

THE UNIVERSITY OF CALGARY

A COMPARATIVE STUDY OF METHODS OF TEACHING FAMILY ASSESSMENT

BY

MAUREEN D. LEAHEY

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

DEPARTMENT OF MEDICAL SCIENCE

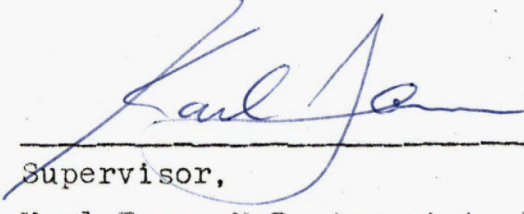
CALGARY, ALBERTA

DECEMBER, 1977

© MAUREEN D. LEAHEY 1977

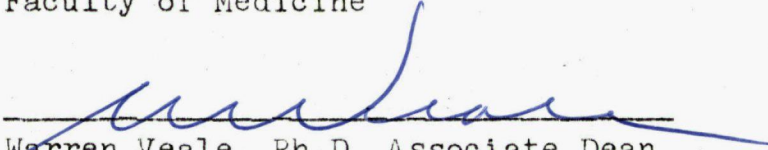
THE UNIVERSITY OF CALGARY
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for Acceptance, a thesis entitled "A Comparative Study of Methods of Teaching Family Assessment" submitted by Maureen D. Leahey in partial fulfillment of the requirements for the degree of Master of Science.

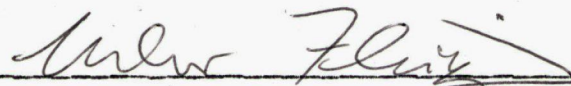


Supervisor,

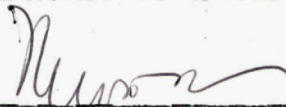
Karl Tomm, M.D. Associate Professor,
Division of Psychiatry
Faculty of Medicine



Warren Veale, Ph.D. Associate Dean
(Research) Faculty of Medicine,
Professor, Medical Physiology



Usher Fleising, Ph.D. Assistant Professor,
Department of Anthropology



M. Watanabe, M.D., Ph.D. Associate Dean
(Education), Faculty of Medicine,
Professor, Internal Medicine



M. Chorny, Ed.D., Professor,
Department of Education Curriculum
and Instruction

12 Dec. 1977

Abstract

The purpose of this study was to determine the effects of three methods of teaching family assessment to medical students. The methods were:

1. Lecture/Demonstration
2. Small Group Discussion/Demonstration
3. Small Group Discussion/Student Presentation

The basic differences between the methods was in terms of two factors:

1. Degree of Student Interaction: In Method 1, the preceptor presented didactic material and did not encourage student to student discussion. Methods 2 and 3 preceptors stimulated discussion between students.
2. Use of Videotape: Methods 1 and 2 used faculty prepared edited videotapes of family assessment by a family psychiatrist. In Method 3, student pairs assessed a family and presented the videotaped interview to their small group for discussion.

A pretest-posttest research design with random assignment of students was used. The subjects were 58 first year medical students at the University of Calgary. Method 1 included 20 subjects; Method 2, 21; and Method 3, 17. A family practitioner and a social scientist/family therapist were paired as preceptors by an independent judge and were randomly assigned to the teaching methods. A senior experienced family therapist was selected and assigned to teach Method 1. The first six two-hour teaching sessions of the

unit, Family in Health and Illness, were included in the study.

Instruments were developed to measure achievement in family assessment. Attempts were made to establish parallelism of pre and posttest forms. Equivalence was established for Part A, the factual knowledge component, but not for Part B, the application component. Interrater reliability coefficients were estimated for the two raters.. These ranged from .76 to .99. Intrarater reliability coefficients were estimated for one investigator. These ranged from .84 to .97. Two-way analysis of variance with repeated measures was used to analyze the data. A .05 level of significance was adopted.

Test results indicated no significant difference between methods in student achievement gain from pre to posttest. The posttest was significantly higher than the pretest at the .00 level of probability. Questionnaire results indicated Method 3 students were significantly less satisfied than other students with their teaching method and spent significantly more time working outside of class than Method 2 students did.

The influence on achievement of background variables, sex, marital status, and previous social service work experience was examined. No significant differences in achievement were noted although women were higher than men on pre- and post-test Part A and Total scores.

Acknowledgements

I am greatly indebted to many people who made the completion of this thesis possible.

I would like to thank especially Dr. Karl Tomm, my Committee Chairman, for introducing me to Family Therapy. For his consistent and conscientious thoroughness in clarifying my thinking and for his assistance in making my Master's program a possibility, I am appreciative. I am grateful to the members of my committee, Dr. Warren Veale and Dr. Usher Fleising, for their support and guidance.

I want also to thank Peter Harasym of the Department of Educational Planning and Assessment for his assistance with the statistical analyses and for his encouragement.

I am indebted to Dr. Mike Tarrant, the Unit Manager, who enthusiastically supported the study; to the administrators who endorsed it; to the families who kindly consented to being videotaped; to the former Block Placement students who participated in the parallelism study; and to the preceptors and students of the Class of 1979, for their cooperation, interest, and time.

I would also like to thank my friends, the Family Therapy Program staff, and Douglas who listened to me, humored me, and gave me empathic support.

