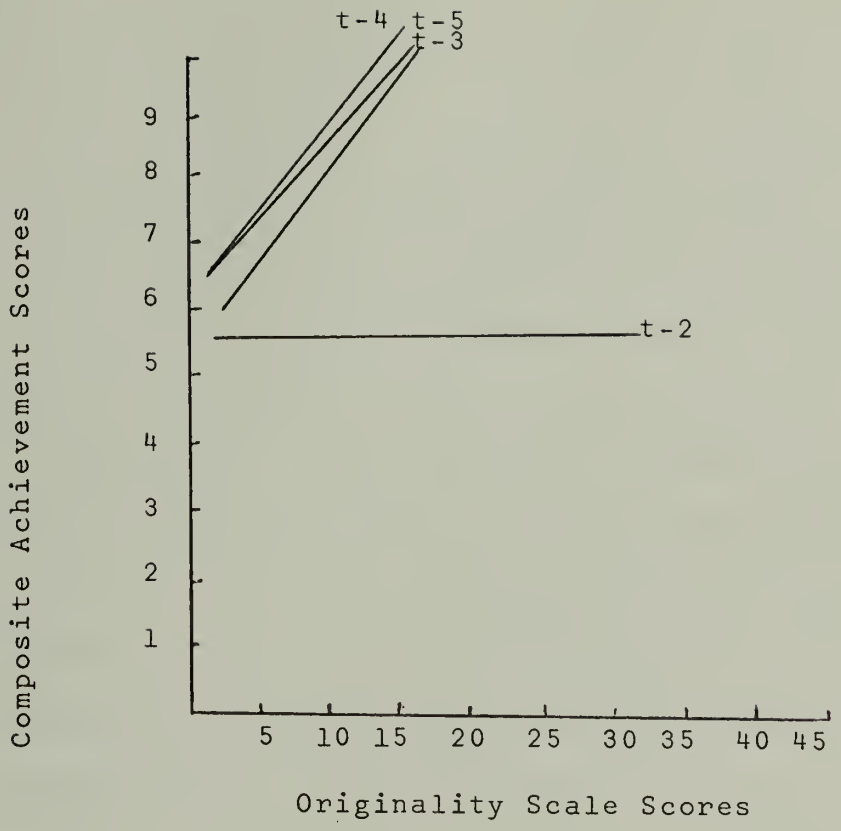


Figure 5.6.3 The regression of composite achievement scores on originality scores for all instructional treatment groups at the sixth grade level.

The pattern of regression of criterion on aptitude in Figure 5.6.2 appear to be disordinal and are probably most exemplary of what one seeks in ATI investigations. Students highest on the capacity for status subscale of the CPI appear to perform best in instructional treatment 4. Students lowest in capacity for status appear to perform best in treatment 5. The other clear case of a disordinal interaction involves treatment 3 and treatment 5.

The capacity for status subscale attempts to measure the personality qualities and attributes which underlie and lead to status. The picture one gets is that the ambitious, active, forceful and insightful individual will perform best in the learning environment facilitated in treatment 4. The shy, conventional and mild individual who may be restricted in outlook and interests is more likely to succeed in treatment 5. Before discussing this further we should consider the results for originality where the criterion again was the composite achievement test scores.

An inspection of Figure 5.6.3 reveals that treatment 4 students who have high originality scores on the Consequences Test again achieve highest in treatment 4, though only slightly better than students in treatments 3 and 5. The interaction, obviously, is ordinal. The results for treatment 2 are suspect. One would hardly expect students from such similar background and experiences to typically produce profiles as divergent as those in Figure 4.6.3. Therefore, comment is withheld pending other data.

