4-25-1974

The Comparison of Two Methods of Instruction in Teaching the Wechsler Intelligence Scale for Children

Kathleen Yost Ritter

TEACHING THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN - RITTER

THE UNIVERSITY OF NEW MEXICO
ALBUQUERQUE, NEW MEXICO 87106

Policy on Use of Theses and Dissertations

Unpublished theses and dissertations accepted for master's and doctor's degrees and deposited in the University of New Mexico Library are open to the public for inspection and reference work. *They are to be used only with due regard to the rights of the authors.* The work of other authors should always be given full credit. Avoid quoting in amounts, over and beyond scholarly needs, such as might impair or destroy the property rights and financial benefits of another author.

To afford reasonable safeguards to authors, and consistent with the above principles, anyone quoting from theses and dissertations must observe the following conditions:

1. Direct quotations during the first two years after completion may be made only with the written permission of the author.

2. After a lapse of two years, theses and dissertations may be quoted without specific prior permission in works of original scholarship provided appropriate credit is given in the case of each quotation.

3. Quotations that are complete units in themselves (e.g., complete chapters or sections) in whatever form they may be reproduced and quotations of whatever length presented as primary material for their own sake (as in anthologies or books of readings) ALWAYS require consent of the authors.

4. The quoting author is responsible for determining "fair use" of material he uses.

This thesis/dissertation by Kathleen Yost Ritter has been used by the following persons whose signatures attest their acceptance of the above conditions. (A library which borrows this thesis/dissertation for use by its patrons is expected to secure the signature of each user.)

NAME AND ADDRESS DATE

This dissertation, directed and approved by the candidate's committee, has been accepted by the Graduate Committee of The University of New Mexico in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

THE COMPARISON OF TWO METHODS OF INSTRUCTION IN TEACHING THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN

Title

KATHLEEN YOST RITTER

Candidate

GUIDANCE AND COUNSELING

Department

John P. Zepper

Dean

April 25, 1974

Date

Committee

Marvin J Hersey

Chairman

George L. Keppers

Wayne R. Maes

THE COMPARISON OF TWO METHODS OF INSTRUCTION IN TEACHING THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN

BY
KATHLEEN YOST RITTER
B.S., University of Utah, 1964
M.S., Kansas State Teachers College, 1967
Ed.S., Kansas State Teachers College, 1970

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy in Pupil Personnel Services
in the Graduate School of
The University of New Mexico
Albuquerque, New Mexico
May, 1974

ACKNOWLEDGMENTS

This dissertation is dedicated to my parents, Wilda and Reuben Yost, whose loving support has been a source of strength for me for all of my 32 years. It is with love and deep gratitude that I dedicate my efforts to them.

My appreciation goes to my chairman, Dr. Marion Heisey, under whose direction this study was developed and conducted. He kindly allowed me to use his two Differential Diagnosis I classes in this study. He spent many extra hours preparing materials for the experiment, and I am grateful to him for his cooperation and faith in me.

The graduate assistant in the Differential course, Bob Guy, gave many hours above and beyond "the call of duty." During the final week of the experiment, he watched 25 WISC demonstrations, in addition to his other responsibilities. I so much appreciate his sacrifice and interest in the study.

A special thanks goes to Dr. Clifford Abe who served on my committee before he left the University of New Mexico. His help, suggestions, and careful reading in the initial phases of the study are deeply appreciated.

And, to my other committee members, Dr. Wayne Maes and Dr. George Keppers, goes my appreciation. Dr. Maes's comments and suggestions were invaluable and readily taken. Dr. Keppers's experience and wise counsel were helpful.

Help in the statistical analysis was provided by James A. Willis in the Computing Center at the University of New Mexico. I am grateful to him for taking time from his other obligations to help me. A special thanks, too, to Richard Rindone in the Department of Educational Foundations at UNM for his programming assistance.

My appreciation is extended to Margaret Weinrod who typed the final copy of this dissertation. Her ability to organize the piles of paper given to her is commended.

I am grateful to my mother-in-law, Anna Ritter. During the nearly nine years that we have been in the same family, her support and reassurance have been constantly felt.

And last, but most importantly, my gratitude and love to those three very special people in my life--John, Alicia, and Mary Kay. John, because he has supported my efforts through the years, and Alicia and Mary Kay, because of who they are. How often they have asked: "Mommy, when are you going to finish that stupid paper so we can have some fun?" Well, dear ones, it's finished. Now let's have that fun!

Help in the statistical analysis was provided by James A. Willis at the Computing Center at the University of New Mexico. I am grateful to him for taking time from his other obligations to help me. A special thanks, too, to Richard Kadono in the Department of Educational Foundations at UNM for his programming assistance.

My appreciation is extended to Margaret Weinrod who typed the final copy of this dissertation. Her ability to organize the piles of paper given to her is commended.

I am grateful to my mother-in-law, Anna Ritter. During the nearly nine years that we have been in the same family, her support and enthusiasm have been constantly felt.

And last, but most importantly, my gratitude and love to those three very special people in my life: John, Alicia, and Mary Kay. John, because he has supported my efforts through the years, and Alicia and Mary Kay, because of who they are. How often they have asked: "Mommy, when are you going to finish that stupid paper so we can have some fun?". Well, dear ones, it's finished. Now let's have that fun!

THE COMPARISON OF TWO METHODS OF INSTRUCTION IN TEACHING THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN

BY
Kathleen Yost Ritter

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy in Pupil Personnel Services
in the Graduate School of
The University of New Mexico
Albuquerque, New Mexico
May, 1974

THE COMPARISON OF TWO METHODS OF INSTRUCTION IN TEACHING THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN

Kathleen Yost Ritter, Ph.D.
Department of Guidance and Counseling
The University of New Mexico, 1974

Statement of the Problem

The study was conducted in an effort to determine if there were any differences in terms of student learning between two methods of instruction in teaching the Wechsler Intelligence Scale for Children (WISC). The scores on two outcome measures--a cognitive examination and a performance examination--were compared between students in an Experimental group who learned by means of independent study and those in a Control group who learned to understand, administer, and score the WISC by means of a traditional method.

Procedures

The study was conducted in the Department of Guidance and Counseling at the University of New Mexico. The subjects were those students enrolled in two sections of an individual intelligence testing course. Initial differences between the groups were ascertained on 18 variables--i.e., age, sex, undergraduate grade point average, and the 15 scales of the Edwards Personal Preference Schedule (EPPS).

The subjects in the Experimental class attended a full-day workshop. Other than this one required attendance, they were not asked to attend any other

class meetings. Ten cassette lectures were available for their check out, informational handouts were distributed, and two texts were recommended. Three WISC videotape demonstrations were shown. They were encouraged to watch others administer the WISC, administer the WISC for observation, "hand-in" protocols for checking, and meet with the instructor and graduate assistant if they wished.

The subjects in the Control class were taught by a fairly traditional method. The regular class meetings consisted of lectures over the same material that was available to the Experimental group on the cassettes, WISC demonstrations, and discussions. They were required to administer 15 WISC tests and hand in 15 scored protocols, observe 5 colleagues administering the WISC, and administer one test for observation. The 3 videotape demonstrations were also optional for this group.

Results

Taking initial differences into consideration, an analysis of covariance procedure indicated no significant differences between the groups on the cognitive examination scores and the performance examination averages. The Control group obtained significantly higher scores on one performance examination item (Accuracy of scoring procedures).

None of the initial variables correlated with either the cognitive examination scores or the performance examination averages in the Experimental and combined groups. For the Control group, both age and the EPPS variable of Change correlated negatively with the cognitive examination scores. The variable

of Nurturance correlated negatively with the performance examination averages.

Significantly high relationships were found between the two outcome measures of the Experimental and combined groups. The Control group data indicated little relationship between the cognitive examination scores and the performance examination averages.

The number of WISC tests administered by the Experimental group correlated significantly with both cognitive examination scores and performance examination averages. A significant correlation was obtained between the number of protocols handed in for evaluation and the cognitive examination scores of the Experimental group.

Conclusions

1. The data essentially confirmed the findings of previous research in the area of comparative teaching methods--that is, that one method of instruction was generally as effective as another in affecting outcome on a final examination. Success on outcome measures is less dependent upon method of instruction than upon the personality of the individual. Although it was not selected as an initial measure in this study, level of motivation was thought to be critical to success.

2. A minimum number of practice WISC administrations are essential in a course in individual intelligence testing.

3. A minimum number of scored protocols should be submitted for evaluation in a course in individual intelligence testing.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

TABLE PAGE

CHAPTER I

THE PROBLEM AND DEFINITION OF TERMS

Introduction

The entire field of individual intelligence testing is fraught with controversy. Fitzgibbon (1973) stated:

> There is little question that testing today is in a position of visibility and public concern such as it has not experienced throughout most of its prior history on the American educational and social scene. . . . It is difficult to recall a period in which test results have been so surrounded by controversy, or when they have generated such intense reaction, either positive or negative, from large segments of the citizenry (p. 1).

During the summer of 1965, people carried pickets and marched in front of the national headquarters of the American Psychological Association in Washington to protest educational testing. The staffs of several congressmen and senators prepared for hearings in which serious charges against certain testing practices would be heard, and the Equal Employment Opportunity Commission was trying to evaluate hundreds of written complaints which involved testing in employment selection and promotion cases (Holmen and Docter, 1972).

Various charges have been brought against the testing industry. Among these are the objections that: (a) psychological tests represent an invasion of privacy, (b) they present a problem of confidentiality, (c) the test content itself

lacks validity, (d) their predictive ability is questionable, (e) they are unfair to culturally disadvantaged groups, (f) they permanently classify individuals, and (g) the labels given to people are misleading (Anastasi, 1967). McClelland spoke of numerous social and psychological implications of testing (1973), and Eysenck's study of racial differences in intelligence (1971) caused considerable controversy, but any cursory review of recent testing literature should convince the reader that the problems and issues are many.

Numerous aspects of intellectual functioning and testing have been studied, lending an increased understanding of various facets of this highly complex area. DuBois (1970) and Linden and Linden (1968) considered testing from an historical perspective, Guilford (1967) took a factor analytic approach, Maccoby (1962) viewed intelligence in light of cognitive abilities, and Edwards (1971, 1972), Matarazzo (1972), and Robb, Bernardoni, and Johnson (1972) attempted a general assessment approach. Clinical interpretations have been provided by Zimmerman and Woo-Sam (1973), Glasser and Zimmerman (1967), and Hewitt and Massey (1969). Ferinden and Jacobson (1969) concentrated on educational interpretation. Annual testing conferences have been held (Educational Testing Service, 1971), and standards and codes of ethics have been written (American Personnel and Guidance Association, 1972; and American Psychological Association, 1966) in order to improve test administration and use.

> Unfortunately, the problem of setting standards for the training and experience expected of test users has not been resolved either through professional associations or through state legislation. We feel that groups attempting to improve

> testing practice in this country should give first priority to the definition of technical competencies essential for various levels of involvement with testing procedures. When these competence standards for different levels of testing work are adequately defined, the next priority should be given to improving training programs and establishing new ones that can help people meet the standards (Holmen and Docter, 1972, p. 162).

In all of the literature cited above, only passing, and infrequent mention was made of the training of examiners in individual intelligence testing situations. Volumes have been written on intellectual functioning and testing, and yet, one of the most critical areas, the training of the examiner, has been neglected. Occasional calls for better training can be found (Zimmerman and Woo-Sam, 1972) and a few training manuals are available (Lutey, 1966; and Pedrini and Pedrini, 1970), but actual experimental studies involving the training of examiners are virtually nonexistent.

Individual intelligence testing is taught at nearly every institution offering graduate programs in psychology or guidance. The traditional, or conventional, method of teaching an individual intelligence testing course involved demonstration of the administration of the instrument by the instructor or other capable person, discussion of responses and the scoring system, practice sessions and review, and presentation of the rationale, constructs, and uses of the test. Little time has been available in class to allow the students to examine the controversial areas and problems in testing--the very areas they will undoubtedly be asked to defend if they continue in the testing field. If more efficient methods were discovered to be equally as effective as traditional methods of teaching students

to understand and administer the tests, more valuable class and instructor time could be spent on the value-laden aspects of testing.

Statement of the Problem

Individual intelligence testing is, today, being asked to account to the public for its practices. Much has been written regarding intellectual functioning and testing which provides responsible and defensible answers to the critics. However, little attention has been paid to the training of examiners, an aspect obviously relevant to professional and ethical practice. In traditional individual intelligence testing courses, most of the time has been devoted to teaching the students to understand and administer the test, thus leaving little or no time for examination of critical issues involving testing. If cognitive understanding and skill in administration could be learned independently by students, valuable class time could thus be used in other ways.

Purpose of the Study

The purpose of the study was to compare two methods of instruction in teaching graduate students to understand, administer, and score the Wechsler Intelligence Scale for Children (WISC). The objective was to obtain experimental data regarding the performance of students who learned to understand, administer, and score the WISC by means of independent study compared with students who learned these skills in a traditional way.

Implications

The results of the study could enable educators to use their time and talents more efficiently when teaching a course in individual intelligence testing, specifically the Wechsler Intelligence Scale for Children. If it could be demonstrated that there is no significant difference in either understanding of the WISC, or administration and scoring of the WISC, or both, between students taught by an independent study method and the traditional method, then instructors could justifiably replace or supplement their existing methods of instruction and include coverage of vital and critical testing issues.

Further, this study should make a contribution to a field in which there is little literature, experimental or otherwise. Individual intelligence testing has been a part of graduate school curricula for nearly half a century and, yet, few investigations have been conducted regarding the efficacy of instruction.

Definition of Terms

Cognitive Examination Score

The score attained by an individual on the experimenter-made paper-and-pencil test of understanding of such aspects of the WISC as administration, scoring, validity, reliability, and interpretation.

EPPS

Edwards Personal Preference Schedule.

Independent Study Method

Cognitive understanding of the WISC and administration and scoring proficiency of the WISC were attained by means of cassette tapes of lectures, handouts, textual readings, videotape viewings of a WISC administration, a full-day WISC administration workshop, and independent practice of WISC administration and scoring.

Performance Examination Average

The score attained after the 12 performance examination item scores, or ratings, were averaged.

Traditional Method

Cognitive understanding of the WISC and administration and scoring proficiency of the WISC were attained in regularly scheduled meetings in which the WISC was administered by the instructor or other capable person, WISC responses and the WISC scoring system were discussed, and the rationale, constructs, and uses of the test were presented and discussed. Fifteen WISC administrations and protocols, and five WISC observations were required of each subject.

Undergraduate Grade Point Average

The average grade earned during the last 60 undergraduate hours.

WISC

Wechsler Intelligence Scale for Children.

Hypotheses

In order to determine specifically the difference between the two groups, the following null hypotheses were formulated and tested:

1. There is no significant difference between the cognitive examination scores of the Experimental and Control groups.

2. There is no significant difference between the performance examination averages of the Experimental and Control groups.

3. There are no significant differences between the performance examination item scores of the Experimental and Control groups.

4. There are no significant relationships between age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, number of WISC tests administered, number of protocols handed in for evaluation, and the cognitive examination scores of the Experimental group.

5. There are no significant relationships between age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, number of WISC tests administered, number of protocols handed in for evaluation, and the performance examination averages of the Experimental group.

6. There is no significant relationship between the two outcome measures, namely, the cognitive examination scores and the performance examination averages of the Experimental group.

7. There are no significant relationships between age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, and the cognitive examination scores of the Control group.

8. There are no significant relationships between age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, and the performance examination averages of the Control group.

9. There is no significant relationship between the two outcome measures, namely, the cognitive examination scores and the performance examination averages of the Control group.

10. There are no significant relationships between age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, and the cognitive examination scores of the combined group.

11. There are no significant relationships between age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, and the performance examination averages of the combined group.

12. There is no significant relationship between the two outcome measures, namely, the cognitive examination scores and the performance examination averages of the combined group.

Assumptions

It was assumed that all subjects honestly answered the questions on the Edwards Personal Preference Schedule (EPPS). They were initially told that their EPPS scores would be used for no purposes other than to determine initial differences between the groups and for later statistical analyses (see Appendix A), but the possibility of a "halo" effect, nevertheless, existed. During the construction of the EPPS, an attempt was made to minimize the influence of social desirability in responses to the statements (Edwards, 1959, pp. 5-6), and this fact made the assumption of honest responses more plausible.

It was further assumed that the subjects' best effort was obtained on both the cognitive and performance examinations. The assumption of uniform testing conditions was made. The researcher personally proctored the cognitive examination sessions of both groups, and an attempt was made to keep the conditions as uniform, and distraction-free, as possible. The researcher was present in the building, but not in the room, at all final performance examination sessions. An effort was made to see that the rooms were free at the scheduled times and no unnecessary interruptions were allowed to occur.

Finally, it was assumed that the subjects in the Experimental group honestly reported the number of WISC tests they administered and the number of WISC administration observations they made. They were assured in the initial handout (see Appendix A), and again when they were asked for their totals during the final class session, that these totals would have no effect upon their grade or course evaluation.

Limitations

The major limitation of the study was considered to be the composition of the groups themselves. Due to class scheduling difficulties, no random assignment of subjects to treatment conditions was possible. The Control group was composed of students which the Guidance Department at the University of New Mexico called the "Block Program." These students enrolled in all their courses together and were separated only in special instances or for "elective" courses. Once enrolled, none of these students dropped out of the course or the program. By the time of the experiment, they had already spent a summer session together.

The Experimental group class was scheduled to meet one evening a week and the majority of the students held full-time jobs. Most of them had not known each other prior to the experiment. The group began the experiment with 14 subjects, but 3 eventually dropped the course, leaving only 11 subjects in the group. One cited "personal problems" as the reason for dropping, and the researcher was unable to contact the other 2.

The researcher attempted to control statistically for any possible initial differences between the groups, but the lack of random assignment to treatment conditions was, nevertheless, considered to be a major limitation of the study.

The results of this study pertained only to the Wechsler Intelligence Scale for Children (WISC). It was beyond the scope of the study to determine if the conclusions could be generalized to the teaching of other individual intelligence testing courses, namely, to other Wechsler scales or to the Stanford-Binet.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

Summaries of research with individual intelligence tests (Zimmerman and Woo-Sam, 1972), and summaries of research on teaching of psychology (Birney and McKeachie, 1955), were readily available, but, in spite of an exhaustive search, no references were found relating to the teaching of an individual intelligence testing course. Calls for better training of examiners were frequent (Miller, Chansky, and Gredler, 1970), but experimental studies of teaching approaches were non-existent.

Not only is there a paucity or lack of studies on the teaching of individual intelligence testing, but the problems of the prospective teacher of such courses are further confounded by the lack of research in areas such as test administration, interpretation. Womer and Wahi (1969) stated:

> The great majority of the research reported in professional journals deals with the tests themselves and with their contributions to educational and psychological research, rather than with the actual administration and scoring of the tests, with test-program development, or with the use and interpretation of the resulting scores. This is not to say there is any paucity of written how-to-do-it statements on test administration or test-program development or test interpretation. There are many, but not all of them are supported by relevant research evidence (p. 1461).

Faced with the lack of published studies on (a) the teaching of individual intelligence testing, and (b) test administration, scoring, etc., instructors have had to devise their own course outlines and guidelines for administration. Two of these eventually were published: Lutey's Individual intelligence testing: A manual (1966), and The Pedrini supplementary aid to the administration of the Stanford-Binet Intelligence Scale (Form L-M): A handbook (1970), but neither contained any reference to experimental studies on teaching upon which the authors based their suggestions. For example, the Pedrini handbook stated: "Training should include 30 to 35 trial tests, carefully supervised (p. 1)." In a letter to this researcher Lutey wrote: "I have no knowledge of any published (or unpublished) formal approaches to teaching individual intelligence testing courses other than what I have presented in my manual (1974)." She continued: ". . . it is very possible that I have missed such information."

Descriptions of measurement courses at various institutions were found (Baker, 1967; and Karmel, 1968), but in only three instances were there specific references to an individual intelligence testing course. Reilley (1969) showed a 30-minute videotape of a WISC administration to 8 graduate students. His discussion of the results of the study consisted exclusively of an elaboration of their reactions to the experience. They were "unanimously favorable (p. 234)." He concluded the article by stating:

> The majority of published reports on the use of video tapes in counselor education is concerned with the counseling function. Very few articles in the professional literature deal with the development of other skills such as testing. Perhaps greater research and discussion in this area is past due (p. 234).

Faced with the lack of published studies on (a) the teaching of individual intelligence testing, and (b) test administration, scoring, etc., instructors have had to devise their own course outlines and guidelines for administration. Two of these eventually were published: [illegible] Intelligence testing: A manual (1966), and the [illegible] supplementary aid to the administration of the Stanford-Binet Intelligence Scale (Form L-M): A workbook (1970), but neither contained any reference to experimental studies or teaching upon which the authors based their suggestions. For example, the [illegible] handbook stated, "Training should include 20 to 25 test trials, carefully supervised (p. 1)." In a letter to this researcher, [illegible] wrote, "I have no knowledge of any published (or unpublished) formal approaches to teaching individual intelligence testing courses other than what I have presented in my manual (1970)." She continued, ". . . it is very possible that I have missed such information."

Descriptions of measurement courses at various institutions were found (Baker, 1967; and Carroll, 1968), but in only three instances were there specific references to an individual intelligence testing course. Kelley (1969) described 30-minute videotapes of a WISC administration to 8 graduate students. His discussion of the results of the study consisted exclusively of an elaboration of their reactions to the experience. They were "unanimously favorable (p. 274)." He concluded the article by stating:

> The majority of published reports on the use of video-tapes in counselor education is concerned with the counseling function. They [illegible] also in the professional literature dealt with the development of other skills such as testing. Perhaps greater investigational attention in this area is overdue (p. 274).

Apart from Reilley's descriptive study, two experimental studies were found, but, in both cases, programmed instruction, rather than the test itself, was the major concern. Burson (1966) stated his purpose as being that of investigating programmed instruction as a method of teaching at the graduate level. He wrote programs which were designed to reduce scoring errors in computing chronological age and to facilitate scoring of vocabulary answers on the Stanford-Binet. The group that used the two programs also attended class. This group was compared with another group which was exposed only to classroom instruction. He found no significant differences between the groups on: (a) vocabulary test scoring errors, and (b) chronological age computational errors.

McGrew, Marcia, and Wright (1966) designed a branching program on the Wechsler Adult Intelligence Scale (WAIS). Students were exposed to various combinations of the program, a televised lecture, and a mimeographed text over the same material. A pretest and a posttest were administered. It was found that the programmed medium had no special superiority over other instructional media. The study was conducted in a two-hour class period, and was part of a Tests and Measurements course. The average time to complete the program was 25 minutes.

Thus, no experimental studies were found that compared different approaches to teaching an entire course especially designed to prepare individual intelligence test examiners. The remainder of this review of literature will concentrate upon other issues related to the question at hand.

Examiner Influence on Test Scores

The literature related to examiner influence in testing situations clearly demonstrated the need for better preparation of examiners. The studies discussed below were fairly conclusive in documenting an experimenter bias or influence in individual intelligence testing.

Sattler and Theye (1967) derived a number of conclusions from their review of research: (a) departures from standard procedure are more likely to affect specialized groups than normal groups, (b) susceptibility to situational variables, especially discouragement, appears more frequently with younger than older subjects, (c) rapport is an important variable in the examiner-subject relationship, (d) examiner influences appear, but reasons for the differences are not clear, (e) examiner experience is usually not a critical variable, and (f) while White examiners may have some subtle deleterious effect on Negro subjects' scores, the evidence is not definitive.

Anastasi (1968) stated: "There is considerable evidence that test results may vary systematically as a function of the examiner (p. 572)." Other investigators also found significant differences between examiners in terms of test results obtained (Cohen, 1965; Sattler, Hillix, and Neher, 1970; Schwartz and Flanigan, 1971; and Smith and May, 1967). Dana and Dana (1969), however, suggested that the examiner-bias effect was weak and may not generalize across tasks, examiners, and subjects. Contrary to the conclusions reached by Sattler and Theye, Auffrey and Robertson (1972) found that the reliability of scoring was significantly affected by scorer's experience level. Cieutat's (1965) examiners

Examiner Influence on Test Scores

The literature related to examiner influence in testing situations clearly demonstrated the need for more investigation of examiners. The studies discussed below are of considerable importance in the assessment of experimenter bias or influence in individual intelligence testing.

Sattler and Theye (1967) derived a number of conclusions from their review of research. (a) departures from standard procedures are more likely to affect specialized groups than normal groups, (b) susceptibility to situational variables, especially discouragement, appears more frequently with younger than older subjects, (c) rapport is an important variable in the examiner-subject relationship, (d) examiner influences appear, but reasons for the differences are not clear, (e) examiner experience is usually not a critical variable, and (f) while White examiners may have had some deleterious effect on Negro subjects' scores, the evidence is not definitive.

Anastasi (1968) stated, "There is considerable evidence that test results may vary systematically as a function of the examiner (p. 37)." Other investigators also found significant differences between examiners in terms of test results obtained (Cohen, 1965; Miller, Holtz, and Baker, 1970; Schwartz and Flanigan, 1971; and Smith and May, 1967). Dunn and Dunn (1969), however, suggested that the examiner-bias effect was weak and may not generalize across tests, examiners, and subjects. Contrary to the conclusions reached by Sattler and Theye, Anthony and Richardson (1972) found that the reliability of scoring was significantly affected by scorer's experience level. Egeland's (1969) examined tested 243 child subjects of both sexes, and he found significant differences among examiners. He noted a cluster of 2 or 3 "high testers," 1 "low tester," and 4 or 5 in the mid-range. The sex of the subject was not a significant factor, but female examiners elicited significantly higher IQ's than male examiners. Examiners elicited higher means from the subjects not of their own sex.

Specific examiner behaviors were found to have significant effects upon the performance of individuals being tested. In a study by Sattler and Martin (1971), 1192 male sixth graders were tested and anxious and nonanxious examiner roles significantly affected performance of the subjects. Significant differences were found when the examiners were approving, disapproving, and neutral (Bornstein, 1967; and Witmer, Bornstein, and Dunham, 1971). WISC scores were obtained from 116 Puerto Rican children, and higher performance levels occurred with examiner behavior that encouraged active participation, verbalization, and repeated effort on the part of the subject (Thomas, Hertzig, Dryman, and Fernandez, 1971). Sattler (1969) found that cues affected performance on specific subtests, and Exner (1966) discovered that rigid treatment on the part of the examiner affected specific subtests.

The studies discussed above focused upon the effect of the examiner on the subject's performance, but Masling (1959) also found the reverse to be true. That is, subject performance had a significant effect upon the examiner's behavior. He coached two subjects to behave in warm and cold ways during Wechsler-Bellevue test administrations. Analysis of audio tapes of the testing situations indicated consistency of either warm or cold behavior on the part of

the subjects. The results indicated that in scoring the responses, the examiners tended to be more lenient with the warm subject than with the cold subject. The warm subject was reinforced more often and given more opportunity to clarify or correct responses. "There was no mistaking the impact of the warm and cold conditions on the examiners (p. 338)."

Not only does the examiner influence the outcome in individual intelligence testing situations, but so also does the manner in which the examiner scores the test protocol. Warren and Brown (1973) checked 240 WISC and Stanford-Binet protocols and found that 37% of them contained scoring errors that affected the reported IQ. Schwartz (1966) found gross scoring differences between judges, primarily on those items which were underinquired. Miller, Chansky, and Gredler (1970) asked 32 trainees to score a WISC protocol, and they found that the Full Scale IQ's ranged from 76 to 93. They urged training institutions to provide extensive experience in all phases of scoring. In another study, Miller and Chansky (1972) called the need "urgent."

General Outcome Research

Numerous studies comparing two or more methods of instruction at the college level have been conducted. Virtually every available approach to teaching has been compared with every other approach, and the results appeared to be fairly conclusive:

> The consensus of studies made since 1920 is that no one mechanical teaching device, in and of itself, is better than another. Teaching by lecture, recitation, discussion, tutorial, reading-study, reading-quiz, cor-

the subjects. The results indicated that in scoring the responses, the examiners tended to be more lenient with the warm subject than with the cold subject. The warm subject was credited more often and given more opportunity to clarify or correct responses. There was no attempt at assessing the impact of the warm and cold conditions on the examiners (p. 259)."

Not only does the examiner influence the outcome in individual intelligence testing situations, but it also does the method in which the examiner scores the test protocol. Warren and Brown (1972) checked 240 WISC and Stanford-Binet protocols and found that 89% of them contained scoring errors that affected the reported IQ. Schwartz (1966) found great scoring differences between judges, primarily on those items which were underqualified. Miller, Chansky, and Gredler (1970) asked 32 trainees to score a WISC protocol, and they found that the Full Scale IQ's ranged from 76 to 93. They urged training institutions to provide extensive experience in all phases of scoring. In another study, Miller and Chansky (1972) called the need "urgent."

General Outcome Research

Numerous studies comparing two or more methods of instruction at the college level have been conducted. Virtually every available approach to teaching has been compared with every other approach, and the results appeared to be fairly conclusive:

> The comparison of studies made since 1924 is that no one method of teaching methods, in and of itself, is better than another. "Teaching by lecture, recitation, discussion, tutorial, audiovisually, reading-desk, corr-

> respondence, or several different laboratory methods . . . has not been demonstrated to be intrinsically better than some other technique (Hatch and Bennet, 1960, p. 10).

Dubin and Taveggia (1968) undertook an analysis of college teaching methods. They examined the data from 91 comparative studies conducted between 1924 and 1965 and concluded:

> These data demonstrate clearly and unequivocally that there is no measurable difference among truly distinctive methods of college instruction when evaluated by student performance on final examinations (p. 35).

In comprehensive reviews of college teaching methods, the conclusions of Eckert and Neale (1965), McKeachie (1963), and Travers (1969) were similar. In still another extensive review, Stephens (1967) found that methods which stressed visual aids or individualized instruction seemed to have a slight, but irregular, advantage. "Between the other methods, however, there was little to choose. One method was about as good as the other and no better (p. 72)."

Carlson's early (1953) study comparing the lecture and discussion methods indicated no significant outcome differences, and similar conclusions have been reached many times since (Weinick, 1960). Many now consider this to be a dead issue. Atherton (1971, 1972) compared lecture, discussion, and independent study methods and found no significant differences between the three methods on the three outcome measures of recall, understanding, and application.