TABLE OF CONTENTS

	<u>Page</u>
Abstract.....	iii
Acknowledgements.....	v
List of Tables.....	viii
List of Figures.....	x
Chapter	
1. Introduction and Review of the Literature	1
Introduction	
Statement of the Problem	
Review of the Literature	
Literature Related to the Teaching of	
Family Assessment	
Literature Related to Teaching Methods	
2. Methodology.....	26
Introduction	
Design	
Selection and Assignment of Subjects	
Selection and Assignment of Preceptors	
3. Instruments and Data Analysis.....	38
Introduction	
Tests	
Questionnaires	
Analysis of Data	
4. Results and Discussion.....	64
Introduction	
Hypotheses	
Results	
Discussion	
Implications	
Limitations	

	<u>Page</u>
5. Summary and Conclusions.....	105
The Problem	
Summary of Methodology	
Findings	
Other Significant Findings	
Conclusions	
Recommendations	
Reference Notes and References.....	117
Appendices.....	
1. Pre and Posttests.....	124
2. Student Questionnaires.....	147
3. Test Data Collection.....	162
4. Raw Data.....	166
5. Distribution of Achievement Scores.....	177
6. Effect of Attendance on Achievement.....	180

LIST OF TABLES

Table	Page
1. Similarities and Differences Between Teaching Methods.....	29
2. Characteristics of Students in Each Teaching Method.....	32
3. Characteristics of Preceptors in Each Teaching Method.....	34
4. Total Number of Test Questions Classified According to Bloom's Taxonomy.....	44
5. Characteristics of Parallelism Study Groups.....	46
6. Probability of No Significant Difference Between Tests, Test-Group Interaction and Groups.....	50
7. Pre and Posttest Variances, Pearson Product-Moment Correlations and t Levels.....	52
8. Estimates of Pretest Interrater Reliability.....	55
9. Estimates of Posttest Interrater Reliability.....	56
10. Estimates of Pretest Intrarater Reliability.....	58
11. Estimates of Posttest Intrarater Reliability.....	59
12. Probability of No Significant Difference Between Tests Part A, Test-Method Interaction, and Methods.....	67
13. Probability of No Significant Difference Between Tests Part B, Test-Method Interaction, and Methods.....	70
14. Probability of No Significant Difference Between Total Tests, Test-Method Interaction, and Methods.....	73

15.	Means for Students' Ratings of Satisfaction with Aspects of Teaching Methods.....	75
16.	Mean Study Hours for Students in Methods 1, 2 and 3.....	77
17.	Method 3 Actual Student Time Expenditure in Video Preparation.....	78
18.	Student Time Expenditure in Video Preparation Outside of Class Time Allotment.....	79
19.	Mean Hours of Student Time Expenditure Outside of Class.....	80
20.	Pre and Posttest Achievement Means for Males and Females.....	81
21.	Probability of No Significant Difference Between Males and Females on Pre and Posttests.....	82
22.	Probability of No Significant Difference Between Tests and Test-Group Interaction.....	83
23.	Pre and Posttest Achievement Means for Married and Unmarried Students.....	84
24.	Probability of No Significant Difference Between Tests, Test-Group Interaction, and Married and Unmarried Groups.....	85
25.	Pre and Posttest Achievement Means for Students Who Have Worked and for Students Who Have Not Worked in the Social Service Field.....	86
26.	Probability of No Significant Difference Between Tests, Test-Group Interaction and Work Experience Groups.....	87

LIST OF FIGURES

Figure	<u>Page</u>
1. Research Design.....	27
2. Part A Pre and Posttest Means.....	48
3. Part B Pre and Posttest Means.....	48
4. Total Pre and Posttest Means.....	49
5. Part A Achievement Means and Standard Deviation for Methods 1, 2, and 3.....	66
6. Part B Achievement Means and Standard Deviation for Methods 1, 2, and 3.....	69
7. Total Achievement Means and Standard Deviation for Methods 1, 2 and 3.....	72

Chapter 1

Introduction and Review of the Literature

Introduction

One of the functions of the physician is to recognize the influence of the family in precipitating, aggravating, and maintaining emotional problems in patients (Tomm, 1973; Weston, 1972). It is therefore important that medical students be educated in the knowledge and skills that will enable them to assess family interaction and its influence on emotional and somatic complaints.

There has been a recent trend in undergraduate medical education to teach students to assess such emotional problems and in some schools to assess family interaction as it affects illness. This was an aim of the University of Calgary in establishing the Continuity Course, in general, and the unit, The Family in Health and Illness, in particular (Continuity Course Objectives, Note 1).

In order to achieve this goal, two questions should be considered:

1. How do medical educators teach students a family assessment model?

2. How do educators evaluate the effectiveness of this teaching of medical students?

The first question is difficult to answer from a review of the literature. Little has been written regarding the specific content or method of educating physicians, medical students, or even family therapists in family assessment. Theoretical models appear to be in an early developmental stage and are frequently insufficiently explicit. Teaching methods are vaguely described and seldom evaluated.

At the University of Calgary the first question has been approached in several ways. Various methods of instruction have been employed in the past four years to help students develop the perceptual and conceptual skills involved in family assessment. These methods have included lectures, small group discussion, role playing, and videotape demonstrations of family sessions by skilled interviewers. Students have assessed families and presented a written evaluation or videotaped interview.