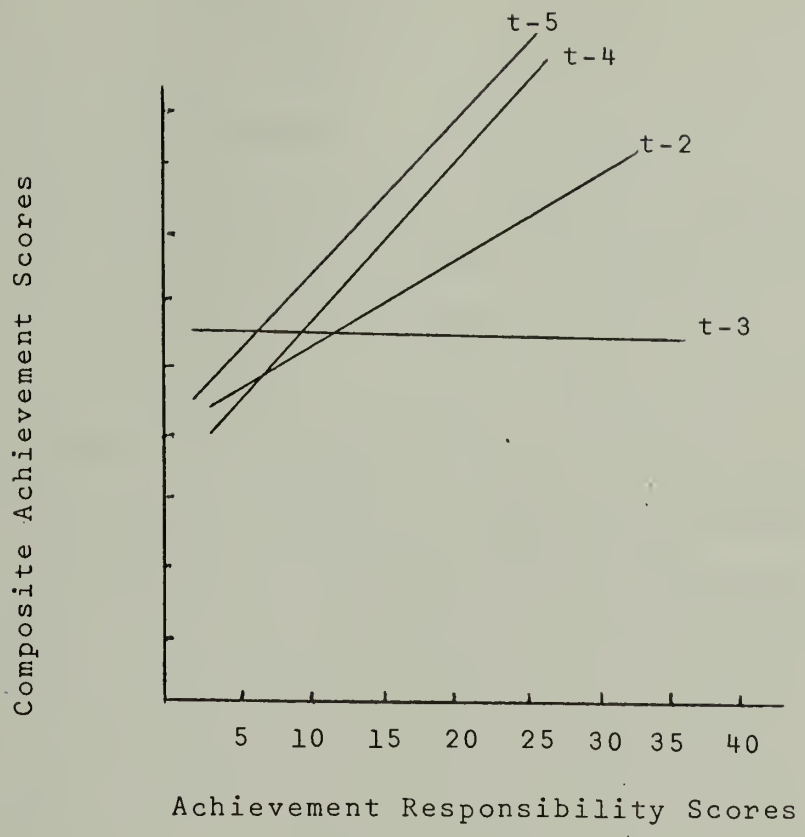


Figure 5.6.4 The regression of composite achievement scores on achievement responsibility scores for all instructional treatment groups at the seventh grade level.

to exercise much less control over students, there are fewer individuals for whom the student can honestly blame for his failure, thus his total IAR score would be low compared to students in other treatments.

The second rejection of the parallelism hypothesis occurred when the Flexibility scale of the CPI was the aptitude and the composite achievement test score was the criterion. Figure 5.6.5 represents the four pairs of treatment groups when the composite achievement test scores were used as the criterion and the CPI Flexibility scores were used as the predictors. Several possible disordinal interactions appear to exist. The two most prominent involve treatment 3 with treatment 5 and treatment 4 with treatment 5. In both cases the less flexible individual appears to achieve best in treatment 5. The most flexible individual would probably achieve highest in treatment 3. One would expect such results for the most open program in the study.

The only significant interaction observed with Study Orientation scores as the criteria was obtained for the seventh grade groups. The relevant aptitude was intellectual efficiency. Figure 5.6.6 shows the regression lines for the four treatments when the composite achievement scores were used as criteria and the CPI Intellectual Efficiency scores were used as predictors. The intellectually efficient individual is clear-thinking, capable, planful, and resourceful. All of these qualities, one would expect to be relevant attributes of persons who have positive attitudes toward learning,

The results for the sixth grade groups would not be noteworthy from a statistical point of view alone since only one of them was significant at better than a .05 level. However, the composite picture is impressive. The three relevant dimensions of aptitude taken together imply that students who are domineering, confident, ambitious, independent or inventive are most likely to achieve best in the highly organized and structured treatment 4 program. The same kind of student shows the least amount of achievement in the more restrictive and traditional treatment 5. The submissive, retiring, shy person appears to thrive in a program like treatment 5.

At the seventh grade level F statistics significant at the .05 level were obtained for three of the predictor variables. The first significant F was for the Intellectual Achievement Responsibility Scale where the criterion was the composite achievement scores. The plots of these treatment groups are presented in Figure 5.6.4. It appears that students in every group except treatment 3 score higher on the criterion the higher their scores are on the IAR. The more students in treatments 2, 4, and 5 tend to attribute their successes to internal forces, the higher their achievement. The plots of treatment groups appear to represent an ordinal interaction. However, instructional treatment group 3's performance may be attributed to the greater freedom allowed students. Since administrators and teachers attempt