Class size appeared to make no difference in terms of performance on a final examination (Nachman and Opochinsky, 1958). The comprehensive review of teaching technology written by Cohen, Clare, and Trent (1973) indicated that

when the effectiveness of media was compared with conventional methods, namely the lecture, no significant differences were found between experimental and control groups on standardized achievement tests or final examinations. On the other hand, Allen (1970) concluded, after his media review, that "the results of . . . studies showed, almost without exception, decided advantages for films and other audio-visual materials over the usual kinds of classroom instruction (p. 6)."

Undoubtedly, the one approach to teaching which consistently facilitated higher outcome ratings when compared with other teaching methods was the Keller plan. This method was formally presented by Keller (1968) and has since received much attention in the research literature. It was essentially an independent-study, mastery-oriented, approach to instruction, based upon small-interval contingency management. It was first used almost exclusively in the field of psychology, but outcome literature has since appeared in other disciplines, particularly science (Kulik, Kulik, and Carmichael, 1974). "In published studies, content learning under the Keller plan always equals, and most often exceeds, content learning under the lecture method (p. 383)." Studies which used this approach will appear later in the Outcome Research in Psychology section.

Tape Recorded Lectures

The Experimental group in the study conducted by this researcher received the majority of its instruction by means of tape recorded lectures. In the comparable studies reviewed, no significant final examination differences were found between the groups taught by tape recorder as opposed to conventional methods when the effectiveness of each was compared with conventional methods, namely the lecture, no significant differences were found between experimental and control groups on standardized achievement tests or final examinations. On the other hand, Allen (1970) concluded, after his media review, that "the results of studies showed, almost without exception, decided advantages to films and other audio visual materials over the usual kinds of classroom instruction (p. 8)."

Undoubtedly, the one approach to teaching which consistently facilitated higher outcome ratings when compared with other teaching methods was the Keller plan. This method was formally presented by Keller (1968) and has since received much attention in the research literature. It was essentially an independent-study, mastery-oriented, approach to instruction, based upon small-interval contingency management. It was first used almost exclusively in the field of psychology, but outcome literature has since appeared in other disciplines, particularly science (Kulik, Kulik, and Carmichael, 1974). "In published studies, content learning under the Keller plan always equals, and most often exceeds, content learning under the lecture method (p. 580)." Studies which used this approach will appear later in the Outcome Research in Psychology section.

Tape Recorded Lectures

The Experimental group in the study conducted by this researcher received the majority of its instruction by means of tape recorded lectures. In the comparable studies reviewed, no significant final examination differences were found between the groups taught by tape recorder as opposed to conventional methods

(Jenkins, 1972; Menne, Hannum, Klingensmith, and Nord, 1969; and Michael, 1965). Popham also found no significant outcome differences between his two groups in each of his two studies, but his research design was not directly comparable to the present study. In his first experiment (1961), the instructor met with the students for 20 minutes for a discussion after they had heard the tape recorded lecture. The students met as a group to listen to the tape. He held this variable constant in his second study (1962), but, rather than the instructor meeting with the students to discuss the tape, they met alone to discuss the tape and had no formal instructor contact during the experiment.

Class Attendance

Class attendance was considered to be a crucial variable in the present study, and, thus, literature related to this phenomenon was reviewed. The results of the review were mixed. The self-directed study group in an experiment conducted by Himmel (1972) performed significantly better than did the traditionally taught control group on the final examination. There was no difference between the group hearing taped lectures and the group listening to live lectures on three outcome measures in a study conducted by Menne, Hannum, Klingensmith, and Nord (1969). Caro (1961), earlier, came to a similar conclusion, although his only outcome measure was a final course examination. Partridge (1968) used 48 subjects and 4 treatment groups in his study and concluded:

> The experiment indicates the importance of lecture participation in promoting student achievement. Dialogue between teacher and student appears to support

> student development of certain intellectual skills and abilities unobtainable by reading the text alone (p. 6).

This statement appeared to receive support from others. Kelly and Thibaut (1969) stated:

> Research by social psychologists has shown that the acquisition and exercise of these various skills (perceptual, behavioral, and cognitive)* is affected by the sheer presence of other persons (p. 2).

McKeachie (1967) concluded a summary of teaching research by commenting:

> If students are to achieve application, critical thinking, or some higher cognitive outcomes, a reasonable assumption is that they should have an opportunity to practice application and critical thinking and to receive feedback on the results (p. 217).

Outcome Research in Psychology

There appeared to be a high level of interest among educators in the field of psychology in effective instruction (APA, Division on the Teaching of Psychology, 1973; and MacLeod, 1965, 1971). Seminar classes in the teaching of psychology have been conducted for prospective teachers (Ericksen, 1968), and descriptions of instructional procedures for teachers to follow were found (Ferster, 1968). The area of highest research interest, currently, appeared to revolve around the Keller plan of individualized, contingency-management, instruction. There was no shortage of experiments in which the students taught

---

* illustration of skills inserted by this researcher; taken from preceeding paragraph (p. 2)

> student development of certain intellectual skills and abilities unobtainable by reading the text alone (p. 8).

This statement seems at to receive support from others . Kelly and Thibaut (1969) stated:

> Research by social psychologists has shown that the acquisition and exercise of these various skills (perceptual, behavioral, and cognitive)* is affected by the mere presence of other persons (p. 3).

McKeachie (1967) concluded a summary of teaching research by commenting:

> If students are to achieve application, critical thinking, or some higher cognitive outcomes, a reasonable assumption is that they should have an opportunity to practice application and critical thinking and to receive feedback on the results (p. 217).

Ongoing Research in Psychology

There appeared to be a high level of interest among educators in the field of psychology in effective instruction (APA, Division on the Teaching of Psychology, 1972, and MacLeod, 1965, 1971). Seminar classes in the teaching of psychology have been conducted for prospective teachers (Ericksen, 1968), and descriptions of instructional procedures for teachers to follow were found (Forster, 1969). The area of highest research interest, currently, appeared to revolve around the Keller plan of individualized, contingency-management instruction. There was no shortage of essay items, in which the students taught

---

*Dimensions of skills touched by this researcher, taken from preceding paragraph (p. 2).

according to this method performed significantly higher on examinations than conventionally-taught students (Born, Gledhill, and Davis, 1972; Cooper, 1973; Cooper and Greiner, 1971; Morris and Kimbrell, 1972; and Withers and Kent, 1972).

Hartley's (1972) annotated bibliography of 49 comparative approaches to teaching psychology contained 29 references to programmed courses and none to individual intelligence testing courses. Despite the popularity of the Keller plan, and variations thereof, no studies were found which applied this method of instruction to courses in individual intelligence testing. Only in the area of the teaching of individual projective testing, did there appear to be experimental research and interest in program development (Jackson and Wohl, 1966; and Shemberg and Keeley, 1970).

Generally speaking, the psychology outcome literature reached the same conclusions as did the general literature in the area of comparative instructional methods (Dowell, 1967; and Robson, 1961). That was, one particular teaching method doesn't consistently and significantly yield higher student performance on examinations. Himmel (1970) suggested that self-directed study methods were as useful in promoting information-learning as traditional methods.

One thread emerged from the literature, however, that merited further exploration. When the outcome indicator was content acquisition as measured by formal examinations, teacher-structured instructional methods appeared to be superior to informal, less-structured, approaches (Coop, 1968; Coop and Brown, 1970; Cunningham, 1964; McKeachie, Lin, Forrin, and Teevan, 1960; and Marr,

Plath, Wakeley, and Wilkins, 1960). McKeachie (1968), however, questioned this somewhat narrow approach to learning and offered, as a summary to current outcome literature:

> . . . when our goals go beyond student acquisition of knowledge to education in the highest sense, small classes, vital discussions, and interpersonal interaction become much more important to education (p. 555).
>
> . . . the research indicates the superiority of discussion over lecture in developing problem-solving skills and probably in affecting motivation and attitudes (p. 552).

Pascal (1971) experimentally validated McKeachie's statements.

Personality and Learning

In the present study, Hypotheses 4, 5, 7, 8, 10, and 11 were designed to determine if any particular learner characteristic was significantly related to higher performance on the two outcome measures. The literature was fairly conclusive in establishing the existence of learning styles (Briggs, 1968; Gropper and Kress, 1965; and Tallmadge and Shearer, 1969), or the idea that personality has an influence upon learning.

Nearly 20 years ago (1955), Mills concluded that different people learn most efficiently by different teaching methods and that the need was to get away from the concept of a "best method" for teaching. McKeachie (1958) discussed the importance of motivation and expectations to learning, and Burton (1967) stated: "The learning process and achievement are materially affected by the level of aspiration set by the learner (p. 105)." MacLeod (1969) mentioned that the require-

Platt, Wakeley, and Kollasch, 1970). McKeachie (1963), however, questioned this somewhat narrow approach to learning and offered, as a summary to current outcome literature:

> . . . when our goals go beyond student acquisition of knowledge to education in the highest sense, small classes, vital discussion, and face-to-face personal interaction become much more important to education (p. 355).
>
> . . . the research indicates the superiority of discussion over lecture in developing problem-solving skills and probably in affecting motivation and attitudes (p. 358).

Bacal (1971) experimentally validated McKeachie's statements.

Personality and Learning

In the present study, hypotheses 4, 5, 7, 8, 10, and 11 were designed to determine if any particular learner characteristic was significantly related to higher performance on the learner outcome measures. The literature was fairly conclusive in establishing the existence of learning styles (Briggs, 1968; Cropper and Kress, 1965; and Fellenbridge and Shearer, 1969), or the idea that personality has an influence upon learning.

Nearly 20 years ago (1950), Wills concluded that different people learn most efficiently by different teaching methods and that the need was to get away from the concept of a "best method" for teaching. McKeachie (1958) discussed the importance of motivation and expectations to learning, and Burton (1962) stated: "The learning process and achievement are materially affected by the level of aspiration set by the learner (p. 103)." Anderson (1959) mentioned that the relationments of the subject and the characteristics of the students must be taken into consideration when selecting a teaching approach: "There is no fixed formula (p. 48)."

Bigelow and Egbert (1968) discovered that the personality traits of intellectual efficiency and responsibility were pertinent to independent study success. Beach (1960) studied the concept of sociability and its relationship to academic achievement. He discovered that there was no difference in achievement between the more and less sociable students in his independent study group. The less sociable students achieved more than the more sociable students in the lecture situation. In the small groups, the more sociable students achieved more than the less sociable students. After an extensive review of the literature related to teaching methods, Cohen, Clare, and Trent (1973) concluded;

> . . . investigations concerning the interaction between teaching method and student characteristics indicated that students who achieved most in conventional lecture situations were characterized by moderate achievement and social needs and low creativity. Students characterized by high creativity or by high social needs tended to perform best in small discussion groups (p. 1028).

This investigator used the Edwards Personal Preference Schedule (EPPS) to yield a measurement of personality, and cognitive examination scores and performance examination averages as the dependent variables in the correlation procedures. Hakel (1966) used scholastic aptitude test scores and two grade-point indices as dependent measures, and found no instances in which any EPPS scale correlated significantly with any of the criterion measures. No studies were found which were somewhat more comparable to the WISC investigation.

sonts of the subject and the characteristics of the students need to be taken into consideration when selecting a teaching approach. There is no fixed formula (p. 40)."

Stanley and Clark (1968) discovered that the personality traits of intellectual efficiency and personal ability were pertinent to independent study success. Beach (1960) studied the concept of sociability and its relationship to academic achievement. He discovered that there was no difference in achievement between the more and less sociable students in his independent study group. The less sociable students achieved more than the more sociable students in the lecture situation. In the small groups, the more sociable students achieved more than the less sociable students. After an extensive review of the literature related to teaching methods, Cohen, Glore, and Best (1973) concluded:

> Investigations concerning the interaction between teaching method and student characteristics indicated that students who achieved most in conventional lecture situations were characterized by moderate achievement and social needs and low creativity. Students characterized by high creativity or by high social needs tended to perform best in small discussion groups (p. 1828).

This investigator used the Edwards Personal Preference Schedule (EPPS) to yield a measurement of personality, and cognitive examination scores and performance examination averages as the dependent variables in the correlation procedures. Hakel (1966) used scholastic aptitude test scores and two grade-point indices as dependent measures, and found no instances in which any EPPS scale correlated significantly with any of the criterion measures. No studies were found which were somewhat more comparable to the EPPS investigation.

Summary

Why another "method" study, when the literature appears to indicate that one method of instruction is generally as effective as any other in influencing student achievement on a final examination?

First, no studies were found to provide evidence that this apparent conclusion could be generalized when comparing two methods or approaches to teaching an individual intelligence testing course. In no studies reviewed was the subject matter similar to that in the present study, and, thus, inference was not considered possible.

Secondly, in no studies reviewed was the performance of a prescribed skill (i.e., intelligence test administration) an issue. No researchers chose a performance examination specifically designed to measure skill in a prescribed task situation as an outcome measure. No evidence was found that would enable this researcher to make a prediction based upon experimental data.

The evidence indicated that the simple fact of learning together would have a different effect upon the group attending class as opposed to those not meeting during the course of the experiment. The specific effect, however, was not known, and in the case of the present study, could not be determined in advance.

The evidence suggested that students learning by tape recorded lectures performed no differently on final examinations than students who attended class. In the WISC study, however, students in the Experimental group attended a full-day WISC workshop. This served to provide some peer group and instructor con-

tact during the course of the experiment. Further, as mentioned previously, they were also evaluated on a performance, as well as a cognitive, examination. Thus, the related literature was considered to be only indirectly comparable.

The evidence relating to an experimenter influence on test scores was decisive. Psychological factors and scoring errors were found to have an influence upon obtained IQ scores. Based upon the evidence, this investigator inferred a need for better, and more data-based methods of examiner preparation.

Finally, the research suggested that the time had come to call a moratorium on comparative methodological studies, and look more toward finding approaches suited to individual learning styles and personalities. Six of the hypotheses tested in this study were designed to correlate personality variables with the two outcome measures. This researcher, however, was not inclined to concur with the general evidence and call a halt to comparative studies involving methods of instruction in teaching individual intelligence testing courses when it appeared that the beginning had not yet been reached.

not during the course of the experiment. Further, as mentioned previously, they were also evaluated on a performance, as well as a cognitive, examination. Thus, the related literature was considered to be only indirectly comparable.

The evidence relative to socioeconomic influences on test scores was decisive. Psychological factors that seem to arise were found to have an influence upon obtained I.Q. scores. Based upon the evidence, this investigator inferred a need for better, and more individualized methods of examinee preparation.

Finally, the research suggested that the time had come to call a moratorium on comparative methodological studies, and look more toward finding approaches suited to individual learning styles and personalities. Six of the hypotheses tested in this study were designed to correlate personality variables with the two outcome measures. This researcher, however, was not inclined to concur with the general evidence and call a halt to comparative studies involving methods of instruction in teaching individual intelligence testing courses when it appeared that the beginning had not yet been reached.

CHAPTER III

DESIGN OF THE STUDY

Introduction

The study was conducted at the University of New Mexico, Albuquerque, New Mexico. Two sections of Differential Diagnosis I (Guidance 512) were the focus of the study. The experiment was conducted during the first ten weeks of the Fall semester, 1973. The concentration during that time was upon the Wechsler Intelligence Scale for Children (WISC), administration, application, and interpretation. The remaining six weeks of the course were taught by the instructor, and consisted essentially of study and administration of the remaining Wechsler scales, other diagnostic instruments, and child case study preparation.

Subjects

The subjects were those students enrolled in two sections of Guidance 512 at the University of New Mexico during the Fall semester of 1973. Those students who enrolled in the Thursday evening section were designated as the Experimental group, and the students who enrolled in the Tuesday morning section were designated as the Control group. For the purposes of the experiment, the Experimental group was referred to as Group I, and the Control group was referred to as Group II. All were graduate students. All of the Control group,

and approximately half of the students in the Experimental group, had been admitted to the Masters degree program in the Guidance Department.

Many of the students in the Experimental group worked off-campus during the day and attended classes at night. Few had known each other before the experiment. The students in the Control group, on the other hand, knew each other previously, and were enrolled in "required" classes together. The program was referred to as the "Block Program" by the Guidance Department. The students had been admitted to the Department the previous June and, at the time of the experiment, had completed a summer session in classes together.

It was not possible to randomly assign students to control and experimental treatments. Because of this limitation in the design of the study, an attempt was made to ascertain initial differences between the groups on 18 variables. The variables selected were age, sex, undergraduate grade point average, and the 15 personality scales of the Edwards Personal Preference Schedule.

The mean, standard deviation, and variance for each variable was computed for each group, and an F_{max} value was computed to determine the homogeneity of the variances of each variable. In every case, the necessary 5.78 tabled F value was not reached, and there was no significant difference between any of the variances. Thus, the pooled variance t model was selected in order to determine initial differences between the groups.

In Table 1, the means of age, sex, and undergraduate grade point average for the Experimental and Control groups were compared. There were no differences found between the means of the groups with reference to these three variables.

and approximately half of the students in the Experimental group, had been admitted to the Masters degree program in the Guidance Department.

Many of the students in the Experimental group worked off-campus during the day and attended classes at night. Few had known each other before the experiment. The students in the Control group, on the other hand, knew each other previously, and were enrolled in "blocked" classes together. The program was referred to as the "Block Program" by the Guidance Department. The students had been admitted to the Department the previous June and, at the time of the experiment, had completed a summer session in classes together.

It was not possible to randomly assign students to control and experimental treatments. Because of this limitation in the design of the study, an attempt was made to ascertain initial differences between the groups on 19 variables. The variables selected were age, sex, undergraduate grade point average, and the 15 personality scales of the Edwards Personal Preference Schedule.

The mean, standard deviation, and variance for each variable was computed for each group, and an F_{max} value was computed to determine the homogeneity of the variances of each variable. In every case, the necessary F_{max} tabled value was not reached, thus there was no significant difference between any of the variances. Thus, the pooled variance t model was selected in order to determine initial differences between the groups.

In Table 1, the means of age, sex, and undergraduate grade point average for the Experimental and Control groups were compared. There were no differences found between the means of the groups with reference to these three variables.

TABLE 1

COMPARISONS BETWEEN THE MEANS OF AGE, SEX, AND UNDERGRADUATE GRADE POINT AVERAGE FOR THE EXPERIMENTAL AND CONTROL GROUPS

VARIABLE	GROUP	N	MEAN	S	S^2	F	T
Age	Exp	11	28.45	8.43	71.06	2.42	0.34
	Cont	14	27.50	5.42	29.34		
Sex (Female = 1 Male = 2)	Exp	11	1.27	0.47	0.22	1.21	-0.78
	Cont	14	1.43	0.51	0.26		
Grade Point Average	Exp	11	3.14	0.41	0.17	1.48	1.45
	Cont	14	2.92	0.34	0.11		

t = 2.069 needed for significance at .05 level with df = 23

Table 2 describes the EPPS variables and the manifest needs associated with each. In Table 3, the means of the 15 variable scores for the Experimental and Control groups were compared. Significant differences were found between the means of the personality variables of Deference, Order, and Autonomy. The Experimental group was found to have significantly higher scores on the variables of Deference, that is, to be more conforming and conventional, and Order, that is, to be more organized and systematic. The Control group's mean Autonomy score was found to be significantly higher than was the mean of the Experimental group. They were, as a group, more independent and non-conforming. No significant differences were found between the means of the other twelve personality variables.

The average age of the Experimental group was 28.45 (range = 22 to 50), compared with 27.50 for the Control group (range = 22 to 41). There were eight females and three males in the Experimental group, compared with eight females and six males in the Control group.

Both groups began the experiment with 14 members, but 3 of the Experimental group dropped the class. One dropped on the day of her final performance examination and cited "personal problems" as the reason. The other two dropped the course and withdrew from the University during the first two weeks. The researcher was unable to contact them.

Table 2 [illegible]

with each. In Table [illegible]

and Control groups were [illegible]

the means of the [illegible]

Experimental group was [illegible]

of Delinquency, that is, [illegible]

is, to be more [illegible]

score was found to be [illegible]

group. They were, [illegible]

nificant differences were [illegible]

variables.

The average age [illegible]

compared with 27.50 [illegible]

eight females and [illegible]

females and six males in [illegible]

Both groups [illegible]

mental group dropped the [illegible]

examination and [illegible]

the course and withdrew [illegible]

researcher was unable [illegible]

TABLE 2

THE EDWARDS PERSONAL PREFERENCE SCHEDULE VARIABLES
AND THE MANIFEST NEEDS ASSOCIATED WITH EACH

1. ach Achievement: To do one's best, to be successful, to accomplish tasks requiring skill and effort, to be a recognized authority, to accomplish something of great significance, to do a difficult job well, to solve difficult problems and puzzles, to be able to do things better than others, to write a great novel or play.

2. def Deference: To get suggestions from others, to find out what others think, to follow instructions and do what is expected, to praise others, to tell others that they have done a good job, to accept the leadership of others, to read about great men, to conform to custom and avoid the unconventional, to let others make decisions.

3. ord Order: To have written work neat and organized, to make plans before starting on a difficult task, to have things organized, to keep things neat and orderly, to make advance plans when taking a trip, to organize details of work, to keep letters and files according to some system, to have meals organized and a definite time for eating, to have things arranged so that they run smoothly without change.

4. exh Exhibition: To say witty and clever things, to tell amusing jokes and stories, to talk about personal adventures and experiences, to have others notice and comment upon one's appearance, to say things just to see what effect it will have on others, to talk about personal achievements, to be the center of attention, to use words that others do not know the meaning of, to ask questions others cannot answer.

5. aut Autonomy: To be able to come and go as desired, to say what one thinks about things, to be independent of others in making decisions, to feel free to do what one wants, to do things that are unconventional, to avoid situations where one is expected to conform, to do things without regard to what others may think, to criticize those in positions of authority, to avoid responsibilities and obligations.

6. aff Affiliation: To be loyal to friends, to participate in friendly groups, to do things for friends, to form new friendships, to make as many friends as possible, to share things with friends, to do things with friends rather than alone, to form strong attachments, to write letters to friends.

7. int Intraception: To analyze one's motives and feelings, to observe others, to understand how others feel about problems, to put one's self in another's place, to judge people by why they do things rather than by what they do, to analyze the behavior of others, to analyze the motives of others, to predict how others will act.

8. suc Succorance: To have others provide help when in trouble, to seek encouragement from others, to have others be kindly, to have others be sympathetic and understanding about personal problems, to receive a great deal of affection from others, to have others do favors cheerfully, to be helped by others when depressed, to have others feel sorry when one is sick, to have a fuss made over one when hurt.

9. dom Dominance: To argue for one's point of view, to be a leader in groups to which one belongs, to be regarded by others as a leader, to be elected or appointed chairman of committees, to make group decisions, to settle arguments and disputes between others, to persuade and influence others to do what one wants, to supervise and direct the actions of others, to tell others how to do their jobs.

10. aba Abasement: To feel guilty when one does something wrong, to accept blame when things do not go right, to feel that personal pain and misery suffered does more good than harm, to feel the need for punishment for wrong doing, to feel better when giving in and avoiding a fight than when having one's own way, to feel the need for confession of errors, to feel timid in the presence of superiors, to feel inferior to others in most respects.

11. nur Nurturance: To help friends when they are in trouble, to assist others less fortunate, to treat others with kindness and sympathy, to forgive others, to do small favors for others, to be generous with others, to sympathize with others who are hurt or sick, to show a great deal of affection toward others, to have others confide in one about personal problems.

12. chg Change: To do new and different things, to travel, to meet new people, to experience novelty and change in daily routine, to experiment and try new things, to eat in new and different places, to try new and different jobs, to move about the country and live in different places, to participate in new fads and fashions.

13. end Endurance: To keep at a job until it is finished, to complete any job undertaken, to work hard at a task, to keep at a puzzle or problem until it is solved, to work at a single job before taking on others, to stay up late working in order to get a job done, to put in long hours of work without distraction, to stick at a problem even though it may seem as if no progress is being made, to avoid being interrupted while at work.

14. het Heterosexuality: To go out with members of the opposite sex, to engage in social activities with the opposite sex, to be in love with someone of the opposite sex, to kiss those of the opposite sex, to be regarded as physically attractive by those of the opposite sex, to participate in discussions about sex, to read books and plays involving sex, to listen to or to tell jokes involving sex, to become sexually excited.

15. agg Aggression: To attack contrary points of view, to tell others what one thinks about them, to criticize others publicly, to make fun of others, to tell others off when disagreeing with them, to get revenge for insults, to become angry, to blame others when things go wrong, to read newspaper accounts of violence.

TABLE 3

COMPARISONS BETWEEN THE MEANS OF THE SCORES OF
THE EDWARDS PERSONNAL PREFERENCE SCHEDULE
FOR THE EXPERIMENTAL AND CONTROL GROUPS

VARIABLE	GROUP	N	MEAN	S	S^2	F	T
Achievement	Exp	11	16.27	3.98	15.82	2.06	1.05
	Cont	14	14.86	2.77	7.67		
Deference	Exp	11	12.64	2.91	8.46	1.14	2.14*
	Cont	14	10.21	2.72	7.41		
Order	Exp	11	11.00	4.00	16.00	2.05	3.37*
	Cont	14	6.43	2.79	7.80		
Exhibition	Exp	11	13.64	3.70	13.65	4.71	-0.46
	Cont	14	14.14	1.70	2.90		
Autonomy	Exp	11	13.73	3.72	13.82	1.38	-2.20*
	Cont	14	17.36	4.36	19.02		
Affiliation	Exp	11	15.27	3.04	9.32	2.09	-0.65
	Cont	14	16.29	4.39	19.30		

TABLE 3 (continued)

VARIABLE	GROUP	N	MEAN	S	S^2	F	T
Intraception	Exp	11	19.18	3.89	15.16	1.18	-0.32
	Cont	14	19.71	4.23	17.91		
Succorance	Exp	11	10.73	4.65	21.62	2.84	0.30
	Cont	14	10.29	2.76	7.61		
Dominance	Exp	11	11.91	5.01	25.09	1.10	-1.10
	Cont	14	14.07	4.78	22.84		
Abasement	Exp	11	11.64	3.91	15.26	1.24	1.36
	Cont	14	9.36	4.34	18.86		
Nurturance	Exp	11	18.55	3.75	14.07	2.04	0.85
	Cont	14	16.93	5.36	28.69		
Change	Exp	11	19.27	3.58	12.82	1.23	-0.15
	Cont	14	19.50	3.98	15.81		
Endurance	Exp	11	12.82	4.31	18.57	1.18	1.14
	Cont	14	10.93	3.97	15.76		

TABLE 3 (continued)

VARIABLE	GROUP	N	MEAN	S	S^2	F	T
Heterosexuality	Exp	11	12.91	6.19	38.29	2.44	-1.94
	Cont	14	16.86	3.96	15.67		
Aggression	Exp	11	10.46	3.42	11.68	1.28	-1.92
	Cont	14	12.93	3.03	9.15		

$t = 2.069$ needed for significance at .05 level with $df = 23$

Procedures

The researcher attended the first class meeting of both groups and explained the procedures that each was to follow during the course of the experiment. A prepared handout detailing these procedures was given to each group and discussed (see Appendix A). Each student was asked to sign a consent form allowing the researcher to check the student's records in order to determine his/her undergraduate grade point average (see Appendix B). Procedures for observing WISC demonstrations at the College of Education training facility, Manzanita Center, were discussed, and each student was issued a pass, allowing him/her to observe the confidential sessions through one-way mirrors.

The students in both groups took the Edwards Personal Preference Schedule (EPPS) during the first class session. They were assured both verbally and in the handout (see Appendix A) that the scores would be used for no purposes other than for ascertaining initial differences between the groups, and for later statistical analyses. Each student was provided with an EPPS question booklet and an IBM answer sheet, which the researcher later scored by hand. The instrument was administered according to the directions in the EPPS Manual (Edwards, 1959, p. 6). The students asked to be given the results, and the researcher consented to attend a meeting of each class at the end of the experiment and discuss the results with them. This took place, in each case, after the students had taken the final cognitive examination. Each person was given a profile of his/her scores and a description of the 15 EPPS scales (see Table 2).

Procedures

The researcher attended the first class meeting of both groups and explained the procedures that each was to follow during the course of the experiment. A prepared handout detailing these procedures was given to each group and discussed (see Appendix A). Each student was asked to sign a consent form allowing the researcher to check the student's records in order to determine his/her undergraduate grade point average (see Appendix B). Procedures for observing WISC demonstrations at the College of Education training facility, Monteville Center, were discussed, and each student was issued a pass, allowing him/her to observe the confidential sessions through one-way mirrors.

The students in both groups took the Edwards Personal Preference Schedule (EPPS) during the first class session. They were assured both verbally and in the handout (see Appendix A) that the scores would be used for no purpose other than for ascertaining initial differences between the groups, and for later statistical analyses. Each student was provided with an EPPS question booklet and an IBM answer sheet, which the researcher later scored by hand. The instrument was administered according to the directions in the EPPS Manual (Edwards, 1959, p. 6). The students asked to be given the results, and the researcher consented to attend a meeting of each class at the end of the experiment and discuss the results with them. This took place, in each case, after the students had taken the final cognitive examination. Each person was given a profile of his/her scores and a description of the 15 EPPS scales (see Table 2).

The researcher had to make a slight alteration in regard to the initial intention of determining an undergraduate grade point average for each student. Many of the subjects had attended several institutions which often had very different methods of determining student achievement, and the process of calculating the grade point average became difficult, and frequently, unreliable. The Graduate School had computed an average based on the last 60 undergraduate hours of coursework for each student, and thus, to insure a comparable frame of reference, these were used as a measure of prior academic achievement.

The experiment, itself, lasted from August 27, 1973, until November 2, 1973. During the final week of the study (October 29 to November 2), the groups each met during the regularly-scheduled class times to take a paper-and-pencil examination. The researcher proctored the administration of the examination, and an attempt was made to keep the testing situations uniform and distraction-free. After all the students had completed the examination, anonymous comments and reactions to the previous ten week experience were requested. Relevant excerpts from these can be found in Appendix G.

Also during the final week of the study, the subjects each administered a WISC at Manzanita Center. The researcher and the graduate assistant assigned to the course observed each demonstration through a one-way mirror. The available times were posted, and each student made an appointment with the secretary at the Center. Each person was responsible for bringing his/her own subject. The procedures for scheduling, test check-out, purchase of materials, and protocol checks by the graduate assistant were described in the initial handouts (see Ap-

pendix A). The researcher was present in the building during all final demonstrations, and an effort was made to see that rooms were available when scheduled and that distractions during testing were minimal.