The second question, regarding evaluation of training programs, is seldom discussed in the literature. Very few quasi-experimental studies of teaching family assessment to medical students have been documented. At the University of Calgary, no formal evaluation of the family assessment teaching program had been carried

out. Both faculty and students, however, expressed varying opinions with respect to the relative efficacy of each of the teaching methods employed. Several of these methods, for example, small group discussion and use of videotaped interviews, are expensive in terms of faculty time involvement, student time expenditure, and audio-visual resource use.

The major objective of this study, therefore, was to delineate three methods of teaching a systems model of family assessment to first year medical students and to evaluate the effectiveness of using these methods.

Statement of the Problem

The primary purpose of this study was to determine the effects of teaching family assessment by comparing achievement scores of students in a lecture method, students in a small group discussion method, and students in a small group discussion/student project method. An attempt was made to answer the following question:

Which method of teaching family assessment to first year medical students effects the greatest achievement score?

In evaluating the effectiveness of teaching methods, one should consider not only student achievement but also student opinion of the method and the time expenditure. Thus, the second purpose of this study was to attempt to answer the following questions:

1. Which method of teaching family assessment evokes the greatest student satisfaction with the method?
2. Which method of teaching family assessment involves the greatest student time expenditure?

The third purpose of this study was to examine the relationship between student background variables and achievement scores. The background variables were: sex, marital status, and previous social service work experience. Thus, some of the questions which this study attempted to answer were:

Is there a significant difference between males' and females' achievement scores?

Is there a significant difference between married and unmarried students' achievement scores?

Are achievement scores of those who have worked in the social service field significantly higher than those who have not worked in this field?

Review of the Literature

This section presents a review of the literature dealing with evaluation of teaching methods. More specifically, the literature is reviewed under two major headings:

1. Literature Related to the Teaching of Family Assessment

2. Literature Related to Teaching Methods

Literature Related to the Teaching of Family Assessment

Little has been written regarding the specific content or method of educating therapists, physicians, or medical students in family assessment. What is written is frequently insufficiently explicit with respect to the theoretical model or the teaching method employed. However, there are several articles which describe the use of videotape in training methods. Thus, the literature relating to the teaching of family assessment will be reviewed under three sub-headings:

- i. Description of Training Programs
- ii. Use of Videotape in Psychotherapy Training
- iii. Evaluation of Training Programs

Description of Training Programs. Little differentiation has been made in the literature between the assessment and treatment models used in training professionals to work with families. Carter, Bandler, and Bakst (1953) described an early experiment in introducing medical students to a patient's family members. No assessment model, however, was indicated. Kark (1959) reported a project at the University of Natal in which medical students attempted to diagnose the state of health of a family. Epstein and Levin (1973) outlined the McMaster approach to teaching

family assessment in the Faculty of Medicine while Ferber and Mendelsohn (1969) outlined the approach employed at the Albert Einstein College of Medicine. Bodin (1969) reviewed the training programs for medical personnel at Temple and North Carolina. Beal (1976) compared 15 programs of teaching family assessment, a few of which included medical students. He concluded that the quality and design of training at most centers suffered from an insufficiently designed conceptual framework.

Use of Videotape in Psychotherapy Training. While the documentation of family assessment training programs is just beginning, the use of videotapes in teaching psychotherapy is already underway.

Videotapes are being increasingly used in teaching for a variety of purposes (Cassie, Collins, & Daggett, 1977; Forgotson & Sweeney, 1977). One purpose is to provide students with an opportunity to observe a wide range of psychiatric syndromes to which they might otherwise not be exposed (Miller & Tupin, 1972). Messner and Schmidt (1974) concur on the usefulness of this exposure to teach students the emotional aspects of family practice. Family therapists also report on the merit of videotape use to give students an orientation to family interaction (Flomenhaft & Carter, 1974).

In addition to providing students with an opportunity for increased exposure to psychiatric care, videotapes

allow them to increase their perceptual skills by observing specific interviewing techniques (Bodin, 1969, b). Randels, Kilpatrick, McCurdy, and Saunders (1976) contrast the usefulness of live interviews with videotaped interviews for this purpose. They found that a live diagnostic interview witnessed by students and discussed later is valuable, but more so for overall observations than for obtaining specific information about how to interview.

This need for detailed observation is particularly true in the area of family assessment. Bodin (1969 b) reports the advantage of having the opportunity to split channels on videotape. That is, students may increase their observational skills by paying attention either to the verbal or non-verbal channels of communication. This attribute of the videotape medium to allow splitting of channels as well as motion, is what distinguishes its usefulness from other media for example, audio tape, in the clinical situation (Levie & Dickie, 1973).

In addition to using videotape to increase students' awareness of interaction and to provide them with an opportunity to enhance their observational skills, many family therapists use tape to provide students and staff with feedback on their own interviewing performance (Beal, 1976; Bodin, 1969 b; Haley, 1975; Perlmutter,

Loeb, Gumpert, O'Hara & Higbie, 1967). Alger (1976) comments on the benefit of this feedback for the therapist and the family. He states "that the shift of role position which occurs when therapist and family step back from the usual hierarchial positions to that of cooperative researchers in a common task, marks another significant move in making therapy a more truly human and mutual adventure" (p.547).