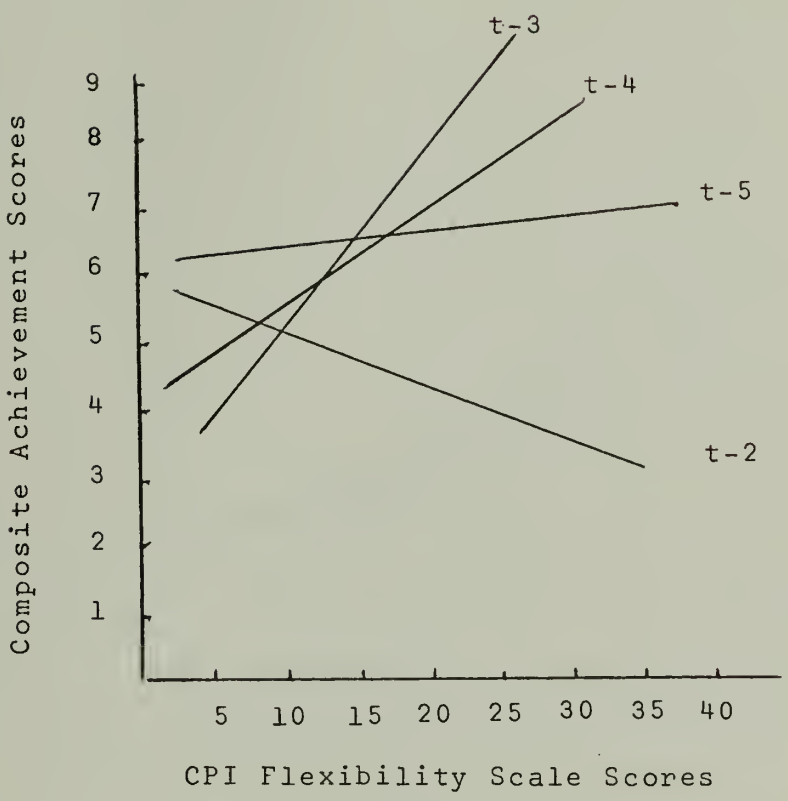


Figure 5.6.5 The regression of composite achievement scores on CPI flexibility scores for all instructional treatment groups at the seventh grade level.

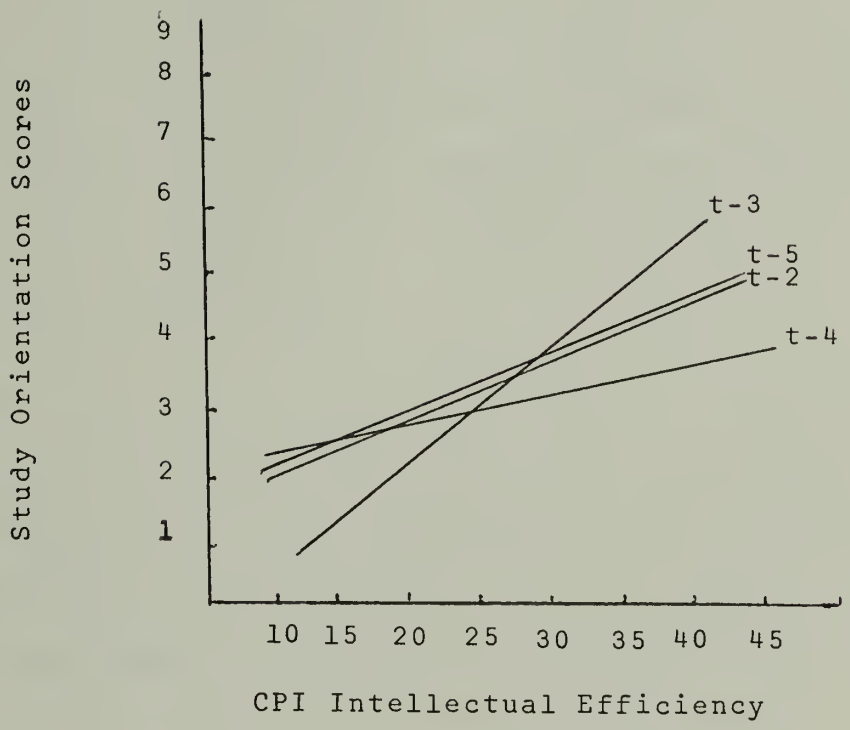


Figure 5.6.6 The regression of study orientation scores on CPI intellectual efficiency scores for all instructional treatment groups at the seventh grade level.

school, and studying in a program which grants 40% of school time to students to use in any fashion they desire. It is certainly not surprising, then, that students in treatment group 3, who score high on the Intellectual Efficiency subscale of the CPI, would also have highest SO scores.

Analysis of covariance conducted on the eighth grade data did not result in the identification of a single significant interaction.

Summary

The ATI analyses resulted in six interactions, only one of which, Dominance, was significant at the .01 level. Trends across grades were not obtained. However, the results were psychologically meaningful and, as a whole, suggested that dominant, enterprising students achieved best in a structured and challenging environment and least in a highly restrictive one. Shy, dependent, submissive individuals appear to excel in highly structured environments where a great deal of direction is provided.