During the intervening weeks of the experiment, the two groups followed different schedules, and these will be discussed separately in the following sections.

Experimental Group

After an initial meeting on August 30, 1973, and a workshop on Saturday, September 8, the group did not meet again until the evening of the final cognitive examination, November 1, 1973. The workshop lasted from 8:30 AM until 4:30 PM, and the entire focus was administration of the WISC. (A schedule for the day can be found in Appendix C.) The morning was spent on the Verbal sub-tests, and the instructor demonstrated the test administration with an eight year old child. After the demonstration, the students broke into dyads and one of the pair acted as the subject, and the other as the examiner. A discussion and question-and-answer session followed. WISC kits were provided for each dyad. The concentration during the afternoon was on the Performance subtests. The same child subject returned (the researcher had taken him home after the morning demonstration), and the instructor demonstrated the administration of the Performance subtests. Following this, the dyads reversed, with the morning "subject" acting as the examiner, and vice-versa. The workshop ended with a discussion and final questioning.

pendix A). The researcher was present in the building during all these demonstrations, and an effort was made to see that rooms were available when scheduled and that distractions during testing were minimal.

During the intervening weeks of the experiment, the two groups followed different schedules, and these will be discussed separately in the following sections.

Experimental Group

After an initial meeting on August 30, 1973, and a workshop on Saturday, September 8, the group did not meet again until the evening of the final cognitive examination, November 1, 1973. The workshop lasted from 8:30 AM until 4:30 PM, and the entire focus was administration of the WISC. (A schedule for the day can be found in Appendix C.) The morning was spent on the Verbal sub-tests, and the instructor demonstrated the test administration with an eight year old child. After the demonstration, the students broke into dyads and one of the pair acted as the subject, and the other as the examiner. A discussion and question-and-answer session followed. WISC kits were provided for each dyad. The concentration during the afternoon was on the Performance subtests. The same child subject returned (the researcher had taken him home after the morning demonstration), and the instructor demonstrated the administration of the Performance subtests. Following this, the dyads reversed, with the morning "subject" acting as the examiner, and vice-versa. The workshop ended with a discussion and final questioning

Other than the workshop experience, the students had no formal contact with the instructor or the graduate assistant assigned to the course. However, they were free to contact these people for conferences or feedback on their own initiative. Handouts were distributed and textbooks were assigned at the first class meeting. The instructor prepared and recorded 10 1-hour lectures on cassette audio tapes, and 8 copies of each tape were available for check-out and use. A cassette tape recorder could be checked out for a 24-hour period. They could keep each tape one week. The Guidance office closed at 5:00 PM daily, and, since several of the people worked until that time, one set of 10 tapes was left in the Learning Materials Center in the College of Education. This was open in the evenings, and a cassette recorder was available for use there.

A videotaped administration of the WISC was shown at three different times during the semester. In order to allow the maximum number of students to see the tape if they wished, 12 noon, 3:00 PM, and 7:00 PM on two separate days were chosen. The researcher was present at all three showings in order to operate the machine and answer questions.

The final demonstration was the only "required" WISC administration, but others were encouraged. They could ask the graduate assistant to evaluate and give feedback on scored protocols, and they were free to schedule a WISC demonstration at Manzanita Center at any time they wanted the graduate assistant's feedback on their administration. They were encouraged to observe others administering the test at the Center, or elsewhere.

These procedures, as they were communicated to the students, can be found in Appendix A (Group I handout).

Control Group

This group met with the instructor every Tuesday morning in the regularly-scheduled class meetings. The group was assigned the same texts as the Experimental group, and received the same handouts. The content that was presented by the instructor during the meetings was the same that was available to the Experimental group on the cassettes. The instructor did not attempt to limit the questioning and discussion evolving from the material as long as he was able to present the planned content in the course of the class period. WISC demonstrations were held during the class meetings, and this group did not attend the September 8 workshop. The 10 cassette tapes were not available to them.

The three videotaped WISC administration showings were open to this group, as well as the Experimental group, and they were encouraged, but not required, to attend. Each member was required to administer 15 WISC tests, and to submit the 15 scored protocols to the graduate assistant for checking. Each student was also required to observe five other people administering the WISC. The WISC videotape could suffice for one of these.

Along with the WISC demonstration at Manzanita Center during the final week of the experiment, they were expected to schedule another demonstration prior to September 28. They were told in the initial handout (see Appendix A, Group II handout) that the purpose of this would be threefold: (a) to allow for

These procedures, as they were communicated to the students, can be found in Appendix A (Group I handout).

<u>Control Group</u>

This group met with the instructor every Tuesday morning at the regularly scheduled class meetings. The group was assigned the same texts as the Experimental group, and received the same handouts. The content that was presented by the instructor during the meetings was the same that was available to the Experimental group on the cassettes. The instructor did not attempt to limit the questioning and discussion evolving from the material as long as he was able to present the planned content in the course of the class period. WISC demonstrations were held during the class meetings, and this group did not attend the September 8 workshop. The 10 cassette tapes were not available to them.

The three videotaped WISC administration showings were open to this group, as well as the Experimental group, and they were encouraged, but not required, to attend. Each member was required to administer 15 WISC tests, and to submit the 15 scored protocols to the graduate assistant for checking. Each student was also required to observe five other people administering the WISC. The WISC videotape could suffice for one of these.

Along with the WISC demonstration at Marcus Clinic during the final week of the experiment, they were instructed to schedule another demonstration prior to September 25. They were told in the initial handout (see Appendix A, Group II handout) that the purpose of this would be threefold: (a) to allow for feedback prior to the final demonstration, (b) to allow others to observe WISC administrations, and (c) to allow for reliability checks of the rating instrument. They brought their own subjects for the demonstrations. The graduate assistant was available for conferences after the demonstration, but often the students had to drive their subject home, and the feedback conference was held at another time.

Instrumentation

In order to acquire indices of achievement on the two outcome measures of (a) cognitive understanding, and (b) administration and scoring proficiency, the researcher constructed two evaluation instruments. These instruments will be discussed separately.

Cognitive Understanding

Before constructing the paper-and-pencil instrument, the researcher listened to the ten cassette tapes, and read the assigned texts and handouts. After a universe from which the items were to be chosen was outlined, the items were then constructed in representative proportion to that universe. The sample items were evaluated by the instructor and the researcher's other two committee members. Irrelevant, duplicative, or misleading items were then eliminated. Fifty-four items were considered appropriate and were retained in the final instrument (see Appendix D). The researcher took all test material only from that which was available to the Experimental group in the study. In order not to let related discussion material invalidate the instrument for the Experimental group, no class feedback prior to the final [illegible]

administrations, and (c) to allow [illegible]

they brought their own subjects [illegible]

was available for conferences prior [illegible]

to drive their subject home, and [illegible]

time.

[illegible]

In order to acquire indices [illegible]

of (a) cognitive understanding, and [illegible]

the researcher constructed two evalu[illegible]

discussed separately.

Cognitive Understanding

Before constructing the [illegible]

listened to the ten cassette tapes, and [illegible]

After a universe from which the [illegible]

were then constructed in representative [illegible]

Items were evaluated by the [illegible]

members. Irrelevant, duplicative [illegible]

four items were considered [illegible]

(see Appendix D). The researcher [illegible]

available to the [illegible]

caution material [illegible]

meetings of the Control group, other than the initial one, were attended by the researcher.

The instrument was not formally pilot tested because no suitable population was available. The course had been taught previously in the Guidance Department, but it had differed each time, depending upon instructor or focus. Since the instructor assembled the material and made the tapes specifically for the experiment, no population with similar exposure was available, and, thus, the final cognitive examination was, in a sense, the pilot test of the instrument.

After the students took the test, a count of the responses to each possible answer was made of the 54 items for Experimental, Control, and combined groups. This count, with the percent of the total responding to each answer, can be found in Appendix E.

For the purposes of statistical analyses, the total number of correct responses to the questions was determined for each student. It was from these figures that group means were calculated.

Administration and Scoring Proficiency

No suitable instrument for measuring WISC administration and scoring proficiency was found, and, thus, the researcher constructed the one that can be found in Table 4. The instrument was designed for use by one, or more, observer(s) who merely circled along the 15-point continuum the level of proficiency observed. Items 1 through 9 were completed immediately following the test administration. Items 10, 11, and 12 were completed after the WISC protocol had been scored meetings of the Control group, other than the initial one, were attended by the researcher.

The instrument was not formally pilot tested because no suitable population was available. The course had been taught previously in the Guidance Department, but it had differed each time, depending upon instructor or focus. Since the instructor assembled the material and made the tapes specifically for the experiment, no population with similar exposure was available, and, thus, the final cognitive examination was, in a sense, the pilot test of the instrument.

After the students took the test, a count of the responses to each possible answer was made of the 56 items for Experimental, Control, and combined groups. This count, with the percent of the total responding to each answer, can be found in Appendix E.

For the purposes of statistical analyses, the total number of correct responses to the questions was determined for each student. It was from these figures that group means were calculated.

Administration and Scoring Proficiency

No suitable instrument for measuring WISC administration and scoring proficiency was found, and, thus, the researcher constructed the one that can be found in Table 4. The instrument was designed for use by one, or more, observer(s) who merely circled along the 10 point continuum the level of proficiency observed. Items 1 through 9 were completed immediately following the test administration. Items 10, 11, and 12 were completed after the WISC protocol had been scored

TABLE 4

WISC OBSERVATION INSTRUMENT

Examiner ________________
Subject ________________
TEST ADMINISTERED FOR OBSERVATION __________ Rated by ________________
Date ________________

(circle one)

1. Adherence to standardization procedure	stand. pro. adhered to	stand. pro. generally adhered to	much deviation from stand. pro.
	15 14 13 12 11	10 9 8 7 6 5	4 3 2 1
2. Maintenance of eye contact with subject	eye contact effectively maintained	some eye contact	little eye contact
	15 14 13 12 11	10 9 8 7 6 5	4 3 2 1
3. Establishment of rapport with subject	excellent rapport	some rapport	little rapport
	15 14 13 12 11	10 9 8 7 6 5	4 3 2 1
4. Overall organization of materials	materials well organized	moderately organized; some fumbling	much fumbling for materials
	15 14 13 12 11	10 9 8 7 6 5	4 3 2 1
5. Presentation of verbal directions to subject	directions clearly and accurately given	some lack of clarity and accuracy	directions were unclear and/or inaccurate
	15 14 13 12 11	10 9 8 7 6 5	4 3 2 1
6. Appropriateness of questioning subject's responses	appropriate and timely questioning	some lack of appropriateness or timeliness in questioning	inappropriate questioning or lack of questioning
	15 14 13 12 11	10 9 8 7 6 5	4 3 2 1

TABLE 4 (continued)

7. Presentation of performance subtests	efficiently presented 15 14 13 12 11	moderately well presented 10 9 8 7 6 5	poorly presented; fumbling 4 3 2 1
8. Effective use of stop-watch for timed subtests	started and stopped at appropriate times 15 14 13 12 11	sometimes started and stopped at appropriate times 10 9 8 7 6 5	rarely started and stopped at appropriate times 4 3 2 1
9. Composure of examiner during test administration	relaxed; assured 15 14 13 12 11	moderate assuredness; some hesitancy 10 9 8 7 6 5	hesitating; ill at ease 4 3 2 1
10. Effectiveness of recording subject's responses	accurately recorded subject's responses 15 14 13 12 11	some inaccuracy in recording 10 9 8 7 6 5	responses poorly recorded 4 3 2 1
11. Accuracy of scoring procedures	accurate scoring 15 14 13 12 11	scoring moderately accurate 10 9 8 7 6 5	scoring poor 4 3 2 1
12. Accuracy of conversion	accurate conversion procedures 15 14 13 12 11	some accuracy in conversion procedures 10 9 8 7 6 5	poor conversion procedures 4 3 2 1

by the examiner and submitted to the observer.

Inter-rater Reliability. The initial testing of the instrument was conducted during the month of September, 1973, when all members of the Control group were required to administer a WISC at Manzanita Center for demonstration. Three of the Experimental group members also chose to administer a WISC during this time. Approximately half of the administrations were watched by two observers, and during the remaining administrations, three observers were present. The initial instrument contained only 5 response ratings, but due to problems of variance, which will be discussed later, the instrument was expanded to include 15 response options for the final demonstrations.

Kendall's rank correlation coefficient (tau, τ) was the statistic chosen to determine measures or degrees of association between two sets of ranks (Bradley, 1968). Due to the small number of response options (5) and, thus, the frequent lack of variance, the statistic was selected because of its ability to allow tied rankings to be randomly broken. The corrected tau coefficients of correlation for the 12 item scores and an overall performance examination tau can be found in Table 5.

Table 5 also indicates measures of the relationships among the rankings of three observers. The statistic chosen to determine the degree of association among them was Kendall's coefficient of concordance (W) (Siegal, 1956). In order to calculate both tau and W, it was first necessary to convert the ratings made by the two and three observers, respectively, to rankings.

by the examiner and submitted

Inter-rater Reliability.

ducted during the month of September,

group were required to administer

Three of the Experimental group

this time. Approximately half of

observers, and during the remaining

The initial instrument contained

variance, which will be discussed

15 response options for the final

Kendall's rank correlation

to determine measures or degrees

1968). Due to the small number of

lack of variance, the statistic was

rankings to be randomly broken.

the 12 item scores and an overall

Table 5.

Table 5 also indicates

of three observers. The statistic

among them was Kendall's coefficient

order to calculate both the

made by the two and three observers

TABLE 5

RANK CORRELATION COEFFICIENTS (TAU) BETWEEN TWO OBSERVERS AND COEFFICIENTS OF CONCORDANCE (W) AMONG THREE OBSERVERS FOR 12 ITEM SCORES AND OVERALL PERFORMANCE EXAMINATION COEFFICIENTS

VARIABLE	CORRECTED TAU	W
Item 1	.82	.76**
Item 2	.87	.63
Item 3	.38	.84**
Item 4	.60	.71*
Item 5	.87	.40
Item 6	.82	.33
Item 7	.59	.90**
Item 8	.94	.67*
Item 9	.65	.61
Item 10	.52	.53
Item 11	.94	.85**
Item 12	.88	.42
Overall	.73	.64

tau = no specific level of significance due to corrections
W = 64.4 needed for significance at .05 level*
75.6 needed for significance at .01 level**

The item tau coefficients ranged from .38 for Item 3 to .94 for Item 8. The Overall tau was .73. Six of the coefficients of concordance were significant at the .05 level (Items 1, 3, 4, 7, 8, and 11), and four (Items 1, 3, 7, and 11) were significant at the .01 level. The item W coefficients ranged from .33 for Item 6 to .90 for Item 7. The Overall W was .64 which approached, but did not reach, significance at the .05 level.

The use of three observers, compared with two, did not appear to increase the coefficients of inter-rater reliability substantially, and a decision was made to use two observers during the final performance examinations. Due to the small amount of variance and the frequency of tied decisions by the observers, the corrected tau statistic was necessary with the 5-point rating scale with two observers. The use of the correction was judged to be too arbitrary and imprecise for future use, and, thus, the rating scale was expanded to 15 response options.

During the month of October, the two observers held practice and discussion sessions which were designed to increase their operational knowledge of the instrument. Both observers were present in the observation booth during all final demonstrations, and each took notes during the WISC administrations. Immediately following all demonstrations, each observer independently marked Items 1 through 9, basing the decisions upon his or her notes. After the student scored the protocol, he/she gave it to the graduate assistant who evaluated it and marked Items 10 through 12 on the instrument. He then gave the protocol to the researcher who, without knowledge of his evaluations, marked Items 10, 11, and 12.

Spearman's rank correlation coefficient (r_s) (Tuckman, 1972) was the statistic chosen to determine the level of agreement between the two observers during the final WISC administrations. All of the ratings for each item assigned by each observer were first changed to rankings before the coefficients were computed. The coefficients can be found in Table 6.

All of the correlation coefficients were significant at the .01 level, indicating a high level of inter-rater reliability. The item coefficients ranged ranged from .73 for Item 5 to .99 for Items 10 and 12. The Overall coefficient of correlation was .88.

Internal Consistency. After the final performance examination data were collected, internal consistency of the instrument was determined. Coefficients of correlation between the 12 item scores and the performance examination averages for the Experimental, Control, and combined groups were calculated. The BMDO2D program was used. This information can be found in Table 7.

For the experimental group, all of the 12 performance item scores showed significant correlations with the performance examination averages. Only one correlation, between Item 8 (Effective use of stopwatch) and the performance examination averages, was significant at the .05 level. The other eleven item scores were significant at the .01 level. A high relationship was demonstrated between performance on the individual items and average examination performance. Each item appeared to be able to predict the performance examination averages to a high degree.

Spearman's rank correlation coefficient (r_s) (Duckworth, 1972) was the statistic chosen to determine the level of agreement between the two observers during the final WISC administrations. All of the ratings for each item assigned by each observer were first changed to rankings before the coefficients were computed. The coefficients can be found in Table 6.

All of the correlation coefficients were significant at the .01 level, indicating a high level of inter-rater reliability. The item coefficients ranged from .73 for Item 3 to .97 for Items 10 and 12. The Overall coefficient of correlation was .98.

Internal Consistency. After the final performance examination data were collected, internal consistency of the instrument was determined. Coefficients of correlation between the 12 item scores and the performance examination averages for the Experimental, Control, and combined groups were calculated. The BMDO2D program was used. This information can be found in Table 7.

For the experimental group, all of the 12 performance item scores showed significant correlations with the performance examination averages. Only one correlation, between Item 8 (Effective use of stopwatch) and the performance examination averages, was significant at the .05 level. The other eleven item scores were significant at the .01 level. A high relationship was demonstrated between performance on the individual items and average examination performance. Each item appeared to be able to predict the performance examination averages to a high degree.

TABLE 6

RANK CORRELATION COEFFICIENTS (r_s) BETWEEN TWO OBSERVERS FOR 12 ITEM SCORES AND OVERALL FINAL PERFORMANCE EXAMINATION COEFFICIENT

VARIABLE	SPEARMAN (r_s) COEFFICIENTS
Item 1	.80**
Item 2	.94**
Item 3	.87**
Item 4	.88**
Item 5	.73**
Item 6	.84**
Item 7	.83**
Item 8	.86**
Item 9	.84**
Item 10	.99**
Item 11	.94**
Item 12	.99**
Overall	.88**

.377 needed for significance at .05 level*

.534 needed for significance at .01 level**

TABLE 7

CORRELATIONS BETWEEN ITEM SCORES AND PERFORMANCE EXAMINATION AVERAGES FOR THE EXPERIMENTAL, CONTROL, AND COMBINED GROUPS

	CORRELATIONS WITH PERFORMANCE EXAMINATION AVERAGES OF:		
VARIABLE	EXP. GROUP	CONTROL GROUP	COMB. GROUP
Item 1	.81**	.81**	.80**
Item 2	.89**	.41	.73**
Item 3	.82**	.38	.67**
Item 4	.88**	.92**	.85**
Item 5	.94**	.60*	.82**
Item 6	.85**	.53	.72**
Item 7	.81**	.62*	.75**
Item 8	.65*	.56*	.58**
Item 9	.94**	.79**	.90**
Item 10	.83**	.12	.61**
Item 11	.82**	.60*	.73**
Item 12	.80**	.53	.63**

Exp = .602 needed for significance at .05 level with df = 9*
.735 needed for significance at .01 level with df = 9**
Cont = .532 needed for significance at .05 level with df = 12*
.661 needed for significance at .01 level with df = 12**
Comb = .396 needed for significance at .05 level with df = 23*
.505 needed for significance at .01 level with df = 23**

For the Control group, three outcome measures correlated at the .01 level with the performance examination averages. Item 1 correlated .81, Item 4 (Overall organization of materials) correlated .92, and Item 9 (Composure of examiner during test administration) correlated .79. Four outcome measures correlated at the .05 level: Item 5 (Presentation of verbal directions), Item 7 (Presentation of Performance subtests), Item 8 (Effective use of stopwatch), and Item 11 (Accuracy of scoring procedures). Item 6 (Appropriateness of questioning) fell .004 short of being significant at the .05 level, and Item 12 (Accuracy of conversion procedures) was .006 short of significance. Thus, seven of the twelve item scores were found able, to a significant degree, to predict the performance examination averages. Two were a fraction of a percentage point from being significant, and three items were found to be no better than chance in predicting performance examination averages for the Control group.

For the combined group, all twelve of the performance examination item scores correlated at the .01 level with the performance examination averages. Each item showed a significant ability to predict the performance examination averages. Thus, an internal consistency was demonstrated within the instrument for the combined group.

All possible intercorrelations between item scores for the Experimental, Control, and combined groups can be found in Appendix F.

For the Control group, three outcome measures correlated at the .01 level with the performance examination averages. Item 1 correlated .81, Item 4 (Overall organization of materials) correlated .82, and Item 9 (Composure of examiner during test administration) correlated .79. Four outcome measures correlated at the .05 level: Item 5 (Presentation of verbal directions), Item 7 (Presentation of Performance subtests), Item 8 (Effective use of stopwatch), and Item 11 (Accuracy of scoring procedures). Item 6 (Appropriateness of questioning) fell .004 short of being significant at the .05 level, and Item 12 (Accuracy of conversion procedures) was .006 short of significance. Thus, seven of the twelve item scores were found able, to a significant degree, to predict the performance examination averages. Two were a fraction of a percentage point from being significant, and three items were found to be no better than chance in predicting performance examination averages for the Control group.

For the combined group, all twelve of the performance examination item scores correlated at the .01 level with the performance examination averages. Each item showed a significant ability to predict the performance examination averages. Thus, an internal consistency was demonstrated within the instrument for the combined group.

All possible intercorrelations between item scores for the Experimental, Control, and combined groups can be found in Appendix F.

Statistical Analyses

The data were analyzed by means of several statistical procedures. Initial differences between the groups were determined by means of a series of t tests. The BMDX70 program was selected for this purpose because of its ability to analyze data with more variables (36) than subjects (25). The BMDX70 data print-out gave a t estimate for both a pooled variance and a separate variance model. In all cases, the pooled variance t was selected because no F value (given in the print-out) reached the necessary 5.78 (13, 10) level of significance. In other words, there were no significant differences between the variances of any of the 18 measurements of initial differences.

The item frequency count for the combined group (see Appendix E) was done with the BMDO4D program. The count was done by hand for the Experimental and Control groups. The researcher did the percentage conversions for all groups by hand.

In the initial testing of the performance examination rating instrument, Kendall's rank correlation coefficient (tau, τ) (Bradley, 1968) was used to determine degrees of association between the ratings of two observers. Kendall's coefficient of concordance (W) (Siegal, 1956) was used for the same purpose with three observers.

Coefficients of inter-rater reliability were established during the final performance examinations. Two observers watched each WISC administration, and Spearman's rank correlation coefficient (r_s) (Tuckman, 1972) was used to determine the levels of agreement.

The BMDO2D program was used to determine internal consistency of the performance examination rating instrument. Intercorrelations between 36 variables for the Experimental, Control, and combined groups were determined by means of this same Biomedical program. A description of the variables and the three correlation matrices can be found in Appendix F.

Hypotheses 1, 2, and 3 were tested by means of an analysis of covariance procedure. The BMDO9V program was selected because of its ability to take three initial differences between the groups into consideration in determining differences in performance between the groups on the two outcome measures and on the 12 performance examination items.

Hypotheses 4 through 12 were tested by means of the BMDO2D correlation program. The use of this program resulted in determining the relationships between the 18 initial variables and the two outcome measures, and between the two outcome measures for the Experimental, Control, and combined groups. The three complete correlation matrices are found in Appendix F.

The BMD02D program was used to determine internal consistency of the performance examination rating instrument. Intercorrelations between 58 variables for the Experimental, Control, and combined groups were determined by means of this same biomedical program. A description of the variables and the three correlation matrices can be found in Appendix F.

Hypotheses 1, 2, and 3 were tested by means of an analysis of covariance procedure. The BMDX82 program was selected because of its ability to take these initial differences between the groups into consideration in determining differences in performance between the groups on the two outcome measures and on the 12 performance examination items.

Hypotheses 4 through 12 were tested by means of the BMD02D correlation program. The use of this program resulted in determining the relationships between the 18 initial variables and the two outcome measures, and between the two outcome measures for the Experimental, Control, and combined groups. The three complete correlation matrices are found in Appendix F.

CHAPTER IV

RESULTS AND DISCUSSION

Introduction

This chapter includes the testing of the hypotheses, analyses of the data, and a discussion of the results. The findings are summarized at the end of the chapter.

Comparisons Between the Means of the Outcome Measures

In this section, the differences between the two groups on the two outcome measures--the cognitive examination and the performance examination--were compared. The twelve items of the performance examination were examined collectively, as well as individually. Tests of initial differences between the groups indicated significantly higher scores on the EPPS variables of Deference and Order for the Experimental group, and Autonomy for the Control group (see Table 3). In determining the level of significance of the difference between the groups on the outcome measures, an analysis of covariance procedure, which allowed initial differences between the groups on these three variables to be taken into consideration, was used.

Hypothesis 1

There is no significant difference between the cognitive examination scores of the Experimental and Control groups.

The 1.3905 F value in Table 8 was not significant and, thus, no significant difference was found between the cognitive examination scores of the Experi-

TABLE 8

COMPARISON BETWEEN THE MEANS OF THE COGNITIVE EXAMINATION SCORES OF THE EXPERIMENTAL AND CONTROL GROUPS

SOURCE	SS	df	MS	F
Between	44.3347	1	44.3347	1.3905
Within	637.6584	20	31.8829	
Total	681.9931	21		

4.35 needed for significance at .05 level

mental and Control groups. Hypothesis 1 was not rejected. The mean of the Experimental group was 41.00, compared with 41.71 for the Control group. When adjusted with regard to initial differences between the groups on the EPPS variables of Order, Deference, and Autonomy, the means were 39.43 and 42.95, respectively. The range of correct items in the Experimental group was from 28 to 50 (standard

Hypothesis I

There is no significant difference between the mean cognitive examination scores of the Experimental and Control groups.

The 1.3905 F value in Table 8 was not significant; thus, no significant difference was found between the cognitive examination scores of the Experi-

TABLE 8

COMPARISON BETWEEN THE MEANS OF THE COGNITIVE EXAMINATION SCORES OF THE EXPERIMENTAL AND CONTROL GROUPS

SOURCE	SS	df	MS	F
Between	44.3347	1	44.3347	1.3905
Within	637.6584	20	31.8829	
Total	681.9931	21		

4.35 needed for significance at .05 level

mental and Control groups. Hypothesis I was not rejected. The mean of the Experimental group was 41.09, compared with 41.71 for the Control group. When adjusted with regard to initial differences between the groups on the EPPS variables of Order, Deference, and Autonomy, the means were [illegible] and [illegible], respectively. The range of correct items in the Experimental group was from [illegible] to [illegible] (standard deviation = 7.28), and from 34 to 47 in the Control group (standard deviation = 3.87).

Hypothesis 2

There is no significant difference between the performance averages of the Experimental and Control groups.

TABLE 9

COMPARISON BETWEEN THE MEANS OF THE PERFORMANCE EXAMINATION AVERAGES OF THE EXPERIMENTAL AND CONTROL GROUPS

SOURCE	SS	df	MS	F
Between	11.4762	1	11.4762	1.7988
Within	127.5984	20	6.3799	
Total	139.0746	21		

4.35 needed for significance at .05 level

No significant difference was found between the performance examination averages of the groups. The 1.7988 F value was not significant and, thus, Hypothesis 2 was not rejected. These data can be found in Table 9. The mean of the Experimental group was 8.55, compared with 8.86 for the Control group. The adjusted means were 7.72 and 9.51. The range of averages for the Experimental group was 1.83 to 12.58 (standard deviation = 3.36), and from 5.58 to 11.42

(standard deviation = 1.76) for the Control group. As mentioned previously in the Hypothesis 1 discussion, the range of cognitive examination scores was also greater for the Experimental group. In both instances, the Experimental subjects scored both higher and lower than did the subjects in the Control group.

In order to assure an acceptable level of WISC administration proficiency, subjects with performance examination averages below 8.00 were required by the instructor to administer another WISC for demonstration. Four people in the Experimental group (36%) fell into this category, compared with three (21%) in the Control group. The four people in the Experimental group raised their performance examination average from 4.75 to 8.29. The three subjects in the Control group raised their average from 6.58 to 9.72. For the purpose of the WISC experiment, however, the averages obtained during the examination week (October 29 to November 2) were used.

Hypothesis 3

There are no significant differences between the performance examination item scores of the Experimental and Control groups.

Sub-hypotheses 3A-L

There are no significant differences between the performance examination scores of the Experimental and Control groups on the following items lettered A-L below:

A - Adherence to standardization procedure
B - Maintenance of eye contact with subject
C - Establishment of rapport with subject
D - Overall organization of materials
E - Presentation of verbal directions to subject
F - Appropriateness of questioning subject's responses
G - Presentation of performance subtests

H - Effective use of stopwatch for timed subtests
I - Composure of examiner during test administration
J - Effectiveness of recording subject's responses
K - Accuracy of scoring procedures
L - Accuracy of conversion procedures

The data in Table 10 indicated no significant differences between the groups on all items, except one--Item 11, Accuracy of scoring procedures. The difference was in favor of the Control group. The 5.0217 F value was significant at the .05 level, and Hypothesis 3, Sub-hypothesis K, was rejected. Hypothesis 3, Subhypotheses A through J and L, were not rejected.

Before adjustments were made with regard to initial differences between the groups on the EPPS variables of Order, Deference, and Autonomy, the unadjusted means of Item 11 were 6.82 for the Experimental group, and 7.29 for the Control group. The unadjusted t value was -.36 with 2.069 needed for significance at the .05 level. After adjustments had been made with regard to initial differences between the groups, the means were 5.32 for the Experimental group, and 8.47 for the Control group. The adjusted t value was -2.2409. Only in this one instance, did the adjustments for initial differences between the groups appear to result in a significant difference between the means of the outcome measures.