Sanborn, Pyke, and Sanborn (1975) support the value of videotape for both the family and the therapist. Self-confrontation via videotape playback allows the individual to know more about his behavior and thus to do something about changing it. This concept of increased awareness leading to increased change is also voiced by Perlberg, Peri, Weinreb, Nitzam and Shimron (1972) who used tapes to improve teaching skills of dentists.

One caution in using tapes during teaching has been voiced by Metzner and Bittker (1973) who recommend interspersing use of videotape with live commentary by the educator and provision of frequent opportunity for student participation. They warn of the disadvantage of overuse of tape in lieu of faculty-student interchanges.

In summary it appears useful to use videotapes judiciously when teaching to expose students to a variety of clinical situations, to provide opportunity for increasing their observational skills, and to give them

constructive feedback on their performance.

Evaluation of Training Programs. If the literature regarding the specific area of family assessment training is sparse, even more so is the literature pertaining to the evaluation of training programs. Few experimental studies have been documented. Schopler, Fox, and Cochrane (1967) evaluated their students' ability to assess a family by devising rating scales to compare students' reactions to individual and family live interviews. Students, however, were not randomly assigned to experimental groups. Flomenhaft and Carter (1975), Stier and Goldenberg (1975), and Goldenberg, Stier, and Preston (1975) were among the first to attempt to measure any changes in their students' work performance, attitude, or role performance. Most of the remaining literature on evaluation of family therapy training is of a descriptive impressionistic nature (Beal, 1976; Cleghorn & Levin, 1973; Cohen, Gross, & Turner, 1976; Ferber & Mendelsohn, 1969; Flint & Rioch, 1963; O'Hare, Heinrich, Kirschner, Oberstone, & Ritz, 1975). Liddle and Halpin (Note 2) cite the need for the development of instruments to evaluate training programs.

Literature Related to Teaching Methods

In reviewing the general literature on teaching methods, one might conclude that any teaching method is

as good as another since few differences exist when various methods are compared. McKeachie (1960) concluded this and stated, "Recent research on the improvement of instruction has not resulted in clear cut conclusions about the relevant effectiveness of varying teaching methods" (p.358). In another analysis, Dubin and Taveggia (1968) looked at the data from many studies of the relationship between student achievement and varying methods of instruction. They 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).

The problem with such conclusions is that specific characteristics of the studies analyzed are frequently not cited. Various subject matter areas are often compared together and specific characteristics of the instructional method are frequently omitted. Lumsdaine (1967) asserts that omitting such considerations has done more to obscure the truth than any other flaw in educational research.

In light of Lumsdaine's comments, the literature relating to teaching methods will be considered under four sub-headings:

- i. Literature Related to the Lecture Method
- ii. Literature Related to the Small Group

Discussion Method

iii. Literature Related to Student Projects

iv. Literature Related to Indices of Student

Evaluation of Teaching Methods

Literature Related to the Lecture Method. One study which is fairly specific in outlining the types of teaching methods, subject matter, and outcome measurements used is Joyce and Weatherall's experiment (1957) which compared four methods of teaching statistics and pharmacology to 53 pre-clinical medical students. They used lectures, discussion groups, practical classes, and unsupervised reading. In lectures, material was presented by the lecturer's spoken word and no deliberate response was required from the 13 students. This contrasted with the other methods in which students either actively discussed the material, received verbal feedback on their experimental work, or engaged in unsupervised reading. The type of outcome measured was knowledge achievement on short answer and multiple choice tests. Results indicated no significant difference between methods on the statistics material, but on the pharmacological material, discussion groups and lectures were more effective than practical classes or reading.

In addition to measuring academic performance, Joyce and Weatherall compared factors relating to differences in teaching methods. The first factor was student

impression of the method. They found that the mean performance of those students who preferred one method did not differ significantly from that of those who liked it least. A second factor was expenditure of time in different methods of teaching. They found that the average time consumed was, for students, lectures 2 hours, discussion groups $2\frac{1}{2}$ hours, and practical classes or reading 3 hours per session or topic. For faculty, the economics of the lecture method was clearly demonstrated. They estimated .05 hours per session per student as compared to .3 hours per session per student for discussion groups and .33 hours per session per student for practical classes of the same scale.

It may be concluded from Joyce and Weatherall's study that discussion groups lead to slightly more gain of knowledge than lectures or practical sessions but in terms of results obtained for amount of work done, lectures ranked highest.

In the subject area of cardiology, Manning, Abrahamson, and Dennis (1968) did not support the efficiency of the lecture method which Joyce and Weatherall had demonstrated. Using 148 physicians, they compared

four teaching methods: programmed text, textbook, lecture/demonstration and lecture/workshop. Subject content was the plotting of mean cardiac vectors for the 12 lead EKG. The same multiple choice instrument was used as pre and posttest and no significant difference among the gain scores was reported. There were differences between the methods, however, in amount of time spent learning the material: standard text book - 48.9 minutes, lectures/demonstration - 90 minutes; programmed text - 72.8 minutes; lecture/workshop - 80 minutes. The authors did not report the amount of time which they spent preparing the program text but did state that this preparation influenced the use of the other teaching modes. Thus, faculty time expenditure could not adequately be measured.