Chapter VI

Conclusions and Implications for Further Research

6.1 General Summary

In view of the importance of comprehensive evaluations of school programs and the methodological weaknesses of existing studies, this investigation was designed to accomplish two goals. One was to compare the effects of five instructional treatments which were thought to differ in philosophical, procedural, and programmatic features. The second objective was to determine whether the instructional programs interacted with student aptitudes to produce higher performance on selected criteria. Data was collected on a number of variables in the areas of achievement, attitudes, environmental perception and creativity. The results suggested that differences existed along each of these dimensions. A summary and discussion of the highpoints of the analysis should further clarify the important findings of the study.

6.2 Comparison Studies

The analysis of the achievement scores revealed interesting results. Students in the two treatments (4 and 5) with the most structured programs scored highest on the composite achievement tests. It could well be that instructional treatments 4 and 5 were yet best suited for producing

high scholastic achievement as measured on standardized achievement tests. Another interpretation of the differences in achievement becomes apparent upon consideration that students in the other three treatments were only spending 60% of their school time in formal instruction. The alternative explanation for differences in achievement is that treatments 1, 2, and 3 were not as efficient as they are likely to become. Since they had been in operation for three years at the most, it might be argued that instructional programs 1, 2, and 3 were still developing, whereas the programs in the other schools had been evolving over many more years and were thus more firmly established.

The results of the analyses of the creativity test scores are contrasted with those of the achievement test data. At the sixth grade level, students in treatment 4 had the highest mean score on a discriminant function characterized by high fluency, flexibility, and to a lesser extent, originality, but low scores on word fluency. Treatment 5 had the lowest mean score on this function which includes many of the major characteristics we normally associate with creative persons (ideational fluency, flexibility, and originality).

Treatment 5 students in the sixth grade had the highest mean score for a second discriminant function characterized by low fluency and originality scores but high flexibility scores. The picture one gets of sixth graders in treatment 5,

as a result of their scores on this function, is that of students who are readily able to recall appropriate material but less adaptible in situations requiring problem solving abilities or innovation.

For the seventh and eighth grades the major difference between students related again to a contrast between ideational fluency and flexibility. In both seventh and eighth grades the students in treatments 4 and 5 could be primarily characterized as being lowest on ideational fluency and originality and highest on flexibility and word fluency.

The above results on achievement and creativity variables seem to have tremendous implications. Most educators agree that in times of rapid change we must be more concerned about producing and stimulating creative thinkers. It is often pointed out that instruction in most American schools is designed to produce students who can excel on achievement and intelligence tests. The implicit but primary purpose of education, given this assumption, is to aid students in developing skills which enable them to achieve this goal. Since available evidence shows that creative and intelligent conceptualizing reflect different cognitive styles, we must now consider the possibility that optimum environments for one will not be optimum for the other. The results of this study certainly seem to indicate that the different programs foster different skills, different thematizing modes and different styles of relating

to problems. Certainly, this possibility is worthy of further study.

The results of the analyses conducted on the sixth grader's study orientation scores showed that treatment 4 students had better study habits and attitudes toward school and school-related activities than students in treatments 3 and 5. For the seventh grade groups treatment 3 students had higher scores on the SO scale than students at other schools. The reversal here may well be due to the transition seventh grade students in treatment 3 experienced moving from the somewhat controlled program offered in the sixth grade to the more open arrangement available in the seventh grade. The fact that the significant differences did not hold up in the eighth grade might be due to the fact that students had become accustomed to the unique program and had begun to take the available freedom for granted. They then might revert to old habits and attitudes toward school. For students just entering the seventh grade the novelty of the program may have stimulated improved study habits and attitudes that apparently did not hold up after a year in the program. A situation like this might well provide the perfect opportunity for studying the way in which humans adapt to new experiences and slowly become impassive.

Significant differences between grades were observed in the study orientation scores of the SSHA. Sixth and seventh grade students, on the average, were found to have

better attitudes towards studying, school related activities, and education in general. Several other researchers have reported changes in attitudes over several school years (Neale and Proshek, 1967; Anttonen, 1967) as well as over a single school year (Neale, Gill and Tismer, 1970). In each case, attitudes toward school were found to decrease. If student attitudes decrease from the beginning of the year to the end and from year to year, then it is reasonable that other student characteristics also change as students progress through school. Anxiety, self concepts, responsibility or conforming behavior may also show differences across students from grade to grade. These questions were not investigated in this study. The results obtained here, however, certainly suggest that such questions should be studied.