Correlations Between Descriptive Measures and Outcome Measures and Between Outcome Measures of the Experimental Group

The discussion in this section will center around correlations, (1) between the 20 descriptive measures and the 2 outcome measures, and (2) between

H – Effective use of stopwatch for timed subtests
I – Correctness of vocabulary during test administration
J – Effectiveness of recording subject's responses
K – Accuracy of scoring procedures
L – Accuracy of conversion procedures

The data in Table 39 indicated no significant differences between the groups on all items, except one—Item 11, Accuracy of scoring procedures. The difference was in favor of the Control group. The 2.0217 t value was significant at the .05 level, and Hypothesis 9, Sub-hypothesis K, was rejected. Hypothesis 9, Subhypotheses A through J and L, were not rejected.

Before adjustments were made with regard to initial differences between the groups on the EPPS variables of Order, Deference, and Autonomy, the unadjusted means of Item 11 were 6.82 for the Experimental group, and 7.29 for the Control group. The unadjusted t value was -.36 with 2.069 needed for significance at the .05 level. After adjustments had been made with regard to initial differences between the groups, the means were 5.32 for the Experimental group, and 8.47 for the Control group. The adjusted t value was -2.2409. Only in this one instance, did the adjustments for initial differences between the groups appear to result in a significant difference between the means of the outcome measures.

Correlations Between Descriptive Measures and Outcome Measures and Between Outcome Measures of the Experimental Group

The discussion in this section will center around correlations, (1) between the 20 descriptive measures and the 2 outcome measures, and (2) between

TABLE 10

COMPARISONS BETWEEN THE MEANS OF THE PERFORMANCE EXAMINATION ITEM SCORES OF THE EXPERIMENTAL AND CONTROL GROUPS

SOURCE	SS	df	MS	F
A.				
Between	24.2716	1	24.2716	2.3981
Within	202.4223	20	10.1211	
Total	226.6939	21		
B.				
Between	17.5122	1	17.5122	1.0028
Within	349.2776	20	17.4639	
Total	366.7898	21		
C.				
Between	20.0178	1	20.0178	1.0678
Within	374.9421	20	18.7471	
Total	394.9599	21		

TABLE 10 (continued)

SOURCE	SS	df	MS	F
D.				
Between	5.5000	1	5.5000	0.4470
Within	246.1011	20	12.3051	
Total	251.6011	21		
E.				
Between	7.2283	1	7.2283	1.0148
Within	142.4577	20	7.1229	
Total	149.6860	21		
F.				
Between	3.0587	1	3.0587	0.3858
Within	158.5631	20	7.9282	
Total	161.6218	21		
G.				
Between	18.5445	1	18.5445	2.1023
Within	176.4192	20	8.8210	
Total	194.9637	21		

TABLE 10 (continued)

SOURCE	SS	df	MS	F
H.				
Between	1.8595	1	1.8595	0.1753
Within	212.1928	20	10.6096	
Total	214.0523	21		
I.				
Between	5.4054	1	5.4054	0.4383
Within	246.6291	20	12.3315	
Total	252.0345	21		
J.				
Between	8.7571	1	8.7571	0.6180
Within	283.4224	20	14.1711	
Total	292.1795	21		
K.				
Between	35.5790	1	35.5790	5.0217*
Within	141.6998	20	7.0850	
Total	177.2788	21		

TABLE 10 (continued)

SOURCE	SS	df	MS	F
L.				
Between	11.7141	1	11.7141	0.5045
Within	464.4082	20	23.2204	
Total	476.1223	21		

4.35 needed for significance at .05 level

the outcome measures, of the Experimental group. The descriptive measures of age, sex, grade point average for the last 60 undergraduate hours, 15 personality variable scores of the Edwards Personal Preference Schedule, number of WISC tests administered, and the number of WISC protocols handed in for evaluation, and the outcome measures of cognitive examination scores and performance examination averages will be considered. The complete correlation matrix, which shows all possible correlations between all variables, can be found in Appendix F. Intercorrelations between the descriptive measures can be found there. The correlations between the descriptive measures and the performance examination item scores are not included in this section, since, out of 216 correlations, only 7 significant correlations were found. This is 4 fewer than could be expected at the .05 level by chance. These correlations can be found in Appendix F. The reader is also referred to Appendix F for correlations between the cognitive examination scores and the performance examination item scores, and between the performance examination item scores.

Two additional descriptive measures, which are included in the matrix in Appendix F but which are not included in the discussion, are the number of WISC observations (students watching other students administer the test), and the number of WISC tests administered for demonstration. The number of observations (range = 1 to 6; mean = 4.18), and the number of demonstrations (range = 1 to 3; mean = 1.36) were considered too small to yield meaningful correlations.

The average number of WISC tests administered by the Experimental group was 10.27 (range = 4 to 17), and the average number of protocols handed the outcome measures, of the Experimental group. The descriptive measures of age, sex, grade point average for the last 60 undergraduate hours, 15 personality variable scores of the Edwards Personal Preference Schedule, number of WISC tests administered, and the number of WISC protocols handed-in for evaluation, and the outcome measures of cognitive examination scores and performance examination averages will be considered. The complete correlation matrix, which shows all possible correlations between all variables, can be found in Appendix F. Intercorrelations between the descriptive measures can be found there. The correlations between the descriptive measures and the performance examination item scores are not included in this section, since, out of 216 correlations, only 7 significant correlations were found. This is 4 fewer than could be expected at the .05 level by chance. These correlations can be found in Appendix F. The reader is also referred to Appendix F for correlations between the cognitive examination scores and the performance examination item scores, and between the performance examination item scores.

Two additional descriptive measures, which are included in the matrix in Appendix F but which are not included in the discussion, are the number of WISC observations (students watching other students administer the test), and the number of WISC tests administered for demonstration. The number of observations (range = 1 to 5, mean = 4.18), and the number of demonstrations (range = 1 to 2, mean = 1.38) were considered too small to yield meaningful correlations.

The average number of WISC tests administered by the Experimental group was 10.57 (range = 4 to 17), and the average number of protocols handed in for evaluation was 4.73 (range = 0 to 14). Three students handed in no protocols. Final tabulations indicated that all Control group members administered 15 WISC tests and handed in 15 protocols, observed 5 administrations, and demonstrated the 2 required times. The specifications of the experimental design were met by the Control group, and thus, it was only necessary to test differences on these four variables with regard to the Experimental group.

Hypothesis 4

There are no significant relationships between age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, number of WISC tests administered, number of protocols handed in for evaluation, and the cognitive examination scores of the Experimental group.

As can be seen in Table 11, significant correlations were found between the number of WISC tests administered and cognitive examination scores (.86), and the number of protocols handed in for evaluation and cognitive examination scores (.73). There was a high positive relationship between the number of WISC tests administered and the scores on the cognitive examination. A lower, but significant, relationship was found between the number of protocols handed in for evaluation and the cognitive examination scores. No significant relationships were found between any of the other descriptive measures and cognitive examination scores. Thus, with the exception of the number of WISC tests adin for evaluation was 4.73 (range ...

Final tabulations indicated that ... WISC ...

tests and handed in 13 protocols, ...

the 2 required times. The ...

the Control group, and thus, ...

four variables with regard to the ...

Hypothesis 4

There are no significant ...

grade point average, fifteen ...

Preference Schedule, number of ...

handed in for evaluation, and the ...

group.

As can be seen in Table 11, ...

the number of WISC tests admin...

and the number of protocols handed ...

scores (.73). There was a high ...

tests administered and the scores ...

significant, relationship was found ...

for evaluation and the cognitive ...

ships were found between any of the ...

examination scores. Thus, with the ...

TABLE 11

CORRELATIONS BETWEEN DESCRIPTIVE MEASURES AND COGNITIVE EXAMINATION SCORES OF THE EXPERIMENTAL GROUP

VARIABLE	CORRELATION WITH COGNITIVE EXAM SCORES
Age	.49
Sex	-.09
GPA	.07
Ach	.28
Def	.08
Order	.33
Exh	.00
Aut	-.55
Aff	-.16
Int	.26
Suc	.13
Dom	-.17
Aba	.27
Nur	.20
Chg	.14
End	.17
Het	.25
Agg	.36
# Adm	.86**
# Prot	.73*

.602 needed for significance at .05 level with df = 9*

.735 needed for significance at .01 level with df = 9**

ministered, and the number of protocols handed in for evaluation, Hypothesis 4 was not rejected.

Hypothesis 5

There are no significant relationships between age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, number of WISC tests administered, number of protocols handed in for evaluation, and the performance examination averages of the Experimental group.

The data in Table 12 indicated a significant correlation (.76) between the number of WISC test administered and the performance examination averages. There was a significant relationship between the number of WISC tests administered and the performance examination averages. There was a high (.58), but not statistically significant, relationship between the number of protocols handed in for evaluation and the performance examination averages. No other significant correlations were found between any of the other descriptive measures and the performance examination averages. With the exception of the one significant correlation found, Hypothesis 5 was not rejected.

Hypothesis 6

There is no significant relationship between the two outcome measures, namely, the cognitive examination scores and the performance examination averages of the Experimental group.

TABLE 12

CORRELATIONS BETWEEN DESCRIPTIVE MEASURES AND PERFORMANCE EXAMINATION AVERAGES OF THE EXPERIMENTAL GROUP

VARIABLE	CORRELATION WITH PERF EXAM AVERAGES
Age	.45
Sex	-.43
GPA	.00
Ach	.14
Def	.28
Order	.34
Exh	-.45
Aut	-.53
Aff	-.08
Int	.17
Suc	.08
Dom	-.35
Aba	.36
Nur	.31
Chg	.36
End	-.01
Het	-.15
Agg	-.30
# Adm	.76**
# Prot	.58

.602 needed for significance at .05 level with df = 9*
.735 needed for significance at .01 level with df = 9**

A significantly high (.86) correlation was found between the cognitive examination scores and the performance examination averages. The correlation was significant at the .01 level. A high degree of predictibility was demonstrated between the cognitive examination scores and the performance examination averages of the Experimental group. Hypothesis 6 was rejected.

Correlations Between Descriptive Measures and Outcome Measures and Between Outcome Measures of the Control Group

The discussion in this section will center around correlations, (1) between the 18 descriptive measures and the 2 outcome measures, and (2) between the outcome measures, of the Control group. The descriptive measures of age, sex, grade point average for the last 60 undergraduate hours, 15 personality variable scores of the Edwards Personal Preference Schedule, and the outcome measures of cognitive examination scores and performance examination averages will be considered.

The correlations between the descriptive measures and the performance examination item scores are not included in this section. Out of 216 correlations, 19 were found to be significant (8 more that could have been expected by chance). Eleven of these correlations were below .60. A correlation of .532 was needed for significance at the .05 level with df = 12.

Intercorrelations between the descriptive measures can be found in Appendix F. The reader is also referred to Appendix F for correlations between the cognitive examination scores and the performance examination item scores, and for intercorrelations between the performance examination item scores.

A significant [illegible]

examination scores and the [illegible]

was significant at the .05 level [illegible]

stated between the cognitive [illegible]

tion averages of the Experimental [illegible]

Correlations [illegible]

The discussion in [illegible]

the 18 descriptive measures [illegible]

come measures, of the [illegible]

point average for the last [illegible]

of the Edwards Personal Preference [illegible]

nitive examination scores and [illegible]

The correlations between [illegible]

examination item scores and [illegible]

19 were found to be significant [illegible]

Eleven of these correlations [illegible]

for significance at the .05 level [illegible]

Intercorrelations between [illegible]

Appendix F. The reader is [illegible]

the cognitive examination [illegible]

and for intercorrelations [illegible]

Hypothesis 7

There are no significant relationships between age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, and the cognitive examination scores of the Control group.

The data in Table 13 indicated significant negative relationships between age and the cognitive examination scores, and between the personality variable of Change and the cognitive examination scores. No similar correlations were noted in reference to the Experimental group (see Table 11). No other significant correlations were noted between the descriptive measures and the cognitive examination scores of the Control group. Thus, with the exception of significant negative correlations between age and the cognitive examination scores, and between the personality variable of Change and the cognitive examination scores, Hypothesis 7 was not rejected.

Hypothesis 8

There are no significant relationships between age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, and the performance examination averages of the Control group.

One significant correlation out of 18 was found between the descriptive measures and the performance examination averages. The personality variable of Nurturance correlated negatively at the .01 level with the performance examination averages. No other significant correlations were found between the

Hypothesis 7

There are no significant relationships between age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, and the cognitive examination scores of the Control group.

The data in Table 10 indicated significant negative relationships between age and the cognitive examination scores, and between the personality variable of Change and the cognitive examination scores. No similar correlations were noted in reference to the Experimental group (see Table 11). No other significant correlations were noted between the descriptive measures and the cognitive examination scores of the Control group. Thus, with the exception of significant negative correlations between age and the cognitive examination scores, and between the personality variable of Change and the cognitive examination scores, Hypothesis 7 was not rejected.

Hypothesis 8

There are no significant relationships between age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, and the performance examination averages of the Control group.

One significant correlation out of 18 was found between the descriptive measures and the performance examination averages. The personality variable of Nurturance correlated negatively at the .01 level with the performance examination averages. No other significant correlations were found between the

TABLE 13

CORRELATIONS BETWEEN DESCRIPTIVE MEASURES AND COGNITIVE EXAMINATION SCORES OF THE CONTROL GROUP

VARIABLE	CORRELATION WITH COGNITIVE EXAM SCORES
Age	-.58*
Sex	-.32
GPA	.44
Ach	-.09
Def	.02
Order	.10
Exh	-.09
Aut	.07
Aff	-.23
Int	.40
Suc	.25
Dom	-.39
Aba	.04
Nur	.33
Chg	-.55*
End	-.09
Het	.09
Agg	.14

.532 needed for significance at .05 level with df = 12*
.661 needed for significance at .01 level with df = 12**

descriptive measures and the performance examination averages. With the exception of the significant negative correlation (-.69) between the personality variable of Nurturance and the performance examination average, Hypothesis 8 was not rejected. Table 14 shows these data.

Hypothesis 9

There is no significant relationship between the two outcome measures, namely, the cognitive examination scores and the performance examination averages of the Control group.

A low negative correlation (-.19) was found between the cognitive examination scores and the performance examination averages. This can be contrasted with the highly significant .86 correlation between the outcome measures of the Experimental group. Little relationship was found between the cognitive examination scores and the performance examination averages of the Control group. Hypothesis 9 was not rejected.

Correlations Between Descriptive Measures and Outcome Measures and Between Outcome Measures of the Combined Group

The discussion in this section will center around correlations (1) between the 18 descriptive measures and the 2 outcome measures, and (2) between the outcome measures, of the combined Experimental and Control groups. The scores of all 25 subjects are included in the analysis. The descriptive measures of age, sex, grade point average for the last 60 undergraduate hours, 15 personality variable scores of the Edwards Personal Preference Schedule, and the outcome descriptive measures and the performance examination averages. With the exception of the significant negative correlation (-.29) between the personality variable of Nurturance and the performance examination average, Hypothesis 8 was not rejected. Table 14 shows these data.

Hypothesis 9

There is no significant relationship between the two outcome measures, namely, the cognitive examination scores and the performance examination averages of the Control group.

A low negative correlation (-.18) was found between the cognitive examination scores and the performance examination averages. This can be contrasted with the highly significant .86 correlation between the outcome measures of the Experimental group. Little relationship was found between the cognitive examination scores and the performance examination averages of the Control group. Hypothesis 9 was not rejected.

Correlations Between Descriptive Measures and Outcome Measures and Between Outcome Measures of the Combined Group

The discussion in this section will center around correlations (1) between the 18 descriptive measures and the 2 outcome measures, and (2) between the outcome measures, of the combined Experimental and Control groups. The scores of all 85 subjects are included in the analysis. The descriptive measures of age, sex, grade point average for the last 60 undergraduate hours, 15 personality variable scores of the Edwards Personal Preference Schedule, and the outcome

TABLE 14

CORRELATIONS BETWEEN DESCRIPTIVE MEASURES AND PERFORMANCE EXAMINATION AVERAGES OF THE CONTROL GROUP

VARIABLE	CORRELATIONS WITH PERF EXAM AVERAGES
Age	.21
Sex	.21
GPA	.12
Ach	.06
Def	-.21
Order	.52
Exh	.21
Aut	.51
Aff	-.45
Int	-.28
Suc	-.06
Dom	.51
Aba	-.38
Nur	-.69**
Chg	.44
End	-.10
Het	.35
Agg	.18

.532 needed for significance at .05 level with df = 12*
.661 needed for significance at .01 level with df = 12**

measures of cognitive examination scores and performance examination averages will be considered. The descriptive measures of number of WISC tests administered, number of protocols handed in for evaluation, number of WISC observations, and number of WISC tests administered for demonstration, are only applicable to the Experimental group, and are excluded from this discussion of combined group correlations.

The correlations between descriptive measures and the performance examination item scores are not included in this section. Out of 216 correlations, 11 were found to be significant, which is approximately the number expected by chance. Five of the significant positive correlations were between age and one of the item scores. These performance examination items were 2, 3, 6, 9, and 10. These data can be found in Appendix F.

Intercorrelations between the descriptive measures are found in Appendix F. The reader is also referred to Appendix F for correlations between the cognitive examination scores and the performance examination item scores, and for intercorrelations between the performance examination item scores.

Hypothesis 10

There are no significant relationships between age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, and the cognitive examination scores of the combined group.

measures of [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]

will be [illegible]. The [illegible] [illegible] of [illegible] [illegible] [illegible] [illegible]

number of [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] and

scores of WISC [illegible]

Expe[illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]

relations.

The correlations between [illegible] [illegible] [illegible] [illegible] [illegible] examination [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] correlations, 11 were found to be [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]cance. Five of the significant [illegible] [illegible] [illegible] [illegible] [illegible] of the [illegible] scores. [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] 11. These data can be found in Appendix [illegible]

Intercorrelations between the [illegible] [illegible] [illegible] [illegible] [illegible]. The reader is also referred to Appendix [illegible] for correlations between [illegible] of the [illegible] scores and the performance [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]

Hypothesis 10

There are no significant relationships [illegible] [illegible] [illegible] [illegible] grade point averages, [illegible] [illegible] [illegible] [illegible] [illegible] Preference Schedule, and the [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] group.

No significant relationships were found between any of the eighteen descriptive measures and the cognitive examination scores. For the Experimental group alone (Hypothesis 4), no significant relationships were found. Two were found to be significant for the Control group, age (-.58) and Change (-.55) (Hypothesis 7). No one individual trait correlated with the cognitive examination scores for the combined group, and Hypothesis 10 was not rejected. These data can be found in Table 15.

Hypothesis 11

There are no significant relationships between age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, and the performance examination averages of the combined group.

No significant relationships were found between any of the descriptive measures and the performance examination averages of the combined group. No descriptive measures correlated higher than .37 with the outcome measure of performance examination averages. No significant correlations were found between the same variables in the Experimental group (Hypothesis 5). The similar Control group data (Hypothesis 8) yielded one significant negative correlation (-.69) between the personality variable of Nurturance and the performance examination averages. No individual trait of the combined group correlated with the performance examination averages, and Hypothesis 11 was not rejected. These data can be found in Table 16.

TABLE 15

CORRELATIONS BETWEEN DESCRIPTIVE MEASURES AND COGNITIVE EXAMINATION SCORES OF THE COMBINED GROUP

VARIABLE	CORRELATION WITH COGNITIVE EXAM SCORES
Age	.16
Sex	-.16
GPA	.17
Ach	.14
Def	.02
Order	.17
Exh	-.01
Aut	-.20
Aff	-.16
Int	.31
Suc	.16
Dom	-.23
Aba	.14
Nur	.22
Chg	-.15
End	-.14
Het	-.11
Agg	-.13

.396 needed for significance at .05 level with df = 23*
.505 needed for significance at .01 level with df = 23**

TABLE 16

CORRELATIONS BETWEEN DESCRIPTIVE MEASURES AND PERFORMANCE EXAMINATION AVERAGES OF THE COMBINED GROUP

VARIABLE	CORRELATIONS WITH PERF EXAM AVERAGES
Age	.37
Sex	-.14
GPA	.02
Ach	.09
Def	.05
Order	.29
Exh	-.29
Aut	-.03
Aff	-.22
Int	-.02
Suc	.03
Dom	.00
Aba	.02
Nur	-.17
Chg	.37
End	-.06
Het	.02
Agg	-.08

.396 needed for significance at .05 level with df = 23*
.505 needed for significance at .01 level with df = 23**

Hypothesis 12

There is no significant relationship between the two outcome measures, namely, the cognitive examination scores and the performance examination averages of the combined group.

A significant (.59) correlation was found to exist between the cognitive examination scores and the performance examination averages of the combined group. This exceeded the necessary .505 needed for significance at the .01 level. This same correlation was significant at the .01 level for the Experimental group (Hypothesis 6), but it was not significant for the Control group (Hypothesis 9). A high degree of predictability was demonstrated between the cognitive examination scores and the performance examination averages of the combined group. Hypothesis 12 was rejected.

Summary

There were no significant differences between the two groups on the outcome measures of cognitive examination scores and performance examination averages. A significant difference was found between the groups on performance examination Item 11, Accuracy of scoring procedures. The difference was in favor of the Control group. No differences were found on the other eleven items.

The experimental group data indicated no significant correlations between the initial variables of age, sex, undergraduate grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, and the outcome measures of cognitive examination scores and performance examination

averages. Significant correlations were found between the number of WISC tests administered, the number of protocols handed in for evaluation, and the cognitive examination scores. Significant correlations were found between the number of WISC tests administered and the performance examination averages. A significant correlation was found between the cognitive examination scores and the performance examination averages.

The Control group data indicated significant negative correlations between age and the cognitive examination scores, and the personality variable of Change and the cognitive examination scores. The correlations between sex, grade point average, the 14 remaining personality variables, and the cognitive examination scores were not significant. A significant negative correlation was found between the personality variable of Nurturance and the performance examination averages. The correlations between age, sex, grade point average, the remaining 14 personality variables, and the performance examination averages were not significant. The correlation between the cognitive examination scores and the performance examination averages was not significant (-.19).

The combined group data indicated no significant correlations between age, sex, grade point average, fifteen personality variable scores of the Edwards Personal Preference Schedule, and the outcome measures of cognitive examination scores and performance examination averages. A significant relationship was found between the cognitive examination scores and the performance examination averages.

CHAPTER V

SUMMARY, CONCLUSIONS, RECOMMENDATIONS, AND IMPLICATIONS

Summary

The study was conducted in an effort to determine if there were any differences in terms of student learning between two methods of instruction in teaching the Wechsler Intelligence Scale for Children (WISC). The scores on two outcome measures--a cognitive examination and a performance examination--were compared between students in an Experimental group who learned by means of independent study and those in a Control group who learned to understand, administer, and score the WISC by means of a traditional method.

Initial differences between the groups were ascertained on 18 variables --age, sex, undergraduate grade point average, and 15 personality variables of the Edwards Personal Preference Schedule. Relationships were determined, (1) between the 18 initial measures and the 2 outcome measures, and (2) between the outcome measures, of the Experimental, Control, and combined groups. The number of WISC tests administered and the number of protocols handed in for evaluation were correlated with performance on the two outcome measures for the Experimental group.

Statistical data indicated initial differences between the two groups on three personality variables. The Experimental group scored higher on the variables

of Deference and Order, and the Control group scored higher on the variable of Autonomy. Taking initial differences into consideration, an analysis of covariance procedure indicated no significant differences between the groups on cognitive examination scores and performance examination averages. Nonsignificant F ratios were obtained on 11 of the 12 performance examination item scores. On Item 11 (Accuracy of scoring procedures), the Control group performed significantly higher than the Experimental group.

None of the initial variables correlated with the cognitive examination scores of the Experimental and combined groups. Both age and the EPPS variable of Change correlated negatively with the cognitive examination scores of the Control group.

No significant correlations were found between the 18 initial measures and the performance examination averages of the Experimental and combined groups. The Control group data indicated a significant negative correlation between the EPPS variable of Nurturance and the performance examination averages.

Significantly high relationships were found between the two outcome measures of the Experimental and combined groups. The Control group data indicated little relationship between the cognitive examination scores and the performance examination averages.

The number of WISC tests administered by the Experimental group correlated significantly with both cognitive examination scores and the performance examination averages. A significant correlation was obtained between the number of Deference and Order, and the Control group scored higher on the variable of Autonomy. Taking initial differences into consideration, an analysis of covariance procedure indicated no significant differences between the groups on cognitive examination scores and performance examination averages. Nonsignificant F ratios were obtained on 11 of the 12 performance examination item scores. On Item 11 (Accuracy of scoring procedures), the Control group performed significantly higher than the Experimental group.

None of the initial variables correlated with the cognitive examination scores of the Experimental and combined groups. Both age and the EPPS variable of Change correlated negatively with the cognitive examination scores of the Control group.

No significant correlations were found between the 18 initial variables and the performance examination averages of the Experimental and combined groups. The Control group data indicated a significant negative correlation between the EPPS variable of Nurturance and the performance examination averages.

Significantly high relationships were found between the two outcome measures of the Experimental and combined groups. The Control group data indicated little relationship between the cognitive examination scores and the performance examination averages.

The number of WISC tests administered by the Experimental group correlated significantly with both cognitive examination scores and the performance examination averages. A significant correlation was obtained between the number of protocols turned in for evaluation and the cognitive examination scores of the Experimental group.

Conclusions

Based upon the results of the experiment, several conclusions were reached:

1. The data essentially confirmed the findings of previous research in the area of comparative teaching methods--that is, that one method of instruction was generally as effective as another in affecting outcome on a final objective examination. In the present study, there was no significant difference between the means of the groups on the paper-and-pencil examination. No research was available, however, to lead the researcher to predict that such results would occur on a performance examination which required the demonstration of a prescribed skill. Based on the findings of the current study, the above inference could be made. Additional research is needed, however, to determine if this was an isolated event, or if this conclusion could be generalized to other studies which compared two methods of instruction in teaching the demonstration of a prescribed skill. The present research suggests that this inference is possible.

2. A minimum number of practice WISC administrations are essential in a course in individual intelligence testing. In spite of the fact that no significant differences were detected between the means of the groups on the two outcome measures, a more precise examination of the data indicated some striking differences in terms of ranges and standard deviations. On both the cognitive

of protocols termed in the evaluation ...

Experimental group.

Based upon the results of the ...

reached:

1. The data essentially ...

the case of cooperative teaching ...

was generally as effective as ...

examination. In the present study, ...

the means of the groups on the ...

available, however, to test the ...

occur on a performance examination ...

scribed skill. Based on the findings ...

could be made. Additional research ...

was an isolated event, or if this ...

which compared two methods of ...

prescribed skill. The present research ...

2. A minimum number of ...

in a course in individual intelligence ...

significant differences were detected ...

outcome measures, a more precise ...

differences in terms of ranges and ...

examination and the performance examination, Experimental group members obtained both lower and higher scores than did the subjects in the Control group. Both the range and the standard deviation were greater. A larger percentage (36%) of the Experimental group (21% in the Control group) scored below the mid-point and were asked to repeat the performance examination. The requirements of the Control group appeared to result in a lower percentage of poor performances on the performance examination, and fewer low scores on the cognitive examination in the Control group.

As indicated in their anonymous comments (Appendix G), the level of discontent was fairly high in the Control group. The class was required for students in the Block Program and not for other Masters degree students in Guidance and some resented this. Several indicated a dislike of testing in general. In spite of these negative feelings, the Autonomous nature of the group, and the general dislike of requirements, no student in the group scored as low on either examination as did the lowest students in the Experimental group. The requirements of the course, at least, seemed to raise the lower levels of performance.

Further, significantly high correlations were obtained between the number of WISC tests administered by Experimental group members and performance on both outcome measures. Those students who administered more WISC tests generally obtained higher cognitive and performance examination scores, and vice-versa. The four students in the Experimental group who were asked to repeat the performance examination had administered the fewest number of practice WISC examination and the performance examination, Experimental group members obtained both lower and higher scores than did the subjects in the Control group. Both the range and the standard deviation were greater. A larger percentage (36%) of the Experimental group (21% in the Control group) scored below the mid-point and were asked to repeat the performance examination. The requirements of the Control group appeared to result in a lower percentage of poor performances on the performance examination, and fewer low scores on the cognitive examination in the Control group.

As indicated in their anonymous comments (Appendix G), the level of discontent was fairly high in the Control group. The class was required for students in the Block Program and not for other Masters degree students in Guidance and some resented this. Several indicated a dislike of testing in general. In spite of these negative feelings, the Autocratic nature of the group, and the general dislike of requirements, no student in the group scored as low on either examination as did the lowest students in the Experimental group. The requirements of the course, at least, seemed to raise the lower levels of performance.

Further, significantly high correlations were obtained between the number of WISC tests administered by Experimental group members and performance on both outcome measures. Those students who administered more WISC tests generally obtained higher cognitive and performance examination scores, and vice-versa. The four students in the Experimental group who were asked to repeat the performance examination had administered the fewest number of practice WISC tests. Thus, if for no other reason than to insure a minimum level of competence, a certain number of WISC administrations ought to be required in a course in individual intelligence testing. This study, however, made no attempt to establish that minimum.

3. A minimum number of scored protocols should be submitted for evaluation in a course in individual intelligence testing. A significant correlation (.73) was found between the number of protocols handed in for evaluation and the cognitive examination scores of the Experimental group. A correlation (.58) which approached, but did not reach, significance was found between the number of scored protocols handed in for evaluation and the performance examination averages. Those students who asked for more feedback in this regard, generally obtained higher scores on the examinations. The mean of the Control group on performance examination Item 11 (Accuracy of scoring procedures) was significantly higher than the mean of the Experimental group. These students were required to submit 15 protocols, whereas the Experimental group students handed in a mean of 4.73 protocols for evaluation. Three students handed in none, and 2 handed in 14. This difference between the groups appeared to increase the accuracy of scoring procedures in favor of the Control group. It was concluded that students should be required to submit a number of scored WISC protocols for evaluation, but this study made no attempt to discern that minimum number.

4. Success on outcome measures is less dependent upon method of instruction than upon the personality of the individual. Several of the Experi-

tests. Thus, if for no other reason than to insure a minimum level of competence, a certain number of WISC administrations ought to be required in a course in individual intelligence testing. This study, however, made no attempt to establish that minimum.