Miller, Allender, and Wolf (1965) also looked at time expenditure as well as achievement in their comparison of programmed text, teaching machine, and lecture methods in physiology. Results indicated the programmed instruction group was significantly higher than other groups on test performance one week after the study was concluded ($p < .05$). The test consisted of sixteen knowledge type questions and fifteen application type questions. In addition to test achievement, a second outcome measured was the amount of study time. Programmed text method saved 20% and teaching machine method 10%

as compared to the lecture method. There were no significant correlations between amount of study time and test scores.

Other studies which compare the lecture method with varying teaching methods have focused on different outcome measures than achievement and time efficiency. Diemer and Mazzocco (1976) in comparing lecture and independent study methods of instruction in dental radiology measured personality effects as well as knowledge achievement. They looked at group dependency vs self-sufficiency in need achievement and found no significant differences between methods on either measurement instrument.

A second study which looked at student attitudes was reported by Elder, Meckstroth, Nice, and Meyers (1964) who compared a linear programmed text, teaching machines, and regular lecture presentation. Subjects were 36 medical school juniors who were engaged in the study of radiology. The students who were not in the lecture group reported a negative attitude towards self-learning.

In reviewing the literature on teaching methods, one finds several studies in which lecture was compared with other methods, but the investigators did not use experimental designs according to Campbell and Stanley's classification (1963). Holt (1975) compared a lecture/demonstration method with self-instruction in the

teaching of audio-visual equipment operation. Results indicated that the self-instruction group was significantly higher in cognitive information at the .05 level of probability. Students, however, were not randomly assigned to groups.

Another study which does not employ an experimental design is reported by Stritter, Burford, Johnson, and Talbert (1973) who compared medical students at the University of North Carolina with National Board of Medical Examiners candidates in two subject areas, pharmacology and obstetrics - gynecology. Lecture, seminar, and textbook methods were grouped together and contrasted with a self-instruction method for University of North Carolina students. No report of the teaching method employed for NBME candidates is given.

The fact that comparative studies which do not employ experimental designs are reported in the literature confirms Lumsdaine's comments (1967):

Both in summaries of research and in original reports, results of relatively rigorous studies are hopelessly confused with results of studies that have employed non experimental or otherwise inferior methods; and methodological requirements for sound inferences about the effects of variables are grossly violated... The most critical need for fruitful research calculated to improve the effectiveness of teaching is to clear up the present state of methodological and related reportorial chaos. I believe it is no exaggeration to say that at present it is difficult indeed to know which reported "findings" are to be believed and which are not. (p.243)

In summary, a review of the literature relating to the lecture method reveals a need for specification of independent variables and experimental studies. In addition, there is a need for multiple outcome measurements as well as achievement test measurements.

Literature Related to the Small Group Discussion Method. In contrast to the setting of the traditional lecture, students in a discussion group setting should be active rather than passive participants in the teaching - learning situation. The opportunity should be provided for the student to clear up hazy points, correct faulty learning, and maximize her/his motivation and interests. In discussion groups, learning is considered to be an event of social interaction, an outcome of personal encounter between teacher and students. Thus, emphasis in the classroom is shifted from the instructional activities of the teacher to the interaction between students and teacher and to procedures for socializing the teaching - learning situation.

This emphasis on interaction between student and teacher is particularly evident in the field of psychiatry because of psychiatrists' interest in group therapy and group dynamics. The Group for Advancement of Psychiatry Report (1958) points out five advantages of using interactive - inquiry discussion groups. First, provision of emotional support for the student is given as the

student discards old modes of thought and tries to learn new unfamiliar ones. This advantage of having emotional support is echoed by O'Hare et al. (1975) in their discussion of learning family therapy by group supervision. They however, also point out a disadvantage of the small group session, namely the rivalry and competition which the setting engenders between members.

A second factor in using the discussion method is the increased opportunity for student involvement in the learning task. Participation is fostered by recognition and encouragement of a range of contributions and ideas, which the group needs for effective inquiry. Cohen et al. (1976) support this need for participation in learning family therapy. In addition, they point out the transitional developmental stages of group participation. That is, initially members rely heavily on faculty input, but then progress to sharing their own opinions with increasing self-confidence. Balint (1969) also noticed this shift from inequality to equality between pupil-teacher relationships in his post-graduate seminars. The negative factor for some students of initial presentation of multiple points of view, however, cannot be overlooked.

The third advantage of using interactive-inquiry groups is the chance for students to experiment and practice new behaviors in a permissive setting. This

is important in learning family assessment because the course content can be applied to students' own small groups. For example, O'Hare et al. (1975) cite the instance in which students were having difficulty exposing their learning problems to other group members. They tended to defend themselves and each other just as a family would. In discussing their own group behavior, they came to an increased understanding of family dynamics.

The opportunity for feedback to the student is the fourth factor which group discussions foster. Phillips (Note 3) in discussing family therapy group supervision vs individual supervision is emphatic on the benefits of group feedback.

If we had to give up one method it would be individual supervision because of the general effectiveness of peers in aiding in the supervisory process. After a few months of deep involvement and high motivation plus the caring for one another that develops, the peer feedback can be most frank, honest and helpful to the point - more so than most individual supervisors can be. The group processes of consensus, support, cohesion and confrontation carry much more weight than the usual one-on-one supervision. Moreover each encounter is a learning experience for most of the group as well as the one presenting the case. The social and financial economy is obvious. (p.4)

The fifth characteristic of discussion group teaching is the lack of "closure" which impels the student to apply new methods she/he has learned to unsolved problems. Discussion groups generally generate several hypotheses to explain what has been observed. Each member thus comes

upon new questions and alternatives. GAP (1958) stresses that this training in evaluation and integration of information is especially important to a physician who must evaluate each patient as a unique problem not to be solved by rote memory of procedures.