One way of determining whether students differ on other measurable attributes across grade levels is to factor analyze responses on the questionnaires for each grade in so-called innovative schools and each grade in conventional schools. Comparison of the obtained factors for each grade could be made between and within instructional programs and the derived components correlated with achievement criteria. Such data on how students differ from grade to grade and how the important dimensions of personality and affect correlate with academic achievement, or how personality and cognitive variables correlate with affective measures, would provide us with a better understanding of personality growth and attitudinal change and their relation to performance.

6.3 ATI Studies

The major reason for conducting the ATI studies was to identify interactions which were congruous with the three defining limitations set forth at the beginning of section 5.6.1. However, while only one significant interaction met all three of those criteria, we have discussed several in some detail. The fact is that the results of the ATI studies were easily interpretable in terms of the defining characteristics of the respective instructional treatments. In keeping with the purposes of the investigation, the results of the ATI studies are tentatively interpreted. The suggestion is that more carefully contrived treatments with better planned programs and more distinctive features might result in more obvious effects. Such aptitudes as dominance, capacity for status, achievement responsibility, flexibility and intellectual efficiency could well be relevant traits on which to base such ATI investigations.

While the traits identified in the ATI studies were psychologically interpretable and meaningful, serious questions still abound. For example, none of the aptitudes occurred across grades. If the programs at the seventh and eighth grade levels were basically the same, as they were said to be, then one would expect to find that some of the same attributes which facilitated performance at the seventh grade level would reappear with eighth grade groups. This did not occur. Obviously, more developmental studies are required to resolve this issue.

On the other hand, the paucity of evidence of ATI's may mean that the wrong aptitudes were used in this investigation, the criteria were too global or the treatments were not variable enough. While the programs differed, the differences may not have been dramatic enough. It is possible that significant ATI's may not be found until there are substantial differences in the instructional treatments. A future ATI investigation, then, might include, for example, a very traditional instructional program as one treatment and an experientially based program as the other.

6.4 Limitations

Several limitations to the studies reported here are in order. The more pervasive limitation has already been discussed. The fact that the experimenters had no control over the programs themselves meant that there was considerable overlap or similarity between the various treatments. While this situation confronts any researcher who attempts to conduct his work in on-going schools, it is less than desirable experimentally. For this reason results can only be discussed as tentative and used as a basis for planning more controlled studies designed to investigate specific questions.

Another major limitation was that the treatment conditions represented a limited sample of all possible treatments. Treatment 5 may have been close to one extreme on a dimension of "structured-unstructured or "open-closed". However, none of the other treatments

could possibly be considered among the most open, flexible, or innovative approaches to instruction that are being practiced in schools today.

The unavailability of pretest data precluded any consideration of growth. Since data was available for three consecutive grades, comparisons across ages was possible, however. The student population represented the upper middle class. Whatever implications are drawn are thus limited to like populations.

Finally, it is recognized that the effectiveness of the testing program was dependent upon the manner in which teachers and administrators went about their tasks. The results of the study are limited to the extent that they failed to foster an atmosopher of seriousness with regards to the testing. There was reason to feel that one or two teachers were not as conscientious as they should have been. Fortunately, the majority of the teachers were very cooperative, and went about their tasks with total dedication.

6.5 Concluding Remarks

The major results of this investigation suggest that instructional treatments shown to be different on several relevant dimensions, also differ in the effects they have on students. It is felt that this result, in and of itself, has important implications. Several prominent investigators have taken the position that there are few, if any, real differences in the effects of different schools, instructional

programs, etc. Coleman (1966) and Jencks et al. (1972) are examples that can be readily cited.

Coleman (1966), in a massive study of socio-economic factors, concluded that the attitudes of the individual student was the most important factor affecting his learning. Specifically, the individual's confidence in his ability to shape his future or control his destiny was considered "to have a stronger relationship to achievement than . . . all other school factors together." (p. 23).

More recently, Christopher Jencks et al. (1972), in a very interesting report, have presented arguments to the effect "that neither the overall level of resources available to a school nor any specific, easily identifiable school policy has a significant effect on a student's cognitive skills or educational attainments." Therefore, they conclude that none of the programs or structural arrangements in common use today have consistently different long-term effects from any other.