3. A minimum number of scored protocols should be submitted for evaluation in a course in individual intelligence testing. A significant correlation (.73) was found between the number of protocols handed in for evaluation and the cognitive examination scores of the Experimental group. A correlation (.38) which approached, but did not reach, significance was found between the number of scored protocols handed in for evaluation and the performance examination averages. Those students who asked for more feedback in this regard, generally attained higher scores on the examinations. The mean of the Control group on performance examination item 11 (Accuracy of scoring procedures) was significantly higher than the mean of the Experimental group. These students were required to submit 15 protocols, whereas the Experimental group students handed in a mean of 4.73 protocols for evaluation. Three students handed in none, and 2 handed in 14. This difference between the groups appeared to increase the accuracy of scoring procedures in favor of the Control group. It was concluded that students should be required to submit a number of scored WISC protocols for evaluation, but this study made no attempt to discern that minimum number.

4. Success on outcome measures is less dependent upon method of instruction than upon the personality of the individual. Several of the test-

mental group members indicated in their anonymous comments (Appendix G) that they missed the group contact they had found in their other courses. Some said they found it frustrating to listen to the tapes alone without having someone there with whom to discuss their reactions. Group feedback seemed to be important to several members, but unless they "set it up" on their own, this contact was not a part of the Experimental group experience. This lack of group contact and, thus, the testing of unprocessed material on the cognitive examination, might have contributed to lower scores on the part of some Experimental group members. On the other hand, some Experimental group members indicated relief at not having to come to the University on Thursday nights. They indicated no frustration at not being able to discuss the material with others. Individual differences, as opposed to method of instruction, appeared to determine success on outcome measures.

None of the initial measures selected for consideration in this study correlated significantly with either of the two outcome measures of the Experimental and combined groups. No personality variable was significantly related to outcomes in these two groups. (The Control group data will be discussed later.) This researcher is of the opinion that individual motivation and expectations (McKeachie, 1958), or level of aspiration (Burton, 1967), are the keys to understanding individual level of attainment. This study, however, made no attempt to ascertain initial measures on these variables. What is known, in light of the results of the present study, is which variables are <u>not</u> related to achievement on outcome measures. That is, there were no significant correlations between mental group members indicated in their anonymous comments (Appendix G) that they missed the group contact they had found in their other courses. Some said they found it frustrating to listen to the tapes alone without having someone there with whom to discuss their practices. Group feedback seemed to be important to several members, but unless they "set it up" on their own, this contact was not a part of the Experimental group experience. This lack of group contact and, thus, the testing of unprocessed material on the cognitive examination, might have contributed to lower scores on the part of some Experimental group members. On the other hand, some Experimental group members indicated relief at not having to come to the University on Thursday nights. They indicated no frustration at not being able to discuss the material with others. Individual differences, as opposed to method of instruction, appeared to determine success on outcome measures.

None of the initial measures selected for consideration in this study correlated significantly with either of the two outcome measures of the Experimental and combined groups. No personality variable was significantly related to outcomes in these two groups. (The Control group data will be discussed later.) This researcher is of the opinion that individual motivation and expectations (McKeachie, 1958), or level of aspiration (Barnes, 1967), are the keys to understanding individual level of attainment. This study, however, made no attempt to ascertain initial measures on these variables. What is known, in light of the results of the present study, is which variables are not related to achievement on outcome measures. That is, there were no significant correlations between the 18 initial personality variables and the 2 outcome measures in the Experimental and combined groups.

It appeared that some Experimental group members were motivated to perform successfully and other weren't. The same people, generally, who scored high on the cognitive examination, scored high on the performance examination. These were generally the same people who administered the highest number of WISC tests, handed in the largest number of protocols for evaluation, and contacted the instructor informally. These people didn't seem to "need" minimum requirements because their internal motivation prompted them to develop competency on their own. It is for those students who didn't, apparently, feel this motivation that the minimum requirements are necessary. This point is entirely speculative, however, since level of motivation was not determined initially in this study.

The Experimental and combined group data yielded no significant correlations between the 18 initial measures and the 2 outcome measures, whereas 3 significant correlations were found between the initial measures and the outcome measures in the Control group data. Younger subjects and subjects with lower Change scores, generally, obtained higher cognitive examination scores. Subjects with lower Nurturance scores, generally, obtained higher performance examination averages. In the Control group, these subjects who scored high on the cognitive examination were, generally, not the same people as those who obtained high ratings on the performance examination. This can be compared to significant (.01) correlations of .86 for the Experimental group and .59 for the combined group between the cognitive examination scores and the perform-the 18 initial personality variables and the 2 outcome measures in the Experimental and combined groups.

It appeared that some Experimental group members were motivated to perform successfully no matter what. The same people, generally, who scored high on the cognitive examination, scored high on the performance examination. These were generally the same people who administered the highest number of WISC tests, handed in the largest number of protocols for evaluation, and contacted the instructor informally. These people didn't seem to "need" minimum requirements because their internal motivation prompted them to develop competency on their own. It is for those students who didn't, apparently, feel this motivation that the minimum requirements are necessary. This point is entirely speculative, however, since level of motivation was not determined initially in this study.

The Experimental and combined group data yielded no significant correlations between the 18 initial measures and the 2 outcome measures, whereas 3 significant correlations were found between the initial measures and the outcome measures in the Control group data. Younger subjects and subjects with lower Change scores, generally, obtained higher cognitive examination scores. Subjects with lower Nurturance scores, generally, obtained higher performance examination averages. In the Control group, those subjects who scored high on the cognitive examination were, generally, not the same people as those who obtained high ratings on the performance examination. This can be compared to significant (.01) correlations of .56 for the Experimental group and .39 for the combined group between the cognitive examination scores and the performance examination averages. An explanation for the Control group data might be in the initial differences between the groups. The Control group was more Autonomous, less Orderly, and less Deferent. An inference of less consistency of behavior among the Control group subjects might be made. As a group, they were less predictable and more independent. The Control group data were less predictable and less consistent than were either the Experimental or combined group data.

Group means, however, do not speak to individual differences, and possibly, those subjects with the lower Change scores (i.e., less need for change) were somewhat more orderly and organized, and thus, scored higher on the cognitive examination. Similarly, those subjects with fewer Nurturance needs (i.e., possibly a lower interest in people and more interest in things) were able to perform significantly higher on the performance examination. These points are entirely speculative, however, and are only intended as plausible explanations for the obtained results. The reader is referred to Appendix F for intercorrelations among the initial variables.

Recommendations

Based upon the results and conclusions reached in this study, the following recommendations are made:

1. It is recommended that the study be repeated and that the subjects be randomly assigned to treatment conditions. This procedure would eliminate many of the limitations of the present study. The initial differences between

ance examination averages. An explanation for the [illegible] in the initial differences between the groups. The Control group was [illegible] Autonomous, less Orderly, and less Deferent. An [illegible] of behavior among the Control group subjects might be [illegible] were less predictable and more independent. The Control group subjects [illegible] predictable and less consistent than were either the [illegible] group data.

Group means, however, do not speak to individual differences, and possibly, those subjects with the lower Change scores [illegible] were somewhat more orderly and organized, and thus, [illegible] nitive examination. Similarly, those subjects with lower [illegible] (i.e., possibly a lower interest in people and more [illegible] to perform significantly higher on the performance examination. [illegible] are entirely speculative, however, and are only intended as plausible explanations for the obtained results. The reader is referred to Appendix [illegible] for intercorrelations among the initial variables.

Recommendations

Based upon the results and conclusions reported in this study, the following recommendations are made:

1. It is recommended that the study be repeated and that the subjects be randomly assigned to treatment conditions. This procedure would eliminate many of the limitations of the present study. The initial differences [illegible]

the groups in terms of personality variables, peer-group influence, and day-time vs. night-time students would be eliminated, and the data would more precisely reflect treatment effect.

2. Further, it is proposed that the study include several different treatment levels. It was concluded that a minimum number of WISC administrations and scored protocols were essential to insure a minimum level of performance, but this minimum level was not established. The recommendation is that the independent study group be retained, but that the subjects also be randomly assigned to groups that are required to administer and score, possibly, 5, 10, and 15 WISC tests, respectively.

3. It is recommended that the initial measures be expanded or revised to include an estimate of motivation. Possibly, a self-report inventory, such as Shostrom's Personal Orientation Inventory (Educational and Industrial Testing Service, 1966) which has an Inner- Other-directed scale, or Cattell and Horn's Motivational Analysis Test (Institute for Personality and Ability Testing, 1964), could be used. Level of motivation appeared to be crucial to success on outcome measures, and, by some procedure, it should be initially considered by future researchers.

Implications

The results of the study demonstrated that there were no significant differences between students taught by an independent study method and a tradithe groups in terms of personality variables, peer-group influences, and day-time vs. night-time students would be eliminated, and the data would more precisely reflect treatment effects.

2. Further, it is proposed that the study include several different treatment levels. It was considered that a minimum number of WISC administrations and scored protocols were essential to insure a minimum level of performance, but this minimum level was not established. The recommendation is that the independent study group be retained, but that the subjects also be randomly assigned to groups that are required to administer and score, possibly, 5, 10, and 15 WISC tests, respectively.

3. It is recommended that the initial measures be expanded or revised to include an estimate of motivation. Possibly, a self-report inventory, such as Shostrom's Personal Orientation Inventory (Educational and Industrial Testing Service, 1966) which has an inner- Other-directed scale, or Cattell and Horn's Motivational Analysis Test (Institute for Personality and Ability Testing, 1964), could be used. Level of motivation appeared to be crucial to success on outcome measures, and, by some subjective, it should be initially considered by future researchers.

Implications

The results of the study demonstrated that there were no significant differences between students taught by an independent study method and a traditional method with regard to either understanding of the WISC or the administration and scoring of the WISC. Based on the results of this experiment, instructors of the WISC could justifiably record their lectures on audiotapes and the WISC demonstrations on videotapes, assign a number of texts and handouts, give a full-day workshop, and expect student performance on two outcome measures to be not significantly different than if they had attended class regularly and presented the information "in person."

The finding that minimal contact hours are needed for the effective transmission of information increases instructor flexibility and allows for the re-deployment of faculty time. Because they need not attend class on a regular basis to present information to students, options that were not before available, are open for their consideration.

Many members of the Experimental group said they missed the opportunity to share their ideas and learning with both peers and the instructor. Some mentioned the isolation they felt listening to the tapes at home, alone. These student re-actions need not be a drawback to the adoption of independent study courses, however. With reallocation of time possible, students could acquire a large amount of information independently, and instructors could meet with them for other pur-poses. Discussions, seminars, conferences, and feedback sessions are options. The possibilities for redeployment of time are limited only by the needs and imagination of the individual instructor.

There was much variability among the Experimental group members with regard to the number of WISC tests administered and the number of scored protocols tional method with regard to either understanding of the WISC or the administration and scoring of the WISC. Based on the results of this experiment, instructors of the WISC could justifiably record their lectures on audiotapes and the WISC demonstration on videotape, assign a number of texts and handouts, give a full-day workshop, and expect student performance on two outcome measures to be not significantly different than if they had attended class regularly and presented the information "in person."

The finding that minimal contact hours are needed for the effective transmission of information increases instructor flexibility and allows for the re-deployment of faculty time. Because they need not attend class on a regular basis to present information to students, options that were not before available, are open for their consideration.

Many members of the Experimental group said they missed the opportunity to share their ideas and learning with both peers and the instructor. Some mentioned the isolation they felt listening to the tapes at home, alone. These student reactions need not be a drawback to the adoption of independent study courses, however. With reallocation of time possible, students could acquire a large amount of information independently, and instructors could meet with them for other purposes. Discussions, seminars, conferences, and feedback sessions are options. The possibilities for redeployment of time are limited only by the needs and imagination of the individual instructor.

There was much variability among the Experimental group members with regard to the number of WISC tests administered and the number of second protocols submitted for evaluation. Those students who administered the fewest number of practice tests and submitted the fewest number of protocols did the poorest on both outcome measures. The conclusion drawn from these findings is that a minimun number of WISC administrations and scored protocols should be required in order to provide extrinsic motivation for those students who apparently aren't intrinsically motivated to reach a minimum level of competency in both cognitive understanding and administration of the WISC. One might ask whether the development of intrinsic motivation in students isn't about as important as how well a student learns to accomplish a certain task. Obviously, success in life demands a great deal of self-reliance, and an educational system which only provides opportunities for extrinsic sources of motivation does little to prepare people to function successfully in life.

Granted, in an area such as individual intelligence testing, which personally affects the lives of so many human beings, competent examiners are essential. Presently, it appears necessary to require a minimum number of practice administrations and protocol checks in order to insure a minimum level competency. However, independent study need not imply a lack of minimum requirements. The instructor employing independent study should expect a certain amount of resistance and shirking of responsibility, due mainly to lack of past exposure to the method on the part of the students. However, with achievement standards or minimum requirements, a basic level of competency can be reached, and yet, students can be provided with opportunities for self-direction and choice.

submitted for evaluation. Those students who adhered to the [illegible] practice tests and submitted the fewest number of protocols [illegible] both outcome measures. The conclusion drawn from this [illegible] mum number of WISC administrations and scored protocols should be required in order to provide extrinsic motivation for those students who are not sufficiently intrinsically motivated to reach a minimum level of competency in the [illegible] understanding and administration of the WISC. One might ask whether the development of intrinsic motivation in students isn't about as important as how well a student learns to accomplish a certain task. Admittedly, [illegible] demands a great deal of self-reliance, and an educational system which [illegible] vides opportunities for extrinsic sources of motivation [illegible] to function successfully in life.

Granted, in an area such as individual intelligence testing, which [illegible] usually affects the lives of so many human beings, competent examiners are essential. Presently, it appears necessary to require a minimum number of WISC administrations and protocol checks in order to insure a minimum level of [illegible] However, independent study need not imply a lack of minimum requirements. The instructor employing independent study should expect a considerable amount of [illegible] and shirking of responsibility, due mainly to lack of past experience [illegible] on the part of the students. However, with achievement of minimum [illegible] requirements, a basic level of competency can be reached, and yet students can be provided with opportunities for self-direction and choice.

Educators are further challenged to allow opportunities for students to reach the maximum level of their capabilities. In the present study, the Experimental group obtained not only lower scores, but also higher scores, on both the cognitive examination and the performance examination. Some Experimental group members administered more WISC tests, handed in more scored protocols for evaluation, observed more other students administering the WISC, and administered more WISC tests for demonstration than the members of the Control group. Perhaps "requirements" limit the upper levels of performance and do not allow students to reach as far as they are capable, or would like to reach. In every learning situation, students should have the opportunity to tap their inner resources and move beyond any minimum requirements that might have been set. The current study demonstrated that students can exceed the requirements if allowed the opportunity to do so.

Certain inferences can also be made concerning differences between part- and full-time students on the basis of this study. The evening students scored higher on the Edwards Personal Preference Schedule variables of Order and Deference, whereas the full-time students scored higher on Autonomy. It isn't known whether these personality differences occur at institutions other than the University of New Mexico, or in other UNM classes. More of the evening students (Experimental group) worked in the daytime, they were slightly older than the Control group students (28.45 vs. 27.50), and were more firmly established in a profession. They might be looked upon as more integrated into the main stream of American society, which often places a premium upon relative

conformity and adherence to regulations. In this context, the higher Order and Deference scores become understandable. A higher percentage of them (36%, compared with 21% for the Control group) repeated the performance examination. They were not quite as able to set their own guidelines when they were not told by someone in authority how to proceed. Their needs for Order and Deference were not met by the Experimental group design, but some did well in spite of this.

The Control group members were full-time students who were participating in a 15 month program of full-time study. Some had held jobs prior to returning to school, but as far as this researcher could determine, none had responsibilities to any outside agencies or institutions at the time of the study. They had, in a sense, severed ties with outside organizational influences in order to further their education. These facts might serve to explain why they, as a group, scored higher on the EPPS variable of Autonomy. Their anonymous comments indicated that they generally disliked requirements and being told what to do.

Teachers in higher education should consider providing additional opportunities for daytime students to fulfill their Autonomy needs through self-directed learning and independent study. The current research did not allow them this opportunity, and as a result, many students in the Control group seemed frustrated. On the other hand, evening students (with high Order and Deference needs) should have opportunities to explore new frontiers in their learning and to develop greater reliance upon their inner resources.

conformity and adherence to [illegible]

Deference scores became [illegible]

compared with 27% for the [illegible]

They were not quite as able [illegible]

by someone in authority how to [illegible]

were not met by the Department [illegible]

this.

The Control group members [illegible]

in a 15 month program of full-time [illegible]

school, but as far as this researcher [illegible]

any outside agencies or institutions [illegible]

severed ties with outside organizations [illegible]

tion. These facts might serve to [illegible]

the EPPS variable of Autonomy. [illegible]

generally disliked routineness [illegible]

Teachers in higher education [illegible]

portunities for daytime students to [illegible]

learning and independent study. [illegible]

opportunity, and as a result, many [illegible]

On the other hand, evening students [illegible]

have opportunities to explore [illegible]

reliance upon their inner resources.

Instructors should possess a strong sense of purpose for it is easy to succumb to students' demands for limits, requirements, etc., in order to reduce their frustration and minimize ambiguity. While a certain amount of structure (or requirements) is called for in order to assure minimum competency standards, the goal should be one of greater reliance on student, rather than instructor, choices and the encouragement of intrinsic motivation on the part of each student.

Possibly the most significant implication of this study has to do with the humanness of students. Technology might be an efficient means of saving valuable educator-time, and requirements may reduce ambiguity on the part of both the instructor and the students and insure a minimum level of competency, however, the person of the student deserves attention and consideration. After all, he/she is what "it's all about." Students dislike isolation and feel a need to share ideas with others, they can tap their intrinisic sources of motivation if given the opportunity, and they dislike feeling stifled by requirements that don't seem to meet their needs. If challenged, they can exceed the "requirements."

APPENDICES

APPENDIX A

Guidance 512--Fall, 1973

GROUP I

Your Guidance 512 class has been selected to take part in an independent study experiment. The study itself will "run" the first ten weeks of the semester or to the end of the week of October 29 to November 2, 1973. After that time, or beginning November 8, 1973, you will meet your class regularly, as originally scheduled. The experiment itself involves the use of the Wechsler Intelligence Scale for Children (WISC) and the concentration of the class until November 2, 1973, will be upon this instrument.

The other section of Guidance 512 will serve as the "control" group in the study. The mean undergraduate grade point averages of the two classes, and the means of scales on the Edwards Personal Preference Schedule will be compared to determine initial differences between the two classes, and for later statistical analyses. Your GPA's and scores will be used for no other purposes.

During the first nine weeks of the semester (August 27 to October 26, 1973) no regular class meetings will be held for your group. Other than a workshop which will be described later, all learning will be dependent upon you and your own initiative. Your class meetings will resume November 1, 1973.

Cognitive Understanding of the WISC. On November 1, 1973, at the regularly scheduled class time of 6:30 p.m., you will meet to take a paper-and-pencil test over the cognitive aspects of the WISC. All information necessary for this

test will be available from the following sources:

1. Handouts issued at the first class meeting.
2. Cassette tapes of lectures available from the Guidance secretary. These tapes and a cassette recorder may be checked out from her.
3. Textbooks and other readings assigned by Dr. Heisey at the first class meeting.
4. Videotaped administration of the WISC. The tape will be shown three times during the nine week period. They will be shown at Manzanita Center, Thursday, September 13 at 12, Thursday, September 13 at 7 p.m., and Wednesday, October 3, at 3 p.m.
5. A workshop to be held Saturday, September 8, 1973, at Manzanita Center from 8:30 a.m. to 4:30 p.m. You will bring: 1) the WISC Manual which you will purchase at the UNM bookstore, 2) 5 WISC record forms and 5 Mazes which you will purchase for 17¢ each from the Manzanita Center Secretary, 3) several pencils with erasers, and 4) a stopwatch. WISC kits will be provided. Please read the WISC Manual before the workshop.

Administration and Scoring of the WISC. During the week of October 29 to November 2, 1973, you will be expected to administer a WISC for demonstration at Manzanita Center. You will bring your own subject for this and all other demonstration tests. You will be evaluated on the quality of your test administration and scoring. The scored protocol will be handed in as soon as possible after the demonstration (preferably the same day). Proficiency may be gained in test will be available from the following sources.

1. Handouts passed at the first class meeting.
2. Cassette tapes of lectures available from the Guidance secretary. These tapes and a cassette recorder may be checked out from her.
3. Textbooks and other readings assigned by Dr. Helsey at the first class meeting.
4. Videotaped administration of the WISC. The tape will be shown three times during the nine week period. They will be shown at Manzanita Center, Thursday, September 12 or 12, Thursday, September 13 at 7 p.m., and Wednesday, October 3, at 3 p.m.
5. A workshop to be held Saturday, September 8, 1973, at Manzanita Center from 8:30 a.m. to 4:30 p.m. You will bring: 1) the WISC Manual which you will purchase at the UNM bookstore, 2) 5 WISC record forms and 5 Mazes which you will purchase for 17¢ each from the Manzanita Center Secretary, 3) several pencils with erasers, and 4) a stopwatch. WISC kits will be provided. Please read the WISC Manual before the workshop.

Administration and Scoring of the WISC. During the week of October 29 to November 2, 1973, you will be expected to administer a WISC for demonstration at Manzanita Center. You will bring your own subject for this and all other demonstration tests. You will be evaluated on the quality of your test administration and scoring. The scored protocol will be handed in as soon as possible after the demonstration (preferably the same day). Proficiency may be gained by any way you choose but the following methods are suggested:

1. Attendance at the workshop on September 8, 1973.
2. Viewing of the videotaped administration of the WISC. As mentioned previously, this tape will be shown three times during the nine week period.
3. Practice in administration and scoring. Although no set number of test administrations are "required" results of previous studies have indicated that people learn by doing. You will only be evaluated on your final performance (week of October 29 to November 2) but practice administrations are highly encouraged.
4. Observations of others administering the WISC. No observations are "required" but it will be to your advantage to watch other students administer the test as often as possible. Check with the Manzanita Center appointment book to see when WISC demonstrations are scheduled.

You will make the arrangements with the Manzanita Center secretary for your final demonstration (week of October 29 to November 2). You are encouraged to do this in advance of this week since all persons enrolled in the two sections of Guidance 512 will be demonstrating during the same week and the time periods available (due to room space and graduate assistant schedules) will be limited.

any way you choose but the following methods are suggested:

1. Attendance at the workshop on September 5, 1973.
2. Viewing of the videotaped administration of the WISC. As mentioned previously, this tape will be shown three times during the nine week period.
3. Practice in administration and scoring. Although no set number of test administrations are "required" results of previous studies have indicated that people learn by doing. You will only be evaluated on your final performance (week of October 29 to November 2) but practice administrations are highly encouraged.
4. Observations of others administering the WISC. No observations are "required" but it will be to your advantage to watch other students administer the test as often as possible. Check with the Mennonite Center appointment book to see when WISC demonstrations are scheduled.

You will make the arrangements with the Mennonite Center secretary for your final demonstration (week of October 29 to November 2). You are encouraged to do this in advance of this week since all persons enrolled in the two sections of Guidance 512 will be demonstrating during the same week and the time periods available (due to room space and graduate assistant schedules) will be limited.

Miscellaneous Procedures.

1. WISC kits may be checked out at Manzanita Center for a 24 hour period. All test forms may be purchased at Manzanita and the Manuals will be purchased by you at the UNM bookstore.
2. Whenever you feel a need for feedback during the initial nine week period, you may schedule a demonstration WISC with the secretary at Manzanita. The graduate assistant will have left his time schedule with her so she will know when he is available to observe your session.
3. Any scored protocols about which you would like feedback should be placed in the "In" Differential Diagnosis box in the Manzanita Conference room. After they are checked, they can be picked up in the "Out" Differential Diagnosis box.
4. Dr. Heisey and the graduate assistant will be available for appointments, conferences, or feedback. Check with them regarding their available times.
5. During the first class meeting, you will be issued a Manzanita Center observation booth pass and a procedure sheet describing the use of the booth.
6. Manzanita Center has a few stopwatches which may be used only for tests administered in the Center. You will need your own stopwatch for tests administered outside of the Center and for the workshop on September 8.

Guidance 512--Fall, 1973
GROUP II

Your Guidance 512 class has been selected to act as the control group in an independent study experiment. The study itself will "run" the first ten weeks of the semester or to the end of the week of October 29 to November 2, 1973. The experiment itself involves the use of the Wechsler Intelligence Scale for Children (WISC) and the concentration of the class until November 2, 1973, will be upon this instrument. Your class will meet as regularly scheduled throughout the semester.

The other section of Guidance 512 will serve as the experimental group in the study. The mean undergraduate grade point average of the two classes and the means of scales on the Edwards Personal Preference Schedule will be compared to determine initial differences between the two classes, and for later statistical analyses. Your GPA's and scores will be used for no other purposes.

Cognitive Understanding of the WISC. On October 30, 1973 at the regularly scheduled class time you will take a paper-and-pencil test over the cognitive aspects of the WISC. All information necessary for this test will be available from the following sources:

1. Handouts issued in class.
2. Textbooks and other reading assigned by Dr. Heisey in class.
3. In-class lectures, discussions, and demonstrations.

Guidance 812—Fall, 1973
GROUP II

Your Guidance 812 class has been selected to act as the control group in an independent study experiment. The study itself will "run" the first ten weeks of the semester or to the end of the week of October 29 to November 2, 1973. The experiment itself involves the use of the Wechsler Intelligence Scale for Children (WISC) and the concentration of the class until November 2, 1973, will be upon this instrument. Your class will meet as regularly scheduled throughout the semester.

The other section of Guidance 812 will serve as the experimental group in the study. The mean undergraduate grade point average of the two classes and the means of scales on the Edwards Personal Preference Schedule will be compared to determine initial differences between the two classes, and for later statistical analyses. Your GPA's and scores will be used for no other purposes. Cognitive Understanding of the WISC. On October 30, 1973 at the regularly scheduled class time you will take a paper-and-pencil test over the cognitive aspects of the WISC. All information necessary for this test will be available from the following sources:

1. Handouts issued in class.

2. Textbooks and other reading assigned by Dr. Heisey in class.

3. In-class lectures, discussions, and demonstrations.

4. Videotaped administration of the WISC. The tape will be shown three times during the nine week period. They will be shown at Manzanita Center Thursday, September 13 at 12, Thursday, September 13 at 7 p.m., and Wednesday, October 3 at 3 p.m.

Administration and Scoring of the WISC. During the week of October 29 to November 2, 1973, you will be expected to administer a WISC for demonstration at Manzanita Center. You will bring your own subject for this and all other demonstration tests. You will be evaluated on the quality of your test administration and scoring. The scored protocol will be handed in as soon as possible after the demonstration (preferably the same day). Administration and scoring proficiency will be gained in the following ways:

1. Viewing a videotaped administration at the WISC. As mentioned previously, this tape will be shown three times during the nine week period.
2. Practice in administration and scoring.
   a. You will be expected to administer a total of 15 WISC tests during the period from August 27 to November 2, 1973. Place the scored protocols in the Differential Diagnosis "In" box in the Manzanita Center Conference Room. After they are checked, they will be placed in the "Out" Differential Diagnosis box and you may pick them up.
   b. You will be expected to administer one additional WISC test at Manzanita for demonstration prior to the week of

4. Videotaped administration of the WISC. The tape will be shown three times during the nine week period. They will be shown at Manzanita Center: Thursday, September 13 at 12, Thursday, September 13 at 7 p.m., and Wednesday, October 2 at 3 p.m.

Administration and Scoring of the WISC. During the week of October 29 to November 2, 1973, you will be expected to administer a WISC for demonstration at Manzanita Center. You will bring your own subject for this and all other demonstration tests. You will be evaluated on the quality of your test administration and scoring. The scored protocol will be handed in as soon as possible after the demonstration (preferably the same day). Administration and scoring proficiency will be gained in the following ways:

1. Viewing a videotaped administration of the WISC. As mentioned previously, this tape will be shown three times during the nine week period.

2. Practice in administration and scoring.

a. You will be expected to administer a total of 15 WISC tests during the period from August 27 to November 2, 1973. Place the scored protocols in the Differential Diagnosis "In" box in the Manzanita Center Conference Room. After they are checked, they will be placed in the "Out" Differential Diagnosis box and you may pick them up.

b. You will be expected to administer one additional WISC test at Manzanita for demonstration prior to the week of

October 29 to November 2. This should be scheduled prior to September 28. The purpose of this administration prior to the final one is three-fold: 1) to allow you to have feedback prior to your final administration, 2) to allow others to observe WISC administrations, and 3) to allow the graduate assistant to check the reliability of his rating instrument. The 2 tests administered for demonstration will "count" toward the 15 required.

3. Observations of WISC administrations. You will be expected to observe 5 WISC administrations. Sign the sheet in the observation booth when you observe. Check the appointment book at Manzanita Center to see when WISC administrations are scheduled.

You will make the arrangements with the Manzanita Center secretary for your final demonstration (week of October 29 to November 2). You are encouraged to do this in advance of this week since all persons enrolled in the two sections of Guidance 512 will be demonstrating during the same week and the time periods available (due to space and graduate assistant schedules) will be limited.

Miscellaneous Procedures.

1. WISC kits may be checked out at Manzanita Center for a 24 hour period. All record forms and Mazes may be purchased for 17¢ each from the Manzanita secretary. The WISC Manuals will be purchased by you at the UNM bookstore.

October 29 [illegible]

prior to September [illegible]

prior to the final [illegible]

have feedback [illegible]

other chance to [illegible] WISC [illegible]

allow the graduate student to [illegible]

rating instrument. [illegible]

tration will [illegible]

2. Observations of WISC administrations. [illegible] to observe 3 WISC administrations [illegible] servation booth when [illegible] of Mercantile Center [illegible] scheduled.

You will make the arrangements with the Mercantile Center [illegible] for your final demonstration (week of October 28 [illegible]) [illegible] couraged to do this in advance of [illegible] sections of Guidance 512 will be [illegible] periods available (due to space and [illegible]

Miscellaneous Procedures.

1. WISC kits may be checked [illegible] period. All [illegible] each from the Mercantile [illegible] purchased by you at the [illegible]

2. Whenever you wish to schedule a WISC administration at Manzanita Center, check with the Manzanita secretary. The graduate assistant will have left his time schedules with her so she will know when he is available to observe your session.

3. Dr. Heisey and the graduate assistant will be available for appointments, conferences, or feedback. Check with them regarding their available times.

4. During the first class meeting, you will be issued a Manzanita Center observation booth pass and a procedure sheet describing use of the booth.