A detailed study related to comparing differences between discussion and lecture methods was carried out by Bloom (1953) who analyzed students' thought processes during the two methods. His sample included college students from 5 lecture classes and 29 discussion classes on the subject areas of social science, biological science, humanities, mathematics, and natural sciences. The investigator made sound recordings during class and played it back to students within two days. The tapes were stopped at various critical points and students were asked to state the thoughts they recalled having during the class. The thoughts recalled from lectures were compared with those they recalled from discussion classes. Results indicated that more irrelevant thoughts, about words or phrases used, occurred during lecture (19%) than during discussion (7.5%). Lecture students also reported more thoughts (12%) about events not occurring in the classroom than did discussion students (7%). With respect to relevant thoughts, Bloom concluded that during the lectures, students evidenced more simple comprehension of subject matter than they do during a discussion.

(22% vs 9%). However, thoughts involving problem solving or synthesis occurred more frequently during discussion (17%) than during lecture (8%).

Following this study, Bloom contrasted discussion and lecture according to educational objectives:

If the objective of education is the development of knowledge about a topic or field, the lecture is a far more efficient method of communicating such knowledge....However, if the objective is development of abilities and skills which are problem-solving in nature, the least efficient discussion is superior to most of the lectures. (p.169)

The problem remains however, that in most studies the achievement objectives of the teaching method is either not clearly delineated or else is a combination of two objectives. Ruja's study (1954) is an example of this latter problem. He compared discussion and lecture methods for psychology and philosophy courses using achievement tests which were a combination of factual knowledge and reasoning with facts. Performance outcomes differed for the two courses. In psychology the lecture method achieved more than the discussion method, while in philosophy both methods achieved approximately the same. Ruja concluded that perhaps the difference in subject matter could be important.

In summary, it seems that discussion facilitates more active student involvement in the learning process than does lecture. Discussion is probably more effective than lecture for teaching cognitive skills such as

interpreting knowledge and solving problems. Discussion is probably less effective than lecture in facilitating the acquisition of factual knowledge.

Literature Related to Student Projects. A few studies have compared the outcomes of courses taught by lectures with those in which students planned and carried out projects. One such study is Ward's (1956) in which he compared lecture/demonstration method with a group study method in physical science instruction for general education college students. These results indicated that project students of high ability showed greater gains in interpreting facts and principles than lecture students of high ability. There were no differences, however, in knowledge of facts and principles regardless of students' ability levels.

The finding that projects or increased student involvement tend to be more advantageous in helping students to achieve cognitive skills such as interpreting data, is consistent with other studies (Costin, 1972). However, the question remains, is it the project method which is the main variable or is it the discussion process inherent in planning and carrying out projects which is the main factor.

Another study which attempted to delineate the differences between projects and lecture was Novak's (1958) in which a comparison was made between a project method and a lecture/demonstration method. Results

showed no significant differences in means on factual knowledge, problem-solving, retention of factual knowledge and gain in scientific attitude. The knowledge acquired during the project work itself, however, was not measured. The same course material was just presented more rapidly for the project group and a six week period was devoted exclusively to individual work.

Literature Related to Indices of Student Evaluation of Teaching Methods.

Most research on comparing teaching methods looks at outcome measures of achievement. Several studies, however, have examined other indices to evaluate differences in methods. Indices which will be reviewed in this section include student preference of method, attitude towards course content, satisfaction with method, and opinion of method effectiveness. Another index, amount of time expenditure on varying methods has been reviewed with previous studies.

James (1962) postulated that personal preference for a method might be a factor in learning. In an experiment using 503 airforce personnel, he obtained individual preferences for learning from lecture or print. He randomly subdivided each preference group so that half received instruction from their preferred mode and half did not. Results indicated no significant difference between groups.

This same conclusion that receiving instruction in the preferred channel appears to have little influence on learning is also supported by Frantz (1976) and Joyce and Weatherall (1957). There may however be a distinction between the student's preferred channel and his best functioning one. Ingersoll (cited in Levie & Dickie, 1973) identified visual attenders and aural attenders by their performance on a bisensory task and found that visual attenders recall more visual stimuli whereas aural attenders recall more aural stimuli. Thus, students who have developed such stable response characteristics may be at a disadvantage in learning from audio-visual presentations when information in both channels is not redundant. One can conclude from this that having channel redundancy is probably an advantage to most people and that method preference is most likely a function of environmental and personal conditioning.

Attitude toward course content is another factor which may contribute to achievement. One study which compared three teaching methods attempted to determine whether student attitudes toward children would be measureably influenced by a course in child development. Leton (1961) found no significant difference between lecture, case-centered and group centered methods either in achievement or on a parent attitude survey.