The results of the investigation reported here are in direct contradiction with the above positions. Here, differences were observed on all of the variables included in the study - achievement, attitudes, environmental perception, fluency, flexibility, and originality. Very few, if any, investigators have attempted to contrast the effects of instructional programs on as many dimensions as were represented in this study. The results suggest a need for

more extensive research in this area. Since some of the measures used here were global in nature, it is possible that even more impressive results could be obtained given more specific variables and more diverse treatments.

Future comparative studies should be limited to two instructional programs so that greater effort can be directed towards defining the process and structural distinctions that exist. In this way differences in effect can be more easily traced to and discussed in terms of specific programmatic features. Such a procedure would all but eliminate one of the most glaring weaknesses of the present study - the fact that there were so many possible explanations of the results. Thus, it is almost impossible to determine whether the differences observed were due to teachers, subject matter, the composition of the classes or some other factors.

The above comment suggests that a more comprehensive method than was used in the present study is needed for describing instructional programs. Generalizability, as well as interpretability, would be facilitated by such a development. One procedure for going about this task would require the collection of extensive data on a number of different instructional programs. The gathering of this kind of data would require extensive observation, the use of interviews, and the administering of relevant questionnaires and tests to students, teachers, and administrators. The data collected should cover a number of variables

dealing with all aspects of schools - administration, instructors, climate, and educational philosophy. The important dimensions of program differences could be determined by factor analysis.

In conclusion, the study reported here was carried out with the intent of generating research questions that might lead to more effective evaluation of schools and more profitable research on innovative instructional programs. The major result was that instructional treatments found to be different on the basis of a semi-objective ranking procedure also differed in their effects on students.

REFERENCES

Abelson, R.P. A note on the Neyman-Johnson technique. Psychometrika, 1953, 18, 213-217.

Amidon, E. & Flanders, N.A. The effects of direct and indirect teacher influence on dependent-prone students learning geometry. Journal of Educational Psychology, 1961, 52, 286-291.

Anderson, G.J. The assessment of learning environments: A manual for the Learning Environment Inventory and the My Class Inventory. Halifax, N.S.: Atlantic Institute of Education, 1971.

Anderson, R.L. & Bancroft, T.A. Statistical theory in research. New York: McGraw-Hill, 1952.

Anttonen, R.G. An examination into the stability of mathematics attitude and its relationship to mathematics achievement from elementary to secondary school level. Unpublished doctoral dissertation, University of Minnesota, 1967.

Auria, C. & Frankiewicz, R.G. Components of school satisfaction among a group of ninth and tenth graders. Reprinted from the Proceedings, 75th Annual Convention, APA, 1967.

Bartlett, M.S. The goodness of fit of a single hypothetical discriminant function in the case of several groups. Annals of Eugenics, 1951, 16, 199-214.

Beach, L.R. Socialibility and academic achievement in various types of learning situations. Journal of Educational Psychology, 1960, 51, 208-212.

Berliner, D.C. & Cahen, L.S. Trait-treatment interactions and learning. In F.N. Kerlinger (Ed.), Review of Research in education. New York: F. E. Peacock Co., in press.

Bloom, B.S. Learning for mastery. UCLA-CSEIP Evaluation Comment, 1, No. 2, 1968.

Bock, R.D. & Haggard, E.A. The use of multivariate analysis of variance in behavioral research. In D.K. Whittla (Ed.), Handbook of measurement and assessment in behavioral sciences. Reading, Mass.: Addison-Wesley, 1968.

Borg, W.R. & Brite, L.R. Teacher's perceptions of team-teaching. California Journal of Educational Research, 1967, 18, 71-81.

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

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

Brown, W.F. & Holtzman, W.H. Manual for the Survey of Study Habits and Attitudes. New York: The Psychological Corporation, 1964.

Burningham, G.L. A study and evaluation of the team-teaching of the fourth grade at Woodstock Elementary School. (Doctoral dissertation. University of Utah), Ann Arbor, Michigan: University Microfilms, 1968, No. 68-11, 905.

Carroll, J.B. A model of school learning, Teachers College Record, 1963, 64, 723-733.

Castenada, A., McCandless, B.R., & Palermo, D.S. The children's form of the manifest anxiety scale. Child Development, 1956, 27, 317-326.

Clark, C.M., Veldman, D.J., & Thorpe, J.S. Convergent and divergent thinking abilities of talented adolescents. Journal of Educational Psychology, 1965, 56, 157-163.