5. Manzanita Center has a few stopwatches which may be used <u>only</u> for tests administered in the Center. You will need your own stopwatch for the tests administered outside of the Center.

APPENDIX B

CONSENT FORM

Date________________________

I hereby give Kathy Ritter permission to look at my admission folder in the Guidance office in order to determine my undergraduate grade point average. If this information is not available in the Guidance office, she has my permission to check with the Admissions Office or the Graduate School. I understand that she will protect my confidence and not disclose this information, or in any way use my name in her final report.

Signed____________________________

APPENDIX C

SCHEDULE FOR THE EXPERIMENTAL GROUP

(Group I) Workshop
September 8, 1973

Each student will bring the following items:

1. WISC Manual
2. 5 WISC record forms and 5 WISC mazes
3. Several pencils with erasers
4. Stopwatch

Morning:

8:30 to 9:30 - WISC demonstration by Dr. Heisey (Verbal subtests). Students will follow in their Manual and score their protocols as the demonstration proceeds.

9:30 to 10:00 - Discussion of previous WISC administration

10:00 to 10:15 - Break

10:15 to 11:15 - Students will break into dyads and one of each pair will administer the Verbal subtests of the WISC to the other. (A kit will be provided for each pair.)

11:15 to 12:00 - Discussion of student WISC administrations

12:00 to 1:00 - Lunch

Afternoon:

1:00 to 2:00 - WISC demonstrations by Dr. Heisey (Performance subtests). Students will follow in their Manual and score their protocols as the demonstration proceeds.

2:00 to 2:30 - Discussion of previous WISC administration

2:30 to 2:45 - Break

2:45 to 3:45 - Dyads will reverse and the original "subtest" will administer the Performance subtests of the WISC to the other.

3:45 to 4:30 - Discussion of the student WISC administrations and final questions

APPENDIX D

FINAL EXAMINATION
WECHSLER INTELLIGENCE SCALE FOR CHILDREN

In each of the following questions, one of the four answers is the most nearly correct. Circle the letter of the correct answer. For example, if "a" is the correct answer for a certain question, the "a" will be circled as follows: a

1. Which of the following is an accurate reflection of the difference between the Binet and Wechsler scales?

   a) Both are considered to be excellent tests for measuring level of intellectual functioning at all ability levels.
   b) The Wechsler tests depend more heavily upon verbal items than does the Binet scale.
   c) The Wechsler scales include separate subtests for measuring separate abilities whereas the Binet does not.
   d) The concept of age level is more central to the Wechsler scales than to the Binet test.

2. Which of the following was NOT a contribution of Alfred Binet?

   a) He devised an instrument which measured individual mental faculties.
   b) He constructed the first individual intelligence test.
   c) He constructed the first instrument which differentiated dull from normal children.
   d) He devised an instrument built around the theory that mental faculties were interwoven and therefore could not be measured independently.

3. A .82 correlation has been found between Binet and Wechsler IQ scores. This means that:

   a) a person who does well on one will tend to do poorly on the other.
   b) if a person took either scale several times in succession, his/her scores would tend to be similar.
   c) if a person took both tests, 82% of the time the scores would differ.
   d) the two tests measure essentially the same things.

4. All of the following, except one, are considered to be assets of the WISC. Which is a limitation?

   a) The WISC covers a broad range of tasks.
   b) The WISC has high Performance reliability.
   c) The WISC provides opportunities for observation of behavior and thought processes.
   d) The WISC standardization sample is based on geographic area of residence, urban-rural residence, and occupational group of fathers of children.

5. Sampling, in the case of an individual intelligence test, refers primarily to:

   a) that evidence of ability that is considered to be reflective of one's total ability.
   b) the ability to generalize precisely from an IQ <u>score</u> to the full range of one's general ability.
   c) the idea that general ability can be broken down into component parts.
   d) those combinations of mental factors that enter into the determination of one's general ability.

6. Which of the following is <u>least</u> likely to depress a person's IQ score?

   a) Home environment.
   b) Genetic and heredity determinants.
   c) School experiences.
   d) Neurological factors.

7. Which of the following is <u>generally</u> considered to be an asset of the Wechsler Performance section?

   a) It is relatively independent of educational and cultural background.
   b) It is less reliable than the Verbal section at all ages.
   c) It is more subject to anxiety than the Verbal section.
   d) It has a greater SE_m than the Verbal section at all age levels.

8. Which is the most accurate statement regarding intellectual functioning?

a) The research results regarding the influence of heredity conclude that intelligence is primarily fixed at birth.
b) Since memory and memory storage are functions of the brain and central nervous system which are present in an individual at birth, one's intellectual functioning is largely determined at the time of birth.
c) Since caffeine and other chemicals can alter intellectual output, heredity and central nervous system influence can be largely overcome.
d) The research results regarding intellectual functioning are complex and far reaching, and the intelligence quotient is influenced by many factors and variables, no one of which is the singular influence.

9. Which is the most accurate statement regarding the intelligence quotient?

a) It is a fixed unit of measurement which remains stable throughout one's lifetime.
b) It cannot be measured directly, but must be obtained by a process of inference from observable performance.
c) It is constantly in a state of fluctuation, but, in most individuals, an optimum performance is reached at approximately age 13.
d) The IQ score is subject to so many influences that its meaning is actually negligible for all practical purposes.

10. Validity refers to:

a) the consistency of a person's scores on a series of measures.
b) the calculated estimate of amount of variation that occurs within a series of measurements.
c) the extent to which a test measures that for which it was designed to measure.
d) whether or not a certain test score is an accurate measure of general ability for an individual person.

11. Which of the following words is <u>un</u>related to the concept of validity?

a) Predictive.
b) Consistency.
c) Concurrent.
d) Content.

12. Reliability refers to:

    a) the consistency of a person's scores on a series of measures.
    b) the amount by which an obtained score differs from a hypothetical true score.
    c) the extent to which a test measures that for which it was designed to measure.
    d) how much the test can actually tell the examiner about a subject's performance.

13. For a 7-1/2 year old child, the SE_m on the Information subtest is 1.75. This means that:

    a) the person's true score differs by no more than 1.75 IQ points from the obtained score.
    b) the reliability is fairly low compared with other subtests.
    c) the chances are about two out of three that a true score on this test is within $\pm$ 1.75 points of the obtained scale score.
    d) chances are about two out of three that the true score has been obtained.

14. Every subtest on the WISC correlates to some degree with every other subtest and with the Verbal, Performance, and Full Scale IQ. This means that:

    a) each subtest measures a truly independent intellectual factor.
    b) common factors underlie all of the WISC subtests.
    c) every subtest measures precisely the same factors as every other subtest.
    d) many different factors make up intelligence and the difference is reflected in the intercorrelations.

15. Information, more so than any other of the WISC subtests, reflects:

    a) visual organization.
    b) planning and organization.
    c) spatial relationships.
    d) early learning and cultural background.

16. The Comprehension subtest minimizes the importance of which one of the following?

    a) Common sense.
    b) Ability to evaluate and use past experience.
    c) Factual recall.
    d) Judgment.

17. The chief advantage of the Comprehension subtest is:

    a) the cultural fairness aspect.
    b) it yields a rich mine of clinically relevant material.
    c) the ease of administration.
    d) it yields the highest correlation with the Full Scale IQ of any of the WISC subtests.

18. The Picture Competion subtest assesses mainly:

    a) verbal learning.
    b) abstract reasoning.
    c) the ability to differentiate essential from non-essential details.
    d) the ability to perceive common elements in terms one is asked to compare.

19. All of the following, except one, could generally be sources of low scores on the Picture Completion subtest. Select that one.

    a) Mental retardation.
    b) Poor visual-motor coordination.
    c) Poor reality testing.
    d) Poor attention and concentration due to anxiety.

20. Which WISC subtest essentially measures one's ability to perceive and analyze forms and/or to analyze the whole into its component parts?

    a) Picture Arrangement.
    b) Picture Completion.
    c) Object Assembly.
    d) Block Design.

21. Which WISC subtest has the highest correlation with general intelligence?

a) Block Design.
b) Coding.
c) Vocabulary.
d) Similarities.

22. Select the subtest that is considered to be "culture fair" and, thus, one of the best single indicators of g for the bilingual child.

a) Block Design.
b) Similarities.
c) Vocabulary.
d) Comprehension.

23. Select the trait <u>not</u> measured by the Mazes subtest.

a) Visual-motor coordination.
b) Factual recall.
c) The ability to follow a visual pattern.
d) The ability to plan ahead.

24. Low scores on the Block Design subtest may be indicative of all of the following, <u>except</u> one. Select that one.

a) Visual-motor difficulties.
b) Color blindness.
c) Limited educational background.
d) Poor perceptual and spatial conceptualization.

25. The WISC subtest that is least vulnerable to emotional disturbance (or most resistant to deterioration in an emotionally disturbed person) is:

a) Vocabulary.
b) Arithmetic.
c) Digit Span.
d) Object Assembly.

21. Which WISC subtest has the highest correlation with general intelligence?

a) Block Design.
b) Coding.
c) Vocabulary.
d) Similarities.

. . . .

22. Select the subtest that is considered to be "culture fair" and, thus, one of the best single indicators of g for the bilingual child.

a) Block Design.
b) Similarities.
c) Vocabulary.
d) Comprehension.

23. Select the trait not measured by the Mazes subtest.

a) Visual-motor coordination.
b) Factual recall.
c) The ability to follow a visual pattern.
d) The ability to plan ahead.

24. Low scores on the Block Design subtest may be indicative of all of the following, except one. Select that one.

a) Visual-motor difficulties.
b) Color blindness.
c) Limited educational background.
d) Poor perceptual and spatial conceptualization.

25. The WISC subtest that is least vulnerable to emotional disturbance (is most resistant to deterioration in an emotionally disturbed person) is:

a) Vocabulary.
b) Arithmetic.
c) Digit Span.
d) Object Assembly.

26. The subtest most subject to anxiety is:

    a) Vocabulary.
    b) Digit Span.
    c) Information.
    d) Mazes.

27. The Coding subtest essentially measures:

    a) rote memory.
    b) anticipation and planning.
    c) concentration.
    d) visual-motor dexterity.

28. The Digit Span subtest essentially measures:

    a) associative thinking.
    b) the ability to separate essential from non-essential elements.
    c) rote memory.
    d) social intelligence.

29. Which definition best fits the Picture Arrangement Subtest?

    a) It is a test which essentially measures one's ability to size up a total situation.
    b) It is a test which assesses the ability to differentiate essential from non-essential details.
    c) It is a test which measures one's ability to perceive the common elements in terms he is asked to compare.
    d) It is a test which assesses the ability to learn and write symbols.

30. Which WISC subtest best measures the ability to bring the terms in each item under a single concept?

    a) Picture Arrangement.
    b) Comprehension.
    c) Similarities.
    d) Block Design.

31. The two WISC subtest scores that provide the most accurate estimate of scholastic achievement are:

    a) Block Design and Information.
    b) Comprehension and Arithmetic.
    c) Vocabulary and Block Design.
    d) Vocabulary and Arithmetic.

32. One factor not measured by the Object Assembly subtest is:

    a) synthesis of parts into an organized whole.
    b) verbal fluency.
    c) perceptual-motor coordination.
    d) spatial relationships.

33. Examiner variables, as described by Glasser and Zimmerman, refer to:

    a) those variations in calculated IQ score which may exist from examiner to examiner.
    b) those reactions of the subject to the testing situation.
    c) those contributions of the individual examiner which may influence the reactions of the subject to the testing situation.
    d) the accuracy of the examiner in following standard testing procedure.

34. Which of the following is NOT a consideration involved in "testing the limits" (extension testing)?

    a) Any response obtained by this method cannot be scored.
    b) It should be done routinely to ascertain the subject's potential or ceiling.
    c) If there is any question as to the validity of the results, one should consider extension testing.
    d) One should use extension testing primarily to discover more about the child than the standardization procedure allows.

35. Select the most accurate statement in regard to brief forms of the WISC.

    a) Careful consideration must be given before using a brief form in that information regarding the child's functioning in various areas could be lost.
    b) Brief forms should be given routinely as a quick, initial screening device.
    c) Dropping one or two subtests is often a useful, time-saving device.
    d) In order to maintain standard procedure, the same subtests must be administered every time a brief form is given.

36. The three WISC subtests which, when combined, are considered the best predictors of the Full Scale IQ, in all but the retarded, are:

    a) Block Design, Picture Arrangement, and Vocabulary.
    b) Arithmetic, Digit Span, and Vocabulary.
    c) Vocabulary, Picture Arrangement, and Arithmetic.
    d) Vocabulary, Arithmetic, and Block Design.

37. If 6 Verbal subtests were given, prorating is done by:

    a) multiplying the sum of the Verbal Raw Scores by 5/4.
    b) multiplying the sum of the Verbal Scaled Scores by 5/4.
    c) multiplying the sum of the Verbal Raw Scores by 5/6.
    d) multiplying the sum of the Verbal Scaled Scores by 5/6.

38. In order to convert Raw Scores to Scaled Scores correctly, one prior determination is essential.

    a) Conversion of Scaled Scores into IQ scores.
    b) Computation of the age of the child.
    c) Prorating the scores.
    d) Accurate determination of the V-P discrepancy.

39. Using the WISC standardization population, the mean discrepancy between the Verbal IQ and the Performance IQ for the entire sample was 0 and the standard deviation was 12.5 IQ points. This means that:

a) In 2/3 of the cases, the V-P discrepancy would be $\pm$ 12.5 IQ points or less.
b) In 95% of the cases, the V-P discrepancy in either direction would average 12.5 IQ points.
c) V-P discrepancy is nonexistent.
d) 12.5% of the time, there would be a difference between Verbal IQ and Performance IQ.

40. Select the statement which is NOT true in regard to Verbal and Performance IQ.

a) A high relationship has been found between academic success and higher Verbal IQ.
b) A V-P discrepancy should be considered significant in every case if the difference is greater than 15 IQ points.
c) In 2/3 of the cases, one would be accurate in assuming neurological impairment if the Verbal IQ is 25 points higher than the Performance IQ.
d) Since one would <u>generally</u> expect those with high IQs to have a higher Verbal IQ and those with low IQ's to have a higher Performance IQ, a 15 point discrepancy would often be true only for those within the middle IQ range.

41. Select the statement that is <u>more</u> accurate for the bilingual child than for others.

a) The bilingual child will often score higher on the Verbal subtests than on the Performance subtests.
b) Bilingual children tend to have higher Performance IQs than Verbal IQs.
c) The bilingual child will often score higher on the Comprehension subtest than on the Block Design subtest.
d) The Performance IQ will often be 25 points or more higher than the Verbal IQ.

42. In the case of an invalid subtest, generally the best procedure to follow is:

    a) to give all 6 subtests in that section and prorate accordingly basing the the IQ on 5 subtests.
    b) to decide if the error is major or minor and, if it's not too large, give the remainder of the subtest and include the subtest score in the sum of scores.
    c) to check to see if the error would affect the IQ score.
    d) to give all 6 subtests in that section and prorate accordingly basing the IQ on the number of subtests correctly administered.

43. Which is NOT an accurate statement in regard to inquiry by the WISC examiner?

    a) Inquiry must be as pointed and specific as possible.
    b) If the child gives a definite 1 point response, the examiner does not try to elicit further elaboration.
    c) Inquiry must be neutral and non-leading.
    d) Too much questioning by the examiner could violate standardization and artificially alter scores.

44. Which is the MOST correct statement in regard to scoring of answers?

    a) Scoring must be done at the time the test is administered in order that the examiner doesn't "forget" the response.
    b) Since one or two raw score points don't usually affect the IQ score, the examiner can often estimate the point value of responses if he/she is uncertain.
    c) In order to increase scoring precision, it is wise to "double check" the score given to each response.
    d) Since "first impressions" are often correct, an examiner should learn to trust his/her judgment in scoring.

45. In which of the following subtests may the examiner NOT repeat the question if the child doesn't understand?

    a) Vocabulary.
    b) Arithmetic.
    c) Digit Span.
    d) Information.

46. Standard Scores are:

a) conversions which have been made of raw scores into scores which describe a raw score's position in a distribution.
b) stanines.
c) scores which have a standard deviation of 2.
d) WISC Scaled Scores.

47. There are several advantages in using a distribution such as a stanine in explaining IQ scores to parents. Which of the following is NOT an advantage?

a) The general ability of the child can be explained in terms of a range as opposed to a specific score.
b) Each stanine unit is equal.
c) Parents may have trouble comprehending the idea of a stanine distribution of scores.
d) Since both the WISC IQ and the stanine are deviation scores based on the normal curve, no serious distortions of basic principles occur when changing from a WISC IQ to a stanine.

48. In writing a report of WISC test results, all of the following, except one, must be included in the Conclusions and Recommendations section. Select that one.

a) Identifying information (name, date of birth, school, etc.).
b) Patterns of strengths and weaknesses.
c) Tentative hypothesis(es).
d) Specific recommendations based upon findings.

49. If, during the Mazes subtest, the subject lifts his/her pencil from the paper, the examiner should:

a) discontinue the subtest.
b) score a failure and move on to the next Maze if the subject doesn't already have 2 consecutive failures (scores of 0).
c) score 1 error and discontinue timing.
d) remind the subject to keep the pencil on the paper and score 1 error.

50. On the Object Assembly subtest, the subject should first see the puzzle pieces:

a) before the directions are given.
b) while the examiner is laying out the pieces.
c) after the directions are given.
d) while they are in the box.

51. If, on the Picture Completion subtest, the subject points to the right place on the card but spoils it by making the wrong verbal response, the examiner should:

a) score 0 on the item.
b) give the subject 1 point credit for the item.
c) ask the subject what important part is missing.
d) ask the subject to explain more fully.

52. Which WISC subtest has the lowest reliability?

a) Vocabulary.
b) Block Design.
c) Information.
d) Digit Span.

53. On the Picture Arrangement subtest, the examiner should start the stopwatch:

a) as soon as the cards are laid out in their order or presentation.
b) while the cards are being laid out.
c) before laying out the cards.
d) while the cards are being taken out of the box.

54. All of the following, except one, are reasons for knowing a testmaker's definition of intelligence. Select the one which doesn't apply.

a) The examiner can decide whether he/she agrees with the testmaker and whether, on that basis, to select the test.
b) Knowledge of the testmaker's definition of intelligence would greatly aid in the interpretation of the results.
c) Knowing the testmaker's definition of intelligence would be helpful in selecting the appropriate test for a particular subject.
d) Knowledge of the testmaker's definition of intelligence would aid the examiner in administering the test according to standardization procedure.

APPENDIX E

COGNITIVE EXAMINATION ITEM COUNT
WITH NUMBER AND PERCENT OF TOTAL GROUP RESPONDING TO EACH RESPONSE OPTION

Item Number		Experimental No.	Experimental %	Control No.	Control %	Combined No.	Combined %
Item 1	a=	3	27	1	7	4	16
	b=	1	9	0	0	1	4
	*c=	7	64	11	79	18	72
	d=	0	0	2	14	2	8
Item 2	0=	0	0	1	7	1	4
	*a=	2	18	4	29	6	24
	b=	1	9	2	14	3	12
	c=	0	0	4	29	4	16
	d=	8	73	3	21	11	44
Item 3	a=	0	0	0	0	0	0
	b=	2	18	1	7	3	12
	c=	0	0	0	0	0	0
	*d=	9	82	13	93	22	88
Item 4	a=	0	0	0	0	0	0
	b=	4	36	2	14	6	24
	c=	0	0	1	7	1	4
	*d=	7	64	11	79	18	72
Item 5	*a=	8	73	6	43	14	56
	b=	0	0	2	14	2	8
	c=	1	9	4	29	5	20
	d=	2	18	2	14	4	16

* correct answer

Item Number		Experimental No.	Experimental %	Control No.	Control %	Combined No.	Combined %
Item 6	0=	0	0	1	7	1	4
	a=	0	0	0	0	0	0
	b=	4	36	7	50	11	44
	*c=	6	55	4	29	10	40
	d=	1	9	2	14	3	12
Item 7	*a=	11	100	13	93	24	96
	b=	0	0	0	0	0	0
	c=	0	0	0	0	0	0
	d=	0	0	1	7	1	4
Item 8	a=	0	0	0	0	0	0
	b=	0	0	1	7	1	4
	c=	0	0	0	0	0	0
	*d=	11	100	13	93	24	96
Item 9	a=	0	0	0	0	0	0
	*b=	6	55	10	71	16	64
	c=	5	45	4	29	9	36
	d=	0	0	0	0	0	0
Item 10	a=	0	0	0	0	0	0
	b=	0	0	0	0	0	0
	*c=	11	100	13	93	24	96
	d=	0	0	1	7	1	4
Item 11	a=	2	18	2	14	4	16
	*b=	6	55	8	57	14	56
	c=	2	18	2	14	4	16
	d=	1	9	2	14	3	12
Item 12	*a=	10	91	7	50	17	68
	b=	0	0	5	36	5	20
	c=	0	0	2	14	2	8
	d=	1	9	0	0	1	4

Item Number		Experimental No.	Experimental %	Control No.	Control %	Combined No.	Combined %
Item 13	a=	2	18	1	7	3	12
	b=	2	18	0	0	2	8
	*c=	7	64	13	93	20	80
	d=	0	0	0	0	0	0
Item 14	a=	0	0	0	0	0	0
	*b=	4	36	6	43	10	40
	c=	0	0	0	0	0	0
	d=	7	64	8	57	15	60
Item 15	a=	0	0	0	0	0	0
	b=	0	0	0	0	0	0
	c=	0	0	0	0	0	0
	*d=	11	100	14	100	25	100
Item 16	a=	0	0	0	0	0	0
	b=	0	0	0	0	0	0
	*c=	11	100	14	100	25	100
	d=	0	0	0	0	0	0
Item 17	a=	2	18	0	0	2	8
	*b=	9	82	14	100	23	92
	c=	0	0	0	0	0	0
	d=	0	0	0	0	0	0
Item 18	a=	0	0	0	0	0	0
	b=	0	0	0	0	0	0
	*c=	10	91	14	100	24	96
	d=	1	9	0	0	1	4
Item 19	a=	2	18	1	7	3	12
	*b=	6	55	12	86	18	72
	c=	3	27	1	7	4	16
	d=	0	0	0	0	0	0

Item Number		Experimental No.	Experimental %	Control No.	Control %	Combined No.	Combined %
Item 20	a=	0	0	0	0	0	0
	b=	0	0	0	0	0	
	c=	4	36	5	36	9	36
	*d=	7	64	9	64	16	64
Item 21	a=	0	0	0	0	0	0
	b=	0	0	0	0	0	0
	*c=	11	100	14	100	25	100
	d=	0	0	0	0	0	0
Item 22	*a=	8	73	14	100	22	88
	b=	1	9	0	0	1	4
	c=	0	0	0	0	0	0
	d=	2	18	0	0	2	8
Item 23	a=	0	0	0	0	0	0
	*b=	11	100	14	100	25	100
	c=	0	0	0	0	0	0
	d=	0	0	0	0	0	0
Item 24	a=	0	0	0	0	0	0
	b=	2	18	0	0	2	8
	*c=	9	92	14	100	23	92
	d=	0	0	0	0	0	0
Item 25	*a=	7	64	10	71	17	68
	b=	1	9	3	21	4	16
	c=	0	0	1	7	1	4
	d=	3	27	0	0	3	12
Item 26	a=	1	9	0	0	1	4
	*b=	9	82	13	93	22	88
	c=	0	0	1	7	1	4
	d=	1	9	0	0	1	4

Item Number		Experimental No.	Experimental %	Control No.	Control %	Combined No.	Combined %
Item 27	a=	2	18	1	7	3	12
	b=	0	0	0	0	0	0
	c=	1	9	2	14	3	12
	*d=	8	73	11	79	19	76
Item 28	0=	0	0	1	7	1	4
	a=	0	0	1	7	1	4
	b=	0	0	0	0	0	0
	*c=	11	100	12	86	23	92
	d=	0	0	0	0	0	0
Item 29	*a=	11	100	14	100	25	100
	b=	0	0	0	0	0	0
	c=	0	0	0	0	0	0
	d=	0	0	0	0	0	0
Item 30	a=	1	9	1	7	2	8
	b=	0	0	0	0	0	0
	*c=	9	82	13	93	22	88
	d=	1	9	0	0	1	4
Item 31	a=	0	0	0	0	0	0
	b=	0	0	0	0	0	0
	c=	1	9	0	0	1	4
	*d=	10	91	14	100	24	96
Item 32	a=	0	0	0	0	0	0
	*b=	11	100	14	100	25	100
	c=	0	0	0	0	0	0
	d=	0	0	0	0	0	0
Item 33	a=	0	0	3	21	3	12
	b=	0	0	0	0	0	0
	*c=	11	100	11	79	22	88
	d=	0	0	0	0	0	0

Item Number		Experimental No.	Experimental %	Control No.	Control %	Combined No.	Combined %
Item 34	a=	3	27	1	7	4	16
	*b=	7	64	7	50	14	56
	c=	0	0	5	36	5	20
	d=	1	9	1	7	2	8
Item 35	*a=	11	100	14	100	25	100
	b=	0	0	0	0	0	0
	c=	0	0	0	0	0	0
	d=	0	0	0	0	0	0
Item 36	a=	1	9	9	64	10	40
	b=	0	0	0	0	0	0
	c=	5	45	4	29	9	36
	*d=	5	45	1	7	6	24
Item 37	a=	0	0	0	0	0	0
	b=	0	0	0	0	0	0
	c=	3	27	1	7	4	16
	*d=	8	73	13	93	21	84
Item 38	a=	0	0	0	0	0	0
	*b=	11	100	14	100	25	100
	c=	0	0	0	0	0	0
	d=	0	0	0	0	0	0
Item 39	*a=	9	82	9	64	18	72
	b=	2	18	2	14	4	16
	c=	0	0	2	14	2	8
	d=	0	0	1	7	1	4
Item 40	a=	0	0	0	0	0	0
	*b=	2	18	2	14	4	16
	c=	2	18	4	29	6	24
	d=	7	64	8	57	15	60

Item Number		Experimental No.	Experimental %	Control No.	Control %	Combined No.	Combined %
Item 41	a=	0	0	0	0	0	0
	*b=	11	100	14	100	25	100
	c=	0	0	0	0	0	0
	d=	0	0	0	0	0	0
Item 42	*a=	2	18	5	36	7	28
	b=	3	27	0	0	3	12
	c=	0	0	0	0	0	0
	d=	6	55	9	64	15	60
Item 43	*a=	7	64	13	93	20	80
	b=	3	27	0	0	3	12
	c=	1	9	1	7	2	8
	d=	0	0	0	0	0	0
Item 44	a=	2	18	0	0	2	18
	b=	0	0	0	0	0	0
	*c=	9	82	14	100	23	92
	d=	0	0	0	0	0	0
Item 45	a=	0	0	0	0	0	0
	b=	0	0	0	0	0	0
	*c=	11	100	14	100	25	100
	d=	0	0	0	0	0	0
Item 46	*a=	6	55	6	43	12	48
	b=	3	27	3	21	6	24
	c=	1	9	1	7	2	8
	d=	1	9	4	29	5	20
Item 47	a=	0	0	0	0	0	0
	b=	2	18	3	21	5	20
	*c=	8	73	11	79	19	76
	d=	1	9	0	0	1	4

Item Number		Experimental No.	Experimental %	Control No.	Control %	Combined No.	Combined %
Item 48	*a=	8	73	12	86	20	80
	b=	0	0	0	0	0	0
	c=	2	18	2	14	4	16
	d=	1	9	0	0	1	4
Item 49	a=	0	0	0	0	0	0
	b=	0	0	1	7	1	4
	c=	0	0	1	7	1	4
	*d=	11	100	12	86	23	92
Item 50	a=	0	0	7	50	7	28
	b=	0	0	0	0	0	0
	*c=	10	91	7	50	17	68
	d=	1	9	0	0	1	4
Item 51	*a=	3	27	1	7	4	16
	b=	4	36	12	86	16	64
	c=	3	27	0	0	3	12
	d=	1	9	1	7	2	8
Item 52	a=	0	0	0	0	0	0
	b=	4	36	0	0	4	16
	c=	0	0	0	0	0	0
	*d=	7	64	14	100	21	84
Item 53	*a=	11	100	14	100	25	100
	b=	0	0	0	0	0	0
	c=	0	0	0	0	0	0
	d=	0	0	0	0	0	0
Item 54	a=	3	27	3	21	6	24
	b=	0	0	0	0	0	0
	c=	0	0	1	7	1	4
	*d=	8	73	10	71	18	72

APPENDIX F

EXPERIMENTAL GROUP CORRELATION MATRIX

	1. AGE	2. SEX	3. GPA	4. ACH	5. DEF	6. ORD	7. EXH	8. AUT	9. AFF	10. INT	11. SUC	12. DOM	13. ABA	14. NUR	15. CHG	16. END	17. HET	18. AGG	19. COG EXAM	20. ITEM 1	21. ITEM 2	22. ITEM 3	23. ITEM 4	24. ITEM 5	25. ITEM 6	26. ITEM 7	27. ITEM 8	28. ITEM 9	29. ITEM 10	30. ITEM 11	31. ITEM 12	32. PERF AV	33. #OBS	34. #ADM	35. #PROT	36. #DEMO	
1.		-49	34	19	26	35	-17	-30	-24	-21	51	-37	71	-14	05	-13	-16	-38	49	00	58	51	10	30	60	31	01	61	58	36	37	45	-40	26	50	79	1.
2.			-06	44	-73	-32	70	39	-06	35	-19	57	-38	-38	-65	-37	-09	54	-09	-47	-39	-31	-46	-41	-52	-44	-41	-42	-14	-21	-26	-43	05	-04	07	-03	2.
3.				-05	-06	75	-18	32	-38	-73	19	01	40	-30	-17	16	-10	00	07	-14	-15	-15	-20	-07	14	08	03	-06	30	27	05	00	-37	01	03	13	3.
4.					-47	-28	25	-18	32	55	49	22	-05	-50	-46	-92	-32	48	28	-14	39	42	-17	-08	29	-32	-46	29	37	26	24	14	-17	11	57	37	4.
5.						18	-57	-42	-02	-14	-07	04	17	39	69	43	-28	-58	08	35	30	22	45	30	11	39	50	20	06	04	-02	28	-28	13	05	02	5.
6.							-21	-03	-18	-53	-05	-11	32	31	08	38	-31	-48	33	20	08	01	29	40	22	43	27	22	47	62	34	34	-42	38	03	15	6.
7.								-03	18	36	03	22	-39	-02	-51	-33	-08	-07	00	-69	-36	-30	-50	-32	-53	-55	-63	-39	-16	-21	-07	-45	17	-02	01	18	7.
8.									-56	-45	-39	19	25	-47	-47	33	30	63	-55	-30	-61	-57	-51	-55	-41	-11	-08	-50	-44	-42	-66	-53	02	-47	-36	-20	8.
9.										52	30	-05	-55	31	-03	-47	-27	-11	-16	-05	05	10	00	-06	-04	-48	-43	-07	-22	12	19	-08	32	-21	-09	-30	9.
10.											-13	30	-41	06	-22	-52	-27	11	26	11	32	31	23	14	-04	-11	-12	21	10	13	24	17	16	20	49	20	10.
11.												-38	13	-19	10	-66	-09	-04	13	-27	35	42	-28	-05	45	-38	-51	24	22	10	23	08	-06	-05	06	16	11.
12.													-39	-27	-20	-13	-50	24	-17	-25	-29	-36	-20	-36	-59	-26	-10	-43	-07	-22	-42	-35	-50	05	16	-05	12.
13.														-18	-17	28	-03	-02	27	09	40	42	08	23	46	53	25	49	30	31	01	36	-15	01	33	62	13.
14.															48	39	-27	-71	20	35	19	17	53	58	-01	32	25	19	03	36	23	31	08	38	-32	-24	14.
15.																26	10	-50	14	53	29	19	53	45	29	33	48	24	17	00	22	36	-20	37	-25	-34	15.
16.																	22	-35	-17	24	-31	-33	26	15	-24	52	60	-17	-25	-09	-23	-01	10	-06	-35	-18	16.
17.																		24	-25	09	-32	-31	-10	-17	10	00	17	-16	-29	-45	05	-15	57	-31	-28	-26	17.
18.																			-36	-10	-22	-13	-38	-48	-06	-29	-20	-22	-26	-28	-36	-30	17	-42	00	-21	18.
19.																				48	79	75	66	84	65	63	43	83	91	77	81	86	-23	86	73	61	19.
20.																					59	49	93	79	64	79	83	67	49	55	57	81	-04	54	23	-10	20.
21.																						96	67	77	82	57	36	96	76	69	64	89	-24	63	64	52	21.
22.																							57	71	78	49	28	90	65	67	58	82	-05	50	60	47	22.
23.																								91	58	83	81	72	59	66	67	88	-09	70	36	08	23.
24.																									68	82	66	84	74	80	76	94	-12	83	39	26	24.
25.																										55	38	89	72	64	75	85	-07	45	46	35	25.
26.																											88	71	59	59	46	81	-18	61	38	32	26.
27.																												45	38	34	35	65	-03	46	26	02	27.
28.																													81	76	72	94	-19	66	62	55	28.
29.																														78	77	83	-51	85	69	57	29.
30.																															73	82	-22	66	50	36	30.
31.																																80	04	66	53	32	31.
32.																																	-17	76	58	39	32.
33.																																		-49	-21	-29	33.
34.																																			44	33	34.
35.																																				80	35.
36.																																					36.