Attitudes toward teaching methods themselves have been evaluated in several studies. Welser, Lewis, and Stockton (1970) asked students in an opinion evaluation to compare an audio-tutorial method with a lecture method in the teaching of canine radiographic anatomy. Their conclusions were that the audio-tutorial method increased involvement of the student in the learning process and changed the role of the instructor "from a pedagogical machine to a diagnostician of the student's learning problems" (p.322). Morris (1976) commented on the usefulness of student feedback in making decisions regarding curriculum content and learning experiences.

DiVesta (1953) also evaluated several teaching methods to determine student opinion of the relevant productiveness of different methods. His ratings, based on questionnaires given to students, indicated that lectures and staff exercises (case conferences) were the most productive activities.

In summary, since learning is a multi-faceted process, it appears useful to obtain several indices of student evaluation in addition to achievement scores when comparing different teaching methods. These evaluations are of interest in themselves and shed a light on the efficacy of the method.

Summary

A review of the literature indicates several problems with studies which attempt to compare different teaching methods. Student samples are often biased in selection and incompletely described so that generalizations from them to the population are not possible. Measurement instruments employed to test achievement are often of unreported reliability and/or validity. Few studies have compared achievement of knowledge with the learning of cognitive skills, such as, interpretation and problem-solving. Tests usually combine the measurement of several learning objectives. Situations in which the experiments were conducted are frequently insufficiently described. A further problem is the incomplete operational description of the teaching methods.

This review also gave emphasis to literature pertaining to the use of videotape. The articles reviewed clearly stressed the usefulness of videotape in family therapy training programs and in teaching general psychiatry in medical schools. However, little work is evident on the use of videotape to teach family assessment in medical schools.

Chapter 2

Methodology

Introduction

The purpose of this study was to determine the effects of three teaching methods of family assessment by comparing scores of groups of first year medical students. This was accomplished by obtaining achievement tests (See Appendix 1.) and questionnaire (See Appendix 2.) data.

Approval for the study was secured from the Ethics Committee, Continuity Course Committee, and the Associate Dean (Education), Faculty of Medicine.

Design

The basic design for the study was a pretest-post-test design with random assignment of students to the experimental treatments. The design is given in Figure 1.

Figure 1

Research Design

		Experimental Treatments					
Weeks		1	2	3	4	5	6
Random Assignment of 72 Students	Q	O R I E N T A T I O N	n=20	Method 1 Lecture/ Demonstration			Q U E S T I O N S A I R E
	U						
	E		Small Group Discussion/ Demonstration n=21	Method 2			
	S						
	P		Small Group Discussion/ Student Presentation n=17	Method 3			I T O N S A I R E
	T						

Note: Total N=58
Cross hatch indicates no class for
Method 3.

Since student opinions about the course were also being sought, questionnaires were administered prior to each test. By administering the questionnaire before the test, the students' responses would not be biased by the test.

The teaching methods to which the students were assigned were: Method 1: Lecture/Demonstration, Method 2: Small Group Discussion/Demonstration, and Method 3: Small Group Discussion/Student Presentation.

The same subject matter topics were scheduled by the unit manager in the same sequence throughout the unit. The unit objectives and suggested reading assignments were the same for all students.

The similarities and differences between the methods are described in Table 1.

Table 1

Similarities and Differences Between Teaching Methods

Factors	Methods		
	1	2	3
	Lecture/ Demonstration	Discussion/ Demonstration	Discussion/ Presentation
1. Group Size	24 students	4 groups of 6 students	
2. Type of Teacher	1 family therapist	Co-preceptors: 1 family physician and 1 social scientist/family therapist	
3. Nature of Classroom Activities	Determined by preceptor Preceptor explains assessment concepts	Determined by preceptor with group consent Preceptor encourages students to develop the assessment	
4. Use of Teaching Aids	Uses slides and blackboard to present factual material	Can use blackboard to further clarify discussed material	
5. Use of Videotape	Preceptors present faculty prepared and edited videotapes of family assessment by a family psychiatrist		Student pairs: -prepare a videotape of a family interview -present tape and assessment to the group
6. Objectives, Core Document, Written Handouts	Same for All Methods		

The main differences between the methods was exemplified in Factor 3: Nature of Classroom Activities and Factor 5: Use of Videotape. Responsibility for classroom activities, for example, subject matter emphasis, was determined by the preceptor in Method 1. She expressed her own opinion regarding unit content and the potential value to the students of the material selected. In Methods 2 and 3, the preceptors expressed their attitudes regarding unit content but would then give students opportunities and responsibility for generating their own criteria for value judgments. In Method 1 the preceptor would invite questions from the students but not encourage talk between the students to clarify concepts. In Methods 2 and 3, the preceptors stimulated student discussion of their reaction to course material.

Use of videotape was another major difference between the teaching methods. In Methods 1 and 2, the preceptors used faculty prepared, edited, videotapes of family assessments which had been done by a family psychiatrist whereas in Method 3, student pairs prepared a videotape of a family interview and presented the tape and assessment to the group. The presenting students had to point out verbal and non-verbal evidence from the tape to support their assessment.

Selection and Assignment of Subjects

The 72 first year medical students at the University of Calgary were selected as subjects for this study. All students in the Class of 1979 who were repeating the unit, Family in Health and Illness, were excluded.

Students were randomly assigned prior to the start of the unit to one of the three teaching methods for the study. Each student was given a code number and a Table of Random Numbers was used to designate the assignment. Method 1: Lecture/Demonstration had 24 students. Methods 2 and 3 each had 24 students who were randomly assigned to four sub-groups of six students for the discussion sessions.