Coleman, J.C. Equality of educational opportunity. Washington: U.S. Department of Health, Education, and Welfare, Office of Education, 1966.

Cramer, E.M. & Bock, R.D. Multivariate analysis. Review of Educational Research, 1966, 36, 604-617.

Crandall, V.E., Katkovsky, W., & Crandall, V.J. Children's beliefs in their own control of reinforcements in intellectual-academic achievement situations. Child Development, 1964, 35, 91-109.

Cronbach, L.J. Intelligence, creativity? A parsimonious re-interpretation of the Wallach-Kogan data. American Educational Research Journal, 1968, 5, 491-512.

Cronbach, L.J. & Snow, R.E. Individual differences in learning ability as a function of instructional variables. Final Report, U.S. Office of Education, Contract No. OEC-4-6-061269-1217, School of Education, Stanford University, 1969.

Gibbons, M. What is individualized instruction? Interchange, 1970, 1, 28-52.

Gill, M.P. & D'oyley, V.R. The construction of an objective measure of self concept. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, 1968.

Glaser, R. Adapting the elementary school curriculum to individual performance. In Proceedings of the 1967 Invitational Conference on Testing Problems. Princeton, N.J.: Educational Testing Service, 1968.

Gough, H.G. Manual for the California Psychological Inventory. Palo Alto: Consulting Psychologists Press, 1957.

Grimes, J.W. & Allinsmith, W. Compulsivity, anxiety, and school achievement. In J.F. Rosenblith & A. W. Allinsmith (Eds.) The causes of behaviors: Readings in child development and educational psychology. Boston: Allyn & Bacon, 1962, 20-433.

Guilford, J.P. Some misconceptions regarding measurement of creative talents. Journal of Creative Behavior, 1971, 5, 77-87.

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

Gulliksen, H. & Wilks, S.S. Regression tests for several samples. Psychometrika, 1950, 15, 91-114.

Hambleton, R.K. A review of testing and decision-making procedures for selected individualized instructional programs. ACT Technical Bulletin. Iowa City, Iowa: American College Testing Program, 1973.

Hambleton, R.K., Rovinelli, R., Sheehan, D., & Newby, J.F. A study of middle school students in different instructional programs. Final report submitted to Superintendents of Schools - Concord, Sudbury and Carlisle, Massachusetts, September, 1972.

Heather, G. Overview of innovations in organization for learning. Interchange, 1972, 3, 47-68.

Hummel, T.J. & Sligo, J.R. Empirical comparison of univariate and multivariate analysis of variance procedures. Psychological Bulletin, 1971, 76, 49-57.

Crockenberg, S.B. Creativity tests: A boon or boondoggle for education. Review of Educational Research, 1972, 42, 27-46.

Cullen, R.J. & Katzenmeyer, C.G. Achievement and ability correlates of components of school attitude among eighth grade students. Paper presented at the Annual Meeting of the American Educational Research Association, Minneapolis, March 1970.

Davis, G.A. & Belcher, T.L. How shall creativity be measured? Torrance Tests, RAT, Alpha Biographical, and IQ. Journal of Creative Behavior, 1971, 5, 153-161.

Diedrich, R.C. Teacher perceptions as related to teacher-student similarity and student satisfaction with school. Unpublished doctoral dissertation, University of Chicago, 1966.

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

Domino, G. Interactive effects of achievement orientation and teaching style on academic achievement. Journal of Educational Psychology, 1971, 62, 427-431.

Doty, B.A. Teaching method effectiveness in relation to certain student characteristics. Journal of Educational Research, 1967, 60, 363-365.

Dowaliby, F.J. and Schumer, H.S. Teacher-centered versus student-centered mode of college classroom instruction as related to manifest anxiety. Journal of Educational Psychology, 1973, 64, 125-132.

Erickson, G. The predictive validity of a battery of creative thinking tests and peer nominations among University of Minnesota high school seniors seven years later. Unpublished master's research paper, University of Minnesota, 1966.

Frymier, J.R. Development and validation of a motivation index. Theory into practice, 1970, 9, 56-88.

Georgiades, W. & Bjelke, J. Evaluation of English achievement in a ninth grade, three-period, team-teaching class. California Journal of Educational Research, 1966, 17, 100-112.

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

Jackson, P.W. & Getzels, J.W. Psychological health and classroom functioning: A study of dissatisfaction with school among adolescents. Journal of Educational Psychology, 1959, 50, 295-300.