.602 needed for significance at .05 level with df = 9
.735 needed for significance at .01 level with df = 9

Note: Decimal points have been omitted.

CONTROL GROUP CORRELATION MATRIX

	1. AGE	2. SEX	3. GPA	4. ACH	5. DEF	6. ORD	7. EXH	8. AUT	9. AFF	10. INT	11. SUC	12. DOM	13. ABA	14. NUR	15. CHG	16. END	17. HET	18. AGG	19. COG EXAM	20. ITEM 1	21. ITEM 2	22. ITEM 3	23. ITEM 4	24. ITEM 5	25. ITEM 6	26. ITEM 7	27. ITEM 8	28. ITEM 9	29. ITEM 10	30. ITEM 11	31. ITEM 12	32. PERF AV	33. #OBS	34. #ADM	35. #PROT	36. #DEMO	
1.		-14	-08	02	-34	-34	-22	06	30	-39	06	21	-17	-19	66	14	35	-52	-58	20	42	36	36	23	34	-12	-11	19	15	-22	-18	21	(not applicable)				1.
2.			05	42	20	34	-43	27	-40	27	-42	27	-07	-38	-19	02	-01	22	-32	39	-22	-15	40	35	32	54	43	27	-75	-20	16	21					2.
3.				05	-33	12	-14	26	-01	00	31	-42	-19	15	11	-29	42	10	44	-12	-13	-10	12	27	-17	06	22	24	-22	30	24	12					3.
4.					-15	-10	-32	10	06	13	-36	02	21	-19	-25	19	-18	-21	-09	-15	-38	-36	08	14	36	00	38	-05	-33	07	46	06					4.
5.						10	-04	-19	-28	48	-15	-14	43	01	-57	49	-65	-31	02	09	08	15	-21	-24	-58	11	-27	-16	-07	-07	-14	-21					5.
6.							02	-13	-36	09	21	07	-08	-21	-03	-36	01	41	10	53	-12	-05	54	32	16	53	51	30	-34	55	52	52					6.
7.								02	-04	-52	04	05	-04	-10	34	00	-08	24	-09	33	26	29	-08	-35	-01	16	-17	05	54	38	00	21					7.
8.									-62	-09	-12	59	-60	-61	13	08	57	08	07	33	39	27	45	50	36	18	22	68	12	00	15	51					8.
9.										-29	05	-44	27	49	25	-18	-23	-11	-23	-36	-32	-26	-43	-16	-18	-28	-20	-49	06	-06	-32	-45					9.
10.											00	-17	15	45	-82	-12	-48	-11	40	-37	-50	-53	-13	05	-20	05	29	-33	-42	-19	18	-28					10.
11.												-04	-39	29	07	-37	19	-24	25	-13	-12	-21	-09	18	-07	-31	-01	-27	21	13	17	-06					11.
12.													-70	-69	17	03	34	14	-39	57	33	20	50	48	63	31	32	34	17	-09	00	51					12.
13.														31	-24	37	-63	-27	04	-38	-17	-05	-42	-52	-44	-28	-44	-26	-03	-03	09	-38					13.
14.															-31	-43	-41	-06	33	-62	-62	-60	-60	-37	-36	-35	-13	-70	-16	-22	-12	-69					14.
15.																-16	53	05	-55	56	44	43	41	21	41	11	-11	44	28	07	-06	44					15.
16.																	-12	-51	-09	-19	57	61	-18	-42	-39	-20	-50	04	28	-09	-30	-10					16.
17.																		21	09	22	48	41	40	28	31	07	04	57	04	-01	-12	35					17.
18.																			14	34	-18	-12	15	08	17	50	38	19	-21	20	-05	18					18.
19.																				-57	-12	-14	-27	-17	-47	-30	02	-05	03	33	16	-19					19.
20.																					32	33	79	46	57	75	41	61	-02	27	24	81					20.
21.																						97	34	-09	-11	02	-36	59	48	13	-33	41					21.
22.																							32	-18	-21	10	-37	57	38	19	-37	38					22.
23.																								64	58	66	64	76	-19	41	42	92					23.
24.																									50	36	58	48	-10	17	45	60					24.
25.																										26	54	30	-12	-07	45	53					25.
26.																											69	37	-31	40	15	62					26.
27.																												20	-42	45	52	56					27.
28.																													00	33	25	79					28.
29.																														15	-07	12					29.
30.																															48	60					30.
31.																																53					31.
32.																																					32.
33.																																					33.
34.																																					34.
35.																																					35.
36.																																					36.

.532 needed for significance at .05 level with df = 12
.661 needed for significance at .01 level with df = 12

Note: Decimal points have been omitted.

COMBINED GROUP CORRELATION MATRIX

	1. AGE	2. SEX	3. GPA	4. ACH	5. DEF	6. ORD	7. EXH	8. AUT	9. AFF	10. INT	11. SUC	12. DOM	13. ABA	14. NUR	15. CHG	16. END	17. HET	18. AGG	19. COG EXAM	20. ITEM 1	21. ITEM 2	22. ITEM 3	23. ITEM 4	24. ITEM 5	25. ITEM 6	26. ITEM 7	27. ITEM 8	28. ITEM 9	29. ITEM 10	30. ITEM 11	31. ITEM 12	32. PERF AV	33. # OBS	34. # ADM	35. # PROT	36. # DEMO	
1.		-32	18	14	03	12	-19	-13	04	-29	36	-12	29	-14	33	00	-01	-43	16	06	50	42	21	27	48	16	-03	47	42	12	11	37	(not applicable)				1.
2.			-04	37	-24	-08	24	35	-26	31	-29	42	-23	-39	-36	-18	01	39	-16	00	-26	-18	01	-02	-05	02	03	-09	-40	-18	-02	-14					2.
3.				05	-05	54	-18	13	-18	-34	24	-26	15	02	-03	01	-01	-07	17	-18	-19	-18	-01	07	-03	07	15	04	06	25	18	02					3.
4.					-20	-05	05	-12	13	31	20	07	13	-26	-35	-34	-33	07	14	-18	05	06	-04	00	29	-20	-06	16	09	16	35	09					4.
5.						34	-37	-40	-22	16	-07	-14	39	20	-04	51	-52	-52	02	05	10	08	13	03	-26	24	15	03	-02	-04	-03	05					5.
6.							-17	-30	-29	-21	07	-15	25	11	00	17	-36	-28	17	16	-10	-13	37	28	10	38	37	18	10	44	41	29					6.
7.								03	08	02	03	17	-24	-06	-15	-22	-04	07	-01	-35	-16	-10	-34	-31	-34	-35	-46	-26	04	-01	-05	-29					7.
8.									-48	-18	-25	47	-36	-58	-08	07	50	42	-20	10	-01	-03	01	02	06	03	03	07	-11	-13	-19	-03					8.
9.										-01	14	-26	-03	41	16	-30	-16	-05	-16	-19	-12	-07	-27	-11	-10	-34	-29	-26	-05	03	-16	-22					9.
10.											-07	04	-08	30	-58	-30	-31	01	31	-12	-06	-09	02	09	-11	-03	10	-03	-14	-03	19	-02					10.
11.												-24	-09	07	08	-50	-02	-13	16	-23	18	18	-19	03	24	-36	-30	08	21	10	20	03					11.
12.													-59	-55	02	-09	-03	25	-23	17	03	-05	14	04	04	-01	09	-09	05	-14	-20	00					12.
13.														19	-22	37	-38	-24	14	-19	06	11	-17	-16	-05	13	-09	12	11	10	09	02					13.
14.															-05	-08	-36	-33	22	-22	-27	-28	-16	02	-22	-04	03	-25	-08	00	02	-17					14.
15.																01	28	-17	-15	53	35	31	46	32	35	22	15	32	22	04	05	37					15.
16.																	-02	-47	14	-01	03	05	05	-12	-33	20	07	-09	-03	-10	-23	-06					16.
17.																		33	-11	20	04	04	07	01	21	-02	06	09	-13	-23	-07	02					17.
18.																			-13	16	-11	-04	-13	-20	09	04	04	-03	-19	-02	-22	-08					18.
19.																				14	50	45	28	49	25	35	26	57	64	61	49	59					19.
20.																					50	45	83	65	62	76	60	64	30	44	37	80					20.
21.																						97	49	43	45	37	02	82	65	47	15	73					21.
22.																							42	35	37	34	-03	77	54	47	08	67					22.
23.																								77	56	74	73	72	24	53	53	85					23.
24.																									60	64	61	71	42	53	59	82					24.
25.																										43	43	66	38	33	56	72					25.
26.																											79	60	27	52	30	75					26.
27.																												33	03	38	45	58					27.
28.																													54	59	48	90					28.
29.																														53	35	61					29.
30.																															58	73					30.
31.																																63					31.
32.																																					32.
33.																																					33.
34.																																					34.
35.																																					35.
36.																																					36.

.396 needed for significance at .05 level with df = 23
.505 needed for significance at .01 level with df = 23

Note: Decimal points have been omitted

APPENDIX G

EXPERIMENTAL GROUP ANONYMOUS COMMENTS

(Due to the extensive nature of some of the comments, only selected portions are included. The spelling, syntax, and punctuation are those of the students.)

I found the first ten weeks of this course most interesting--and most time-consuming. Actually I wish I had kept account of the number of hours I spent on it! . . . I have felt very positive about the course the entire time. The feeling of setting my own time to do these things was stimulating, and I feel that I probably worked harder and learned more than if I'd been in the regular class.

Over all I think it was ok.

Overall I think this has been an effective way to teach the WISC. I do feel that it is important to have interactions (criticisms, evaluations, etc.) by ________. I got more out of ________'s criticisms on one of my demonstrations than from reading the manual over and over. . . . Overall though, I feel this has been a very educational experience and I really felt good about the self-teaching methods used.

My reactions to the first 10 weeks of this course are mixed. For one, It was nice not to have to come to class every Thursday evening but, at the same time, listening to tapes was boring because there was no feedback or opinions on different theories from others members of the class. Having to listen to tapes on

EXPERIMENTAL GROUP ANONYMOUS COMMENTS

(Due to the extensive nature of some of the comments, only selected portions are included. The spelling, syntax, and punctuation are those of the students.)

I found the first ten weeks of this course most interesting--and time-consuming. Actually I wish I had kept account of the number of hours I spent on it! . . . I have felt very positive about this course the entire time. The feeling of setting my own time to do these things was refreshing, and I feel that I probably worked harder and learned more than if I'd been in the regular class.

Over all I think it was ok.

Overall I think this has been an effective way to teach the WISC. I do feel that it is important to have interactions (criticisms, evaluations, etc.) . . . . I got more out of ______'s criticisms on one of my administrations than from reading the manual over and over. . . . Overall though, I feel this has been a very educational experience and I really felt good about the self-teaching methods used.

My reactions to the first 10 weeks of this course are mixed. At one point it was nice not to have to come to class every Thursday evening but, at the same time, listening to tapes was boring because there was no feedback or exchange of different theories from other members of the class. Having the time to . . .

my own time was not good, because I did not feel pressured to listen to them. Besides I like to see my speaker and not just hear him. Also, when listening to the tapes, I did not concentrate or give my undivided attention to them so I did not learn as much. As far as the one WISC demonstration, I would have liked to have been required to do two because I need a practice run, and then a final run. I know I could have done more demonstrations but, when there are no pressures on me, it is hard to produce.

I wish we could have complete independent study classes--I liked it.

Enjoyed the freedom to work independently from a lot of structure. However, I missed feedback and interaction with the instructor and other students.

I think the concept of a free learning set is a good idea. My only objection to it is the loss of discussion. I did hear a number of people from group 2 talking about the questions which had come up in their class, and I felt somewhat cheated. Problems which came up and were written down to be discussed when I could get in touch with _______ sometimes lost their flavor by the time of contact. . . . Everyone was cooperative, yet I did miss the opportunity to hear and exchange ideas.

I have enjoyed the freedom of independent study. However, I feel one or two opportunities to get together during the 10 weeks and rap would have been helpful.

my own time was not good, because I did not feel pressured to listen to them. Besides I like to see my teacher and not just hear him. Also, when listening to the tapes, I did not concentrate or give my undivided attention to them so I did not learn as much. As far as the live WISC demonstration, I would have liked to have been required to do two because I need a practice run, and then a final run. I know I could have done more demonstrations but, when there are no pressures on me, it is hard to produce.

I wish we could have complete independent study classes—I liked it.

Enjoyed the freedom to work independently from a lot of structure. However, I missed feedback and interaction with the instructor and other students.

I think the concept of a free learning set is a good idea. My only objection to it is the loss of discussion. I did hear a number of people from group 2 talking about the questions which had come up in their class, and I felt somewhat cheated. Problems which came up and were written down to be discussed when I could get in touch with ______ sometimes lost their fervor by the time of contact.

Everyone was cooperative, but I did miss the opportunity to hear and exchange ideas.

I have enjoyed the freedom of independent study. However, I feel one or two opportunities to get together during the 10 weeks and rap would have been helpful.

In terms of gaining information, I feel I gained as much as if I had been in the other group. I was convenient to listen to the tapes, rather than to come up here for class lectures. The tapes were easy to follow and helpful. The added benefit was that you could re-play some part if you were not sure you understood. A disadvantage would be not having the opportunity to ask questions. . . . Where I feel this group was slighted would be not having direct guidance, and more importantly, not having interaction with other students. The freedom to choose your own course of direction was nice. However, it was somewhat frustrating trying to decide if I was doing enough to be prepared for the test and the observation.

I think that the tapes were an excellent way of getting additional information mainly because of their accessability. . . . Overall, I like the way the class was set up and feel that it was a valuable learning experience. I do feel that for an unmotivated person this set up could have been a total waste. That would have been his own fault though.

I found that I missed the stimulation of the group in motivating interest in the task of learning how to administer the WISC. Therefore I did not get involved in the course until the end of the 10-week period and did not take much advantage of valuable feedback opportunities which were offered. The visible and social stimulation of "class" were missing and were missed.

In terms of gaining information, I feel I gained as much as if I had been in the other group. It was convenient to listen to the tapes, rather than to come up here for class lectures. The tapes were easy to follow and helpful. The added benefit was that you could re-play some part if you were not sure you understood. . . . A disadvantage would be not having the opportunity to ask questions. . . . A thing I feel this group was slighted would be not having direct guidance, and more importantly, not having interaction with other students. The freedom to choose your own course of direction was nice. However, it was somewhat frustrating trying to decide if I was doing enough to be prepared for the test and the observation.

I think that the tapes were an excellent way of getting assistance or information mainly because of their accessibility. . . . Overall, I like the way the class was set up and feel that it was a valuable learning experience. I do feel that for an unmotivated person this set up could have been a total waste. That would have been his own fault though.

I found that I missed the stimulation of the group in motivating interest in the task of learning how to administer the WISC-R. Therefore I did not get involved in the course until the end of the 10-week period and did not take much advantage of valuable feedback opportunities which were offered. The visible and social stimulation of "class" were missing and were missed.

CONTROL GROUP ANONYMOUS COMMENTS

(Due to the extensive nature of some of the comments, only selected portions are included. The spelling, syntax, and punctuation are those of the students.)

I feel the Differential course has been a very valuable experience for me. I have faith in testing and want to learn more. The structure of the course was clear and well done, although I feel the examination was not essential and does not prove knowledge or lack of knowledge on the WISC. I feel 15 WISC administrations were helpful, and observing also. The overall knowledge from this was exceptionally beneficial to me.

My reactions to the 10 week experience are: excitement that I'm getting closer to understanding and effectively administering a test that is highly regarded, irritation and annoyance that class time was not spent more dynamically and fruitfully, and irritation that more excitement and interest was not generated by and among students. I have also experienced some anxiety over getting all the demos, practice WISC's, observations and reflective study done well within a rather short and busy time schedule.

The ten week course, though boring and stressful at times, was invaluable. What I had expected, i.e., too much work, turned out to be a real experience. The administration of the WISC became actually enjoyable when I began to understand the real importance of each subtest.

CONTROL GROUP RANDOM COMMENTS

[Due to the extensive nature of some of the comments, only selected portions are included. The spelling, syntax, and punctuation are those of the students.]

I feel the Differential course has been a very valuable experience for me. I have faith in testing and want to learn more. The structure of the course was clear and well done, although I feel the examination was not essential and does not prove knowledge or lack of knowledge on the WISC. I feel 15 WISC administrations were helpful, and observing also. The overall knowledge from this was exceptionally beneficial to me.

My reactions to the 10 week experience are: excitement that I'm getting closer to understanding and effectively administering a test that is highly regarded, irritation and annoyance that class time was not spent more dynamically and fruitfully, and irritation that more excitement and interest was not generated by and among students. I have also experienced some anxiety over getting all the demos, practice WISC's, observations and reflective study done well within a rather short and busy time schedule.

The ten-week course, though boring and stressful at times, was invaluable. What I had expected, i.e., too much work, turned out to be a real experience. The administration of the WISC became naturally enjoyable when I began to understand the real importance of each subtest.

An advantage to giving 15 WISC's is that I got to know 15 children that I would ordinarily have never met. I got to see the siblings of my friends and get to know them. . . . I enjoyed giving the WISC most of the time. The amount of time taken was a little harder to accept. I do not like having to observe 5 WISC demonstrations but I can see some validity in the observations. I feel this has been a learning experience for me and will help me in testing in the future. I would have preferred less time on one test and more time on several tests.

There could have been more discussion of individual protocols and the inference made as well as conclusions and recommendations. Perhaps each student could have made a class presentation of a protocol. The discussions and criticisms following might have been helpful.

. . . one needs to administer the test a number of times before one can be familiar both with the test and ones actions and reactions to the test. However, I felt that I reached this level after 10 administrations. Therefore, maybe the number of required WISCs should be more flexible. The demonstrations were interesting, but I feel that 3 observations was more than sufficient for my purposes. Repetition is good while one is learning, however, after a while administration becomes tedious and nonproductive.

I thought the method of presentation was lacking. I found a change of attitude in myself, becoming more interested and aware of the test and its meanings. Was not an enjoyable experience.

An advantage to giving 10 WISCs is that I got to know 10 children that I would ordinarily have never met. I got to see the uniqueness of each and get to know them. . . . I enjoyed giving the WISC most of the time. The amount of time taken was a little harder to accept. I do not like having to observe 5 WISC demonstrations but I can see some validity in the observations. I feel this has been a learning experience for me and will help me in testing in the future. I would have preferred less time on one test and more time on several tests.

There could have been more discussion of individual protocols and the inference made as well as conclusions and recommendations. Perhaps each student could have made a class presentation of a protocol. The discussion and criticisms following might have been helpful.

. . . . one needs to administer the test a number of times before one can be familiar both with the test and ones actions and reactions to the test. However, I felt that I reached this level after 10 administrations. Therefore, maybe the number of required WISCs should be more flexible. The demonstrations were interesting, but I feel that 5 observations was more than sufficient for my purposes. Repetition is good while one is learning, however, after a while administration becomes tedious and nonproductive.

I thought the method of presentation was lacking. I found a change of attitude in myself, becoming more interested and aware of the test and its meaning. Was not an enjoyable experience.

This experience--time consuming, repetitive, frustrating, boring at times. Generally beneficial--sometimes it's hard being a guinea pig. Anyway--it was worthwhile _if_ it proved that this is or isn't a better way--so generally, OK.

I enjoyed giving WISCs and learning how to interpret and score, etc. But I believe that a course as demanding as this one should not be taken along with 3 other classes. I feel more proficient as the result of administering so many tests, but the mental anguish of "how can I write 7 papers and give 15 WISCs this weekend?" spoiled my feelings of achievement somewhat. I don't think it is necessary to give as many as 15.

I do not feel I needed to administer 15 WISCs! That was bothersome busy-work which got in the way of my learning the really important aspects of the WISC. My energy the first 10 weeks went into improving upon my administering abilities, rather than my understanding of what I was administering.

If the rest of our schedules had not been so busy I would have said that people could have reached that particular level of competency on the WISC by themselves. However, with all the other things we had to do, maybe such structure was necessary in order to help us make best use of our time. I endorse this teaching method with some reserve. I would like to know more statistics, more about testing, more about the concept of intelligence. I know enough about the WISC.

This experience--time consuming, repetitive, frustrating, boring at times. Generally beneficial--sometimes it's hard being a guinea pig. Anyway--it was worthwhile if it proves that this is or isn't a better way--so generally, OK.

. . . .

I enjoyed giving WISCs and learning how to interpret and score, etc. but I believe that a course as demanding as this one should not be taken along with 3 other classes. I feel more proficient as the result of administering so many tests, but the mental anguish of "how can I write 7 papers and give 15 WISCs this weekend ?" spoiled my feelings of achievement somewhat. I don't think it is necessary to give as many as 15.

I do not feel I needed to administer 15 WISCs ! That was bothersome busy-work which got in the way of my learning the really important aspects of the WISC. My energy the first 10 weeks went into improving upon my administering abilities, rather than my understanding of what I was administering.

If the rest of our schedules had not been so busy I would have said that people could have reached that particular level of competency on the WISC by themselves. However, with all the other things we had to do, maybe such structure was necessary in order to help us make best use of our time. I endorse this teaching method with some reserve. I would like to know more statistics, more about testing, more about the concept of intelligence. I know enough about the WISC.

I don't believe it is necessary to give any arbitrary number of tests to become proficient. I resented the demands made upon my time by the tests, observations, observed tests, etc. I have mixed feelings about the class in general--I recognize the importance of testing and the responsibility to do it accurately but I was terribly bored at times and found it hard to force myself to look at another WISC.

I thought the subject matter was dry and generally could be . . . learned by my reading at home. The class could have been held half as often as just a place to ask questions about problems, . . .

The past ten weeks in this class have been the most boring, dehumanizing, waste of time I have ever experienced in school. I do not respect myself for having remained in this class. . . . I would have dropped it the first day had it not been a requirement.

REFERENCES

REFERENCES

Allen W. H. Instructional media research: Past, present, and future. Audio Visual Communication Review, 1970, 19, 5-18.

American Personnel and Guidance Association. The responsible use of tests: A position paper of AMEG, APGA, & NCME. Measurement and Evaluation in Guidance, 1972, 5, 385-388.

American Psychological Association. Division on the Teaching of Psychology Newsletter. University of Missouri, Columbia, Mo., December, 1973.

American Psychological Association, American Educational Research Association, & National Council on Measurement in Education, Standards for educational & psychological tests and manuals. Washington, D.C.: APA, 1966.

Anastasi, A. Psychology, psychologists, & psychological testing. American Psychologist, 1967, 22, 297-306.

Anastasi, A. Psychological testing. (3rd ed.) New York: Macmillan, 1968.

Atherton, C. R. College teaching methods: A comparison of three common strategies. (Doctoral dissertation, The University of Illinois at Urbana-Champaign) Ann Arbor, Mich.: University Microfilms, 1971. No. 72-12,076.

Atherton, C. R. Lecture, discussion, and independent study instructional methods revisited. The Journal of Experimental Education, 1972, 40, 24-28.

Auffrey, J., & Robertson, M. Case history information & examiner experience as determinants of scoring validity on Wechsler intelligence tests. Proceedings of the 80th Annual Convention of the American Psychological Association, 1972, 7(Pt. 2), 553-554. (Summary)

Baker, C. D. Training counselors in the use of tests and measurement. Paper presented at the annual convention of the American Personnel and Guidance Association (session no. 214), Dallas, 1967.

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

REFERENCES

Allen W. H. Instructional media research: Past, present, and future. Visual Communication Review, 1970, 19, 5-18.

American Personnel and Guidance Association. The responsible use of tests: A position paper of AMEG, APGA, & NCME. Measurement and Evaluation in Guidance, 1972, 5, 385-388.

American Psychological Association. Division on the Teaching of Psychology. Newsletter. University of Missouri, Columbia, Mo., December, 1971.

American Psychological Association, American Educational Research Association, & National Council on Measurement in Education. Standards for educational & psychological tests and manuals. Washington, D.C.: A.P.A., 1966.

Anastasi, A. Psychology, psychologists, & psychological testing. American Psychologist, 1967, 22, 297-306.

Anastasi, A. Psychological testing. (3rd ed.). New York: Macmillan, 1968.

Atherton, C. R. College teaching methods: A comparison of three classroom strategies. (Doctoral dissertation, The University of Illinois at Urbana-Champaign) Ann Arbor, Mich.: University Microfilms, 1971. No. 72-12,076.

Atherton, C. R. Lecture, discussion, and independent study in teaching methods revisited. The Journal of Experimental Education, 1972, 40, 24-28.

Auffrey, J., & Robertson, M. Case history information & examiner experience as determinants of scoring validity on Wechsler intelligence tests. Proceedings of the 80th Annual Convention of the American Psychological Association, 1972, 7(Pt. 2), 553-554. (Summary)

Baker, C. D. Training counselors in the use of tests and measurement. Paper presented at the annual convention of the American Personnel and Guidance Association (session no. 214), Dallas, 1967.

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

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

Birney, R., & McKeachie, W. The teaching of psychology: A survey of research since 1942. Psychological Bulletin, 1955, 52, 51-68.

Born, D. G., Gledhill, S. M., & Davis, M. L. Examination performance in lecture-discussion & personalized instruction courses. Journal of Applied Behavior Analysis, 1972, 5, 33-43.

Bornstein, A. V. The effects of examiner approval and disapproval upon the performance of subjects on the performance scale of The Wechsler Intelligence Scale for Children. (Doctoral dissertation, The Florida State University) Ann Arbor, Mich.: University Microfilms, 1967. No. 68-339.

Bradley, J. V. Distribution-free statistical tests. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Briggs, L. J. Learner variables and educational media. Review of Educational Research, 1968, 38, 160-176.

Burson, G. E. A comparative study of the effects of programmed presentations of selected portions of the Stanford-Binet intelligence test on student examiners. (Doctoral dissertation, Oklahoma State University) Ann Arbor, Mich.: University Microfilms, 1966. No. 67-7202.

Burton, W. H. Basic principles in a good teaching-learning situation. In H. W. Bernard & W. C. Huckins (Eds.), Readings in educational psychology. Cleveland: World Publishing, 1967. Pp. 102-114.

Carlson, C. R. A study of the relative effectiveness of lecture and directed discussion methods of teaching tests and measurements to prospective Air Force instructors. (Doctoral dissertation, University of Minnesota) Ann Arbor, Mich.: University Microfilms, 1953. No. 6375.

Caro, P. W., Jr. The effect of class attendance and "time structured" content on achievement in general psychology. (Doctoral dissertation, The University of Tennessee) Ann Arbor, Mich.: University Microfilms, 1961. No. 61-6724.

Cieutat, V. J. Examiner differences with the Stanford-Binet IQ. Perceptual & Motor Skills, 1965, 20, 317-318.

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

Birney, R., & McKeachie, W. The teaching of psychology: A survey of research since 1942. Psychological Bulletin, 1955, 52, 51-68.

Born, D. G., Gledhill, S. M., & Davis, M. L. Examination performance in lecture-discussion & personalized instruction courses. Journal of Applied Behavior Analysis, 1972, 5, 33-43.

Bornstein, A. V. The effects of examiner approval and disapproval upon the performance of subjects on the performance scale of the Wechsler Intelligence Scale for Children. (Doctoral dissertation, The Florida State University) Ann Arbor, Mich.: University Microfilms, 1967. No. 68-339.

Bradley, J. V. Distribution-free statistical tests. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Briggs, L. J. Learner variables and educational media. Review of Educational Research, 1968, 38, 160-176.

Burton, G. E. A comparative study of the effects of programmed presentations of selected portions of the Stanford-Binet intelligence test on student examiners. (Doctoral dissertation, Oklahoma State University) Ann Arbor, Mich.: University Microfilms, 1966. No. 67-7702.

Burton, W. H. Basic principles in a good teaching-learning situation. In H. W. Bernard & W. C. Huckins (Eds.), Readings in educational psychology. Cleveland: World Publishing, 1967. Pp. 102-114.

Carlson, C. R. A study of the relative effectiveness of lecture and directed discussion methods of teaching tests and measurements to prospective Air Force instructors. (Doctoral dissertation, University of Minnesota) Ann Arbor, Mich.: University Microfilms, 1953. No. 6275.

Carr, P. W., Jr. The effect of class attendance and "time structured" content on achievement in general psychology. (Doctoral dissertation, The University of Tennessee) Ann Arbor, Mich.: University Microfilms, 1961. No. 61-5224.

Claxton, V. J. Examiner differences with the Stanford-Binet IQ. Perceptual & Motor Skills, 1965, 20, 317-318.

Cohen, A. M., Clare, R., & Trent, J. W. Teaching technology and methods. In M. W. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand, McNally, 1973. Pp. 1023-1071.

Cohen, E. Examiner differences with individual intelligence tests. Perceptual & Motor Skills, 1965, 20, 1324.

Coop, R. H. The effects of cognitive style and teaching method on categories of achievement. (Doctoral dissertation, Indiana University) Ann Arbor, Mich.: University Microfilms, 1968. No. 66-17,260.

Coop, R. H., & Brown, L. D. Effects of cognitive style and teaching method on categories of achievement. Journal of Educational Psychology, 1970, 61, 400-405.

Cooper, J. L. Learning theory and effective instruction. Journal of Higher Education, 1973, 44, 217-234.

Cooper, J. L., & Greiner, J. M. Contingency management in an introductory psychology course produces better retention. Psychological Record, 1971, 21, 391-400.

Cunningham, W. H. Achievement in a first course in psychology in relation to methods of teaching. (Doctoral dissertation, Indiana University) Ann Arbor, Mich.: University Microfilms, 1964. No. 65-388.

Dana, J. M., & Dana, R. H. Experimenter-Bias and the WAIS. Perceptual & Motor Skills, 1969, 28, 694.

Dowell, B. M. The effects of various teaching methods upon the self-concept, attitude toward psychology, anxiety level, dogmatism level, and level of achievement in general psychology of first semester Cedarville College freshmen. (Doctoral dissertation, The University of Tennessee) Ann Arbor, Mich.: University Microfilms, 1967. No. 68-3737.

Dubin, R., & Taveggia, T. C. The teaching-learning paradox: A comparative analysis of college teaching methods. Eugene, Oregon: University of Oregon Center for the Advanced Study of Educational Administration, 1968.

DuBois, P. H. A history of psychological testing. Boston: Allyn & Bacon, 1970.

Eckert, R. E., & Neale, D. C. Teachers and teaching. Review of Educational Research, 1965, 35, 304-317.

Cohen, A. M., Clore, R., & Trent, J. W. Teaching technology and methods. In M. W. Travers (Ed.), *Second handbook of research on teaching.* Chicago: Rand, McNally, 1973. Pp. 1023-1071.

Cohen, E. Examiner differences with individual intelligence tests. *Perceptual & Motor Skills,* 1965, *20,* 1324.

Coop, R. H. *The effect of cognitive style and teaching method on categories of achievement.* (Doctoral dissertation, Indiana University) Ann Arbor, Mich.: University Microfilms, 1968. No. 69-17,260.

Coop, R. H., & Brown, L. D. Effects of cognitive style and teaching method on categories of achievement. *Journal of Educational Psychology,* 1970, *61,* 400-405.

Cooper, J. L. Learning theory and effective instruction. *Journal of Higher Education,* 1973, *44,* 217-234.

Cooper, J. L., & Greiner, J. M. Contingency management in an introductory psychology course produces better retention. *Psychological Record,* 1971, *21,* 391-400.

Cunningham, W. H. *Achievement in a first course in psychology in relation to methods of teaching.* (Doctoral dissertation, Indiana University) Ann Arbor, Mich.: University Microfilms, 1964. No. 65-888.

Dana, J. M., & Dana, R. H. Experimenter-bias and the WAIS. *Perceptual & Motor Skills,* 1969, *28,* 694.

Dowell, B. M. *The effects of various teaching methods upon the self-concept, attitude toward psychology, anxiety level, dogmatism level, and level of achievement in general psychology of first semester Cedarville College freshmen.* (Doctoral dissertation, The University of Tennessee) Ann Arbor, Mich.: University Microfilms, 1967. No. 68-5752.

Dubin, R., & Taveggia, T. C. *The teaching-learning paradox: A comparative analysis of college teaching methods.* Eugene, Oregon: University of Oregon Center for the Advanced Study of Educational Administration, 1968.

DuBois, P. H. *A history of psychological testing.* Boston: Allyn & Bacon, 1970.

Eckert, R. E., & Neale, D. C. Teachers and teaching. *Review of Educational Research,* 1965, *35,* 304-317.

Educational Testing Service. Proceedings of the 1970 Invitational Conference on Testing Problems. New York, 1971. Available from EDRS (ED052264).

Edwards, A. J. Individual mental testing. Scranton, Pa.: Intext, 1971 (Pt. 1), 1972 (Pt. 2).

Edwards, A. L. Manual: Edwards Personal Preference Schedule. New York: The Psychological Corporation, 1959.

Ericksen, S. C. (Ed.) Development and experiment in college teaching. Committee on Institutional Cooperation Panel on Research & Development of Instructional Resources, Big Ten Universities & The University of Chicago, 1968. Available from EDRS (ED020133).

Exner, J. E., Jr. Variations in WISC performances as influenced by differences in pretest rapport. The Journal of General Psychology, 1966, 74, 299-306.

Eysenck, H. J. The IQ argument. New York: Library Press, 1971.

Ferinden, W. E., & Jacobson, S. Educational interpretation of the Wechsler Intelligence Scale for Children (WISC). Linden, New Jersey: Remediation Associates, 1969.

Ferster, C. B. Individualized instruction in a large introductory psychology college course. The Psychological Record, 1968, 18, 521-532.

Fitzgibbon, T. J. The use of standardized instruments with urban and minority-group pupils. New York: Harcourt Brace Jovanovich, 1973.

Glasser, A. J., & Zimmerman, I. L. Clinical interpretation of the Wechsler Intelligence Scale for Children. New York: Grune & Stratton, 1967.

Gropper, G. L., & Kress, G. C., Jr. Individualizing instruction through pacing procedures. Audio Visual Communication Review, 1965, 13, 165-182.

Guilford, J. P. The nature of human intelligence. New York: McGraw-Hill, 1967.

Hakel, M. D. Prediction of college achievement from the Edwards Personal Preference Schedule using intellectual ability as a moderator. Journal of Applied Psychology, 1966, 50, 336-340.

Hartley, J. New approaches in the teaching of psychology; An annotated bibliography. Bulletin of the British Psychological Society, 1972, 25, 291-304.

Educational Testing Service. Proceedings of the 1970 Invitational Conference on Testing Problems. New York, 1971. Available from EDRS (ED082294).

Edwards, A. J. Individual mental testing. Scranton, Pa.: Intext, 1971 (Pt. 1), 1972 (Pt. 2).

Edwards, A. L. Manual: Edwards Personal Preference Schedule. New York: The Psychological Corporation, 1959.

Ericksen, S. C. (Ed.) Development and experiment in college teaching. Committee on Institutional Cooperation Panel on Research & Development of Instructional Resources, Big Ten Universities & the University of Chicago, 1968. Available from EDRS (ED020133).

Exner, J. E., Jr. Variations in WISC performances as influenced by differences in pretest rapport. The Journal of General Psychology, 1966, 74, 299-306.

Eysenck, H. J. The IQ argument. New York: Library Press, 1971.

Farinden, W. E., & Jacobson, S. Educational interpretation of the Wechsler Intelligence Scale for Children (WISC). Linden, New Jersey: Remediation Associates, 1969.

Ferster, C. B. Individualized instruction in a large introductory psychology college course. The Psychological Record, 1968, 18, 521-532.

Fitzgibbon, T. J. The use of standardized instruments with urban and minority-group pupils. New York: Harcourt Brace Jovanovich, 1972.

Glasser, A. J., & Zimmerman, I. L. Clinical interpretation of the Wechsler Intelligence Scale for Children. New York: Grune & Stratton, 1967.

Gropper, G. L., & Kress, G. C., Jr. Individualizing instruction through pacing procedures. Audio Visual Communication Review, 1965, 13, 165-182.

Guilford, J. P. The nature of human intelligence. New York: McGraw-Hill, 1967.

Hakel, M. D. Prediction of college achievement from the Edwards Personal Preference Schedule using intellectual ability as a moderator. Journal of Applied Psychology, 1966, 50, 336-340.

Hartley, J. New approaches in the teaching of psychology. An annotated bibliography. Bulletin of the British Psychological Society, 1972, 25, 291-304.

Hatch, W. R., & Bennett, A. Effectiveness in teaching: New dimensions in higher education, No. 2, U.S. Office of Education, 1960, p. 10. Cited by J. W. Brown & J. W. Thornton: College teaching: Prospectives and guidelines. New York: McGraw-Hill, 1963. Pp. 134-135.

Hewitt, P., & Massey, J. O. Clinical clues from the Wechsler Intelligence Scale for Children with special sections on testing Black and Spanish-speaking children. Palo Alto, Calif.: Consulting Psychologists Press, 1969.

Himmel, C. E. A critical review and analysis of self-directed learning methods utilized in the teaching of undergraduate psychology courses. (Doctoral dissertation, University of Illinois at Urbana-Champaign) Ann Arbor, Mich.: University Microfilms, 1970. No. 70-20,978.

Himmel, C. E. College learning with and without formal classroom instruction --a comparison. Psychology in the Schools, 1972, 9, 272-277.

Holmen, M. G., & Docter, R. Educational and psychological testing: A study of the industry and its practices. New York: Russell Sage Foundation, 1972.

Jackson, C. W., Jr., & Wohl, J. A survey of Rorschach teaching in the university. Journal of Projective Techniques and Personality Assessment, 1966, 30, 115-134.

Jenkins, J. R. A study of small group vs. individual achievement via the audio-tutorial instructional system. (Doctoral dissertation, Purdue University) Ann Arbor, Mich.: University Microfilms, 1972. No. 72-30,911.

Karmel, L. J. Preparing counselors to use tests. Counselor Education and Supervision, 1968, 7, 319-320.

Keller, F. S. "Good-bye, teacher . . . ." Journal of Applied Behavior Analysis, 1968, 1, 79-89.

Kelly, H. H., & Thibaut, J. W. Group problem solving. In G. Lindzey & E. Arnson (Eds.), The handbook of social psychology. (2nd ed.) Vol. 4. Reading, Mass.: Addison-Wesley, 1969. Pp. 1-102.

Kulik, J. A., Kulik, C., & Carmichael, K. The Keller plan in science teaching. Science, 1974, 183, 379-383.

Linden, K. W., & Linden, J. D. Modern mental measurement: A historical perspectus. Guidance Monograph Series, Series III: Testing. Boston: Houghton Mifflin, 1968.

Lutey, C. L. Individual intelligence testing: A manual. Greeley, Colo.: University of Northern Colorado, 1966.

Maccoby, E. E. Differential cognitive abilities. Cooperative Research Project No. 1040: The Cooperative Research Program of the Office of Education. Washington, D.C.: U.S. Dept. of HEW, 1962.

MacLeod, R. B. The teaching of psychology and the psychology we teach. American Psychologist, 1965, 20, 344-352.

MacLeod, R. B. Teaching of psychology. In The quest for relevance: Effective college teaching. Vol. III. The social sciences. Washington, D.C.: Office of Education, Bureau of Research, 1969. Contract No. OEC-4-7-062999-3492. Pp. 43-52. Available from EDRS (ED027857).

MacLeod, R. B. The teaching of psychology. American Psychologist, 1971, 26, 245-249.

Marr, J. N., Plath, D. W., Wakeley, J. H., & Wilkins, D. M. The contribution of the lecture to college teaching. Journal of Educational Psychology, 1960, 51, 277-284.

Masling, J. The effects of warm and cold interaction on the administration and scoring of an intelligence test. Journal of Consulting Psychology, 1959, 23, 336-341.

Matarazzo, J. D. Wechsler's measurement and appraisal of adult intelligence. (5th ed.) Baltimore: Waverly Press, 1972.

McClelland, D. C. Testing for competence rather than for "intelligence." American Psychologist, 1973, 28, 1-13.

McGrew, J. M., Marcia, J. E., & Wright, C. K. Branching program, text, and lecture: A comparative investigation of instructional media. Journal of Applied Psychology, 1966, 50, 505-508.

McKeachie, W. J. Students, groups, and teaching methods. The American Psychologist, 1958, 13, 580-584.

McKeachie, W. J. Research on teaching at the college and university level. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand, McNally, 1963. Pp. 1118-1172.

McKeachie, W. J. Research in teaching: The gap between theory and practice. In C. B. T. Lee (Ed.), Improving college instruction. Washington, D.C.: American Council on Education, 1967. Pp. 211-239.

Lutey, C. L. Individual intelligence testing: A manual. Greeley, Colo.: University of Northern Colorado, 1966.

Maccoby, E. E. Differential cognitive abilities. Cooperative Research Report No. 1040. The Cooperative Research Program of the Office of Education. Washington, D.C.: U.S. Dept. of HEW, 1962.

McLeod, R. B. The teaching of psychology and the psychology we teach. American Psychologist, 1965, 20, 344-352.

Macleod, R. B. Teaching of psychology. In The quest for relevance: Effective college teaching. Vol. III. The social sciences. Washington, D.C.: Office of Education, Bureau of Research, 1969. Contract No. OEC-4-7-062999-3082. Pp. 48-62. Available from EDRS (ED027837).

Macleod, R. B. The teaching of psychology. American Psychologist, 1971, 26, 245-249.

Marr, J. N., Plath, D. W., Wakeley, J. H., & Wilkins, D. M. The contribution of the lecture to college teaching. Journal of Educational Psychology, 1960, 51, 277-284.

Masling, J. The effects of warm and cold interaction on the administration and scoring of an intelligence test. Journal of Consulting Psychology, 1959, 23, 336-341.

Matarazzo, J. D. Wechsler's measurement and appraisal of adult intelligence. (5th ed.). Baltimore: Waverly Press, 1972.

McClelland, D. C. Testing for competence rather than for "intelligence." American Psychologist, 1973, 28, 1-13.

McGrew, J. M., Marcia, J. E., & Wright, C. K. Branching program, text, and lecture: A comparative investigation of instructional media. Journal of Applied Psychology, 1966, 50, 505-508.

McKeachie, W. J. Students, groups, and teaching methods. The American Psychologist, 1958, 13, 580-584.

McKeachie, W. J. Research on teaching at the college and university level. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 1118-1172.

McKeachie, W. J. Research in teaching: The gap between theory and practice. In C. B. T. Lee (Ed.), Improving college instruction. Washington, D.C.: American Council on Education, 1967. Pp. 211-239.

McKeachie, W. J. Psychology at age 75: The psychology teacher comes into his own. American Psychologist, 1968, 23, 551-557.

McKeachie, W. J., Lin, Y. G., Forrin, B., & Teevan, R. Individualized teaching in elementary psychology. Journal of Educational Psychology, 1960, 51, 285-291.

Menne, J. W., Hannum, T. E., Klingensmith, J. E., & Nord, D. Use of taped lectures to replace class attendance. Audio Visual Communication Review, 1969, 17, 42-46.

Michael, W. B. A short evaluation of the research reviewed in educational and psychological testing. Review of Educational Research, 1965, 25, 92-99.

Miller, C. K., & Chansky, N. M. Psychologists' scoring of WISC protocols. Psychology in the Schools, 1972, 9, 144-152.

Miller, C. K., Chansky, N. M., & Gredler, G. R. Rater agreement on WISC protocols. Psychology in the Schools, 1970, 7, 190-193.

Mills, R. E. An evaluation of techniques for the teaching of word recognition. (Doctoral dissertation, The University of Florida) Ann Arbor, Mich.: University Microfilms, 1955. No. 14,321.

Morris, C. J., & Kimbrill, G. M. Performance and attitudinal effects of the Keller method in an introductory psychology course. The Psychological Record, 1972, 22, 523-530.

Nachman, M., & Opochinsky, S. The effects of different teaching methods: A methodological study. Journal of Educational Psychology, 1958, 49, 245-249.

Partridge, P. V. An experiment in the evaluation of instruction at the junior college. Unpublished research project, University of California, Los Angeles, 1968. Available from EDRS (ED037211).

Pascal, C. E. Offering course options: Personality, option preference and course outcomes, NSPI Journal, 1971, 10, 8-9, 17, 18, 21.

Pedrini, D. T., & Pedrini, L. N. The Pedrini supplementary aid to the administration of the Stanford-Binet Intelligence Scale (Form L-M): A handbook. Los Angeles: Western Psychological Services, 1970.

Popham, W. J. Tape recorded lectures in the college classroom. Audio Visual Communication Review, 1961, 9, 109-118.

McKeachie, W. J. Psychology at age 75: The psychology teacher comes into his own. American Psychologist, 1968, 23, 551-557.

McKeachie, W. J., Lin, Y. G., Forrin, B., & Teevan, R. Individualized teaching in elementary psychology. Journal of Educational Psychology, 1960, 51, 285-291.

Menne, J. W., Hannum, T. E., Klingensmith, J. E., & Nord, D. Use of taped lectures to replace class attendance. Audio Visual Communication Review, 1969, 17, 42-46.

Michael, W. B. A short evaluation of the research reviewed in educational and psychological testing. Review of Educational Research, 1965, 35, [illegible]

Miller, C. K., & Chansky, N. M. Psychologists' scoring of WISC protocols. Psychology in the Schools, 1972, 9, 144-152.

Miller, C. K., Chansky, N. M., & Gredler, G. R. Rater agreement on WISC protocols. Psychology in the Schools, 1970, 7, 190-193.

Mills, R. E. An evaluation of techniques for the teaching of word recognition (Doctoral dissertation, The University of Florida). Ann Arbor, Mich.: University Microfilms, 1955. No. 14,321.

Morris, C. J., & Kimbrell, G. M. Performance and attitudinal effects of the Keller method in an introductory psychology course. The Psychological Record, 1972, 22, 523-530.

Nachman, M., & Opochinsky, S. The effects of different teaching methods: A methodological study. Journal of Educational Psychology, 1958, 49, 245-249.

Partridge, P. V. An experiment in the evaluation of instruction at the junior college. Unpublished research project, University of California, Los Angeles, 1968. Available from ERIC (ED039211)

Pascal, C. E. Offering course options: Personality, option preferences, and course outcomes, NSPI Journal, 1971, 10, 8-9, 12, 18, 23.

Pedrini, D. T., & Pedrini, L. N. The Pedrini supplementary aid to the administration of the Stanford-Binet Intelligence Scale (Form L-M): A handbook. Los Angeles: Western Psychological Services, 1970.

Popham, W. J. Tape recorded lectures in the college classroom. Audio Visual Communication Review, 1961, 9, 109-118.

Popham, W. J. Tape recorded lectures in the college classroom--II. Audio Visual Communication Review, 1962, 10, 94-101.

Reilley, R. R. Student reactions to use of video taped WISC instruction. Counselor Education and Supervision, 1969, 8, 233-235.

Robb, G. P., Bernardoni, L. C., & Johnson, R. W. Assessment of individual mental ability. Scranton, Pa.: Intext, 1972.

Robson, C. J. An experimental study of the efficacy of two instructional methods in introductory college psychology. (Doctoral dissertation, University of Minnesota) Ann Arbor, Mich.: University Microfilms, 1961. No. 61-5864.

Sattler, J. M. Effects of cues and examiner influence on two Wechsler subtests. Journal of Consulting and Clinical Psychology, 1969, 33, 716-721.

Sattler, J. M., Hillix, W. A., & Neher, L. A. Halo effect in examiner scoring of intelligence test responses. Journal of Consulting and Clinical Psychology, 1970, 34, 172-176.

Sattler, J. M., & Martin, S. Anxious and nonanxious examiner roles on two WISC subtests. Psychology in the Schools, 1971, 8, 347-349.

Sattler, J. M., & Theye, F. Procedural, situational, and interpersonal variables in individual intelligence testing. Psychological Bulletin, 1967, 68, 347-360.

Schwartz, M. L. The scoring of WAIS Comprehension responses by experienced and inexperienced judges. Journal of Clinical Psychology, 1966, 22, 425-427.

Schwartz, R. H., & Flanigan, P. J. Evaluation of examiner bias in intelligence testing. American Journal of Mental Deficiency, 1971, 76, 262-265.

Shemberg, K., & Keeley, S. Psychodiagnostic training in the academic setting: Past and present. Journal of Consulting and Clinical Psychology, 1970, 34, 205-211.

Siegal, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Smith, H. W., & May, T. Individual differences among inexperienced psychological examiners. Psychological Reports, 1967, 20, 759-762.

Popham, W. J. Tape recorded lectures in the college classroom—II. Audio-Visual Communication Review, 1962, 10, 94-101.

Reilley, R. R. Student reactions to use of video tape [illegible]. Counselor Education and Supervision, 1969, 8, 222-225.

Robb, G. P., Bernardoni, L. C., & Johnson, R. W. Assessment of individual mental ability. Scranton, Pa.: Intext, 1972.

Robson, C. J. An experimental study of the efficacy of [illegible] in introductory college psychology. (Doctoral dissertation, University of Minnesota) Ann Arbor, Mich.: University Microfilms, [illegible] 5564.

Sattler, J. M. Effects of cues and examiner influence on two Wechsler subtests. Journal of Consulting and Clinical Psychology, 1969, 33, 716-721.

Sattler, J. M., Hillix, W. A., & Neher, L. A. Halo effect in examiner scoring of intelligence test responses. Journal of Consulting and Clinical Psychology, 1970, 34, 172-176.

Sattler, J. M., & Martin, S. Anxious and nonanxious examiner roles on two WISC subtests. Psychology in the Schools, 1971, 8, 347-349.

Sattler, J. M., & Theye, F. Procedural, situational, and interpersonal variables in individual intelligence testing. Psychological Bulletin, 1967, 68, 347-360.

Schwartz, M. L. The scoring of WAIS Comprehension responses by experienced and inexperienced judges. Journal of Clinical Psychology, 1966, 22, 425-427.

Schwartz, R. H., & Flanigan, P. J. Evaluation of examiner bias in intelligence testing. American Journal of Mental Deficiency, 1971, 76, 262-265.

Shemberg, K., & Keeley, S. Psychodiagnostic training in the academic setting: Past and present. Journal of Consulting and Clinical Psychology, 1970, 34, 205-211.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Smith, H. W., & May, T. Individual differences among inexperienced psychological examiners. Psychological Reports, 1967, 20, 759-762.

Stephens, J. M. The process of schooling: A psychological examination. New York: Holt Rinehart & Winston, 1967.

Tallmadge, G. K., & Shearer, J. W. Relationships among learning styles, instructional methods, and the nature of learning experiences. Journal of Educational Psychology, 1969, 60, 222-230.

Thomas, A., Hertzig, M. E., Dryman, I., & Fernandez, P. Examiner effect in IQ testing of Puerto Rican working-class children. American Journal of Orthopsychiatry, 1971, 41, 809-821.

Travers, R. M. W. Educational psychology. In R. L. Ebel (Ed.), Encyclopedia of educational research. (4th ed.) New York: Macmillan, 1969. Pp. 413-420.

Tuckman, B. W. Conducting educational research. New York: Harcourt Brace Jovanovich, 1972.

Warren, S. A., & Brown, W. G., Jr. Examiner scoring errors on individual intelligence tests. Psychology in the Schools, 1973, 10, 118-122.

Weinick, G. D. The comparative effectiveness of two teaching methods in attaining specific course objectives: an evaluation of two different methods of teaching a course in psychology with respect to the students' acquisition of course content and changes in personal adjustment. (Doctoral dissertation, New York University) Ann Arbor, Mich.: University Microfilms, 1960. No. 61-351.

Withers, D. R., & Kent, G. W. Teaching without lecturing: Evidence in the case for individualized instruction. The Psychological Record, 1972, 22, 169-175.

Witmer, J. M., Bornstein, A. V., & Dunham, R. M. The effect of verbal approval and disapproval upon the performance of third and fourth grade children on four subtests of the Wechsler Intelligence Scale for Children. Journal of School Psychology, 1971, 9, 347-356.

Womer, F. B., & Wahi, N. K. Test use. In R. L. Ebel (Ed.), Encyclopedia of educational research. (4th ed.) New York: Macmillan, 1969. Pp. 1461-1469

Zimmerman, I. L., & Woo-Sam, J. Research with the Wechsler Intelligence Scale for Children: 1960-1970. Psychology in the Schools, 1972, 9, 232-271.

Stephens, J. M. The process of schooling: A psychological examination. New York: Holt Rinehart & Winston, 1967.

Tallmadge, G. K., & Shearer, J. W. Relationship among learning styles, instructional methods, and the nature of learning experiences. Journal of Educational Psychology, 1969, 60, 222-230.

Thomas, A., Hertzig, M. E., Dryman, I., & Fernandez, P. Examiner effect in IQ testing of Puerto Rican working-class children. American Journal of Orthopsychiatry, 1971, 41, 809-821.

Travers, R. M. W. Educational psychology. In R. L. Ebel (Ed.), Encyclopedia of educational research. (4th ed.) New York: Macmillan, 1969. Pp. 413-420.

Tuckman, B. W. Conducting educational research. New York: Harcourt Brace Jovanovich, 1972.

Warren, S. A., & Brown, W. G., Jr. Examiner scoring errors on individual intelligence tests. Psychology in the Schools, 1973, 10, 118-122.

Weinick, G. D. The comparative effectiveness of two teaching methods in attaining specific course objectives: an evaluation of two different methods of teaching a course in psychology with respect to the student's acquisition of course content and changes in personal adjustment. (Doctoral dissertation, New York University) Ann Arbor, Mich.: University Microfilms, 1960. No. 61-351.

Withers, D. R., & Kent, G. W. Teaching without lecturing: Evidence in the case for individualized instruction. The Psychological Record, 1972, 22, 169-175.

Witmer, J. M., Bornstein, A. V., & Dunham, R. M. The effects of verbal approval and disapproval upon the performance of third and fourth grade children on four subtests of the Wechsler Intelligence Scale for Children. Journal of School Psychology, 1971, 9, 347-356.

Womer, F. B., & Wahl, N. K. Test use. In R. L. Ebel (Ed.), Encyclopedia of educational research. (4th ed.) New York: Macmillan, 1969. Pp. 1461-1469.

Zimmerman, I. L., & Woo-Sam, J. Research with the Wechsler Intelligence Scale for Children: 1960-1970. Psychology in the Schools, 1972, 9, 232-271.

Zimmerman, I. L., & Woo-Sam, J. M. Clinical interpretation of the Wechsler Adult Intelligence Scale. New York: Grune & Stratton, 1973.

Zimmerman, I. L., & Woo-Sam, J. M. Clinical interpretation of the Wechsler Adult Intelligence Scale. New York: Grune & Stratton, 1973.

CURRICULUM VITAE

Kathleen Yost Ritter was born April 9, 1942, in Nampa, Idaho. She attended public schools in Caldwell, Idaho, and graduated from Caldwell High School in 1960. She attended Clarke College, Dubuque, Iowa, in 1960-61, and transferred to the University of Utah in the Fall of 1961. During the Summer of 1962, she attended the College of Idaho in Caldwell. She graduated with honors from the University of Utah in 1964 with a degree in social studies. From 1964 to 1966, she taught social studies at St. Agnes Regional High School in Springfield, Missouri.

In 1967, she received her Masters degree in Guidance from Kansas State Teachers College in Emporia. From 1968 until 1970, she held a graduate assistantship in the Department of Counselor Education at Kansas State Teachers College. During the summers of 1968, 1969, and 1970, she served as a lecturer in that Department. A Specialist in Education degree in Guidance was awarded from Kansas State Teachers College in 1970.

From 1970 to 1972, she was a lecturer at Kansas State Teachers College. During the 1970-71 academic year, she taught prospective secondary teachers in the Department of Curriculum and Instruction. In 1971-72, she taught in the Department of Research and Laboratory Experiences and in the Department of Counselor Education.

In 1972, she began work toward her Doctoral degree at the University of New Mexico. In 1972-73, she held a special assistantship in the College of Edu-

CURRICULUM VITAE

Kathleen Yost Ritter was born April 9, 1942, in Pocatello, Idaho. She attended public schools in Caldwell, Idaho, and graduated from Caldwell High School in 1960. She attended Clarke College, Dubuque, Iowa, in 1960-61, and transferred to the University of Utah in the Fall of 1961. During the summer of 1962, she attended the College of Idaho in Caldwell. She graduated with honors from the University of Utah in 1964 with a degree in social studies. From 1964 to 1966, she taught social studies at St. Agnes Regional High School in Springfield, Missouri.

In 1967, she received her Masters degree in Guidance from Kansas State Teachers College in Emporia. From 1968 until 1970, she held a graduate assistantship in the Department of Counselor Education at Kansas State Teachers College. During the summers of 1968, 1969, and 1970, she served as a lecturer in that Department. A Specialist in Education degree in Guidance was awarded from Kansas State Teachers College in 1970.

From 1970 to 1972, she was a lecturer at Kansas State Teachers College. During the 1970-71 academic year, she taught prospective secondary teachers in the Department of Curriculum and Instruction. In 1971-72, she taught in the Department of Research and Laboratory Experiences and in the Department of Counselor Education.

In 1972, she began work toward her Doctoral degree at the University of New Mexico. In 1972-73, she held a special assistantship in the College of Law.

cation. In 1973-74, she held a graduate assistantship in the Department of Guidance and Counseling and supervised counseling practicum students.

cation. In 1973-74, she held a graduate assistantship in the Department of
Guidance and Counseling and supervised counseling practicum students.