Jackson, P.W. & Lahaderne, H.M. Scholastic success and attitude toward school in a population of sixth graders. Journal of Educational Psychology, 1967, 58, 15-18.

Jencks, C., Smith, M., Aclan, H., Bone, M.J., Cohen, D., Gintis, H., Haynes, B., & Michelson, S. Inequality: A reassessment of the effects of family and schooling in America. New York: Basic Books, 1972.

Johnson, P.O. & Neyman, J. Tests of certain linear hypotheses and their application to some educational problems. In J. Neyman and E.S. Pearson (Eds.), Statistical Research Memoirs, 1936, 1, 57-93.

Khan, S.B. Affective correlates of academic achievement. Journal of Educational Psychology, 1969, 60, 216-221.

Khan, S.B. & Roberts, D.M. Relationships among study habits and attitudes, aptitude and grade 8 achievement. Educational and Psychological Measurement, 1969, 29, 951-955.

Knox, D.M. An experimental study of the effect of a team teaching program upon certain selected variables. (Doctoral dissertation, Western Reserve University), Ann Arbor, Michigan: University Microfilms, 1966, No. 66-8011.

Merrifield, P.R., Gardner, S.F. & Cox, A.B. Aptitude and personality measures related to creativity in seventh-grade children. Reports from the Psychological Laboratory, University of Southern California, No. 28, 1964.

McKeachie, W.J. Anxiety in the college classroom. Journal of Educational Research, 1951, 45, 153-160.

Neale, D.C., Gill, N., & Tismer, W. Relationship between attitudes toward school subjects and school achievement. Journal of Educational Research, 1970, 63, 232-237.

Neale, D.C. & Proshek, J.M. School-related attitudes of culturally disadvantaged elementary school children. Journal of Educational Psychology, 1967, 58, 238-244.

Olsen, G.H. Mulgen: Fortran program for the multivariate general linear hypothesis. Tallahassee: Department of Educational Research of the Florida State University, August, 1970.

Overall, J.E. & Klett, C.J. Applied multivariate analysis. New York: McGraw-Hill, 1972.

Patton, J.A. A study of the effects of student acceptance of responsibility and motivation on course behavior. Unpublished doctoral thesis, University of Michigan, 1955.

Phillips, B.N. An analysis of causes of anxiety among children in school. Final Report, Project No. 2616, U.S.O.E. Cooperative Research Branch, Austin: The University of Texas, 1966.

Plowden, B. Children and their primary schools: A report of the central advisory council for education. Redwood City, California: Pendragon House Inc., 1966.

Potthoff, R.F. On the Johnson-Neyman technique and some extensions thereof. Psychometrika, 1964, 29, 241-256.

Rao, C.R. Linear statistical inference and its applications. New York: Wiley, 1965.

Stern, G.G. Environments for learning. In N. Sanford (Ed.), The american college. New York: Wiley, 1962, 690-730.

Smith, D.E., Wood, R.L., Downer, J.W., & Raygor, A.L. Reading improvement as a function of student personality and teaching method. Journal of Educational Psychology, 1956, 47, 47-59.

Torrance, E.P. Guiding creative talent. Englewood Cliffs, New Jersey: Prentice-Hall, 1962.

Torrance, E.P. Education and the creative potential. Minneapolis: University of Minnesota Press, 1963.

Torrance, E.P. & Myers, R.E. Teaching gifted elementary pupils how to do research. Minneapolis: Perceptive Publishing Co., 1962.

Traub, R.E., Weiss, J., Fisher, C.W., & Musella, D. Closure on openness: Describing and quantifying open education, Interchange, 1972, 3, 69-84.

Traub, R.E., Weiss, J., Fisher, C.W., & Musella, D. Closure on openness in Education. Experimental session presented at the Annual Meeting of the American Educational Research Association, New Orleans, 1973.

Tatsuoka, M.M. Multivariate analysis: Techniques for educational and psychological research. New York: John Wiley & Sons, Inc., 1971.

Wallach, M.A. & Kogan, N. Modes of thinking in young children. New York: Holt, Rinehart & Winston, 1965.

Williams, F.E. Practice and reinforcement as training factors in creative performance. Unpublished doctoral dissertation, University of Utah, 1965.

Yamamoto, K., & Chimbidis, M.E. Achievement, intelligence, and creative thinking in fifth grade children: A correlational study. Merrill-Palmer Quarterly, 1966, 12, 233-241.