Attrition due to failure to take both tests, illness, refusal to participate, and/or scheduling problems, reduced the sample from 72 to 58 students. Method 1 had 20 students; Method 2 had 21 students and Method 3 had 17 students. Characteristics of this sample are given in Table 2.

Table 2
 Characteristics of Students
 in Each Teaching Method

Characteristics	Method		
	1	2	3
Sex			
Males	14	11	14
Females	6	10	3
Marital Status			
Married	9	3	6
Never Married	11	17	11
Relevant Work Experience			
Yes	13	7	6
No	7	14	11

Selection and Assignment of Preceptors

Preceptors for this study included 12 physicians and 12 social scientists/family therapists who were paired by the Unit Manager, who was not involved in the study. These 12 pairs of preceptors were then randomly assigned by means of a Table of Random Numbers to the three teaching methods. The eight pairs for Methods 2 and 3 taught during the study. The four pairs for Method 1 taught in the remainder of the course after

the research was completed. During the research period, a senior experienced family therapist was designated as the preceptor for Method 1.

Characteristics of the preceptors are given in Table 3.

Table 3
 Characteristics of Preceptors
 in Each Teaching Method

Characteristics	Method		
	1	2	3
Mean Age	32	33.1	36.1 ^a
Sex			
Males	0	5	4
Females	1	3	4
Profession			
Family Therapist	1	2	0
Social Worker	0	2	4
Family Practitioner	0	4	3
Pediatrician	0	0	1
Teaching Experience in this Unit			
Experienced	0	4	3
Inexperienced	1	4	5
Mean Number of Years Experience	0	2	3
Experience With Family Interviewing			
Experienced	1	8	7
Inexperienced	0	0	1
Mean Number of Family Interviews			
In past three years	1,000	507	297 ^a
In past four weeks	29	17	15 ^a

^aBased on seven out of eight respondents

Procedure

The first six teaching sessions of the unit, Family in Health and Illness, were included in this study. Each session was held once a week for two hours on Wednesday morning from 0830 to 1030. These six sessions comprised one third of the unit and focused on the topic of family assessment.

A detailed account of the collection of data follows:

1. Prior to the first teaching session, the investigator met with the 25 preceptors assigned to teach the unit. The research study was explained to them; similarities and differences between the teaching methods were discussed. Preceptors were requested to fill out a questionnaire for the purpose of gathering demographic data.
2. At Teaching Session 1 the unit manager gave a 10 minute introduction and overview of the unit to 31 students who were assembled in the lecture theatre. The objectives of the first six weeks of the unit were also reviewed. Students then answered a 10 minute impersonally administered questionnaire and took a 75 minute pretest. Students spent the next 15 minutes in their respective method groups. This time was devoted to a discussion of course logistics. Students were asked not to discuss their method's

teaching content with other students. The remaining 27 students in this sample did not attend Teaching Session 1 because of an unexpected party the night before the test. Eighteen took the pretest later that same day and the remainder took it within eight days. (See Appendix 3 for a detailed account of test data collection.)

3. Prior to each teaching session all preceptors were instructed in family assessment theory by a family psychiatrist.
4. During Session 2, Method 1 students met with the family therapist for lecture/demonstration, while Method 2 students met with their respective preceptor pairs for small group discussion and a demonstration of a prepared videotape. Method 3 students did not have class but were given that time to prepare their video of a family assessment.
5. Students in all three methods met with their respective preceptors to learn family assessment during Sessions 3, 4 and 5. Each week, student pairs in Method 3 presented their videotape of a family assessment.
6. For Session 6, students assembled in the lecture theater to answer a 10 minute impersonally administered questionnaire and take a 75

minute posttest. (See Appendix 3 for a detailed account of test data collection.)

Chapter 3

Instruments and Data Analysis

Introduction

Two instruments, a pretest and a posttest, were used to measure student achievement. To measure student opinions, a pre-unit and post-unit questionnaire were devised. Faculty opinions were also measured by a pre-unit and post-unit questionnaire. In this section, tests and questionnaires will be discussed separately.

Tests

In examining the pre and posttests, eight aspects will be considered: type, design, time limitations, scores, validity, parallelism, interrater reliability, and intrarater reliability.

Type. The entire test was written and composed of short answer and essay restricted type questions. The advantage of the essay type of test for medical students has been documented by Steele (1975) who found the test did not cue students to observe specific psychiatric symptoms. Physicians do not generally experience a cued situation in their practice but rather must gather data themselves and interpret them. Thus essay type tests, although of questionable reliability, do tend to have more validity in that students must show the capacity to use facts constructively.

Joyce and Weatherall (1957) contend that there are

several disadvantages to using essay type tests for medical students. Reliability is never very high nor has validity been established by comparing students' test marks with their subsequent professional record.

The problem of predictive validity, however, is not limited to essay type tests. Thurnblad, Muslin, and Loesch (1973) found that after two years of pre-clinical work students did well in a multiple choice test of psychiatric skills. After that no further increment was found. Students at the clinical level performed no better than students who had limited patient contact. Thus, the predictive validity of the test was questionable. Cline and Garrard (1973) also support this finding in their evaluation of the SAID teaching program in psychiatry. The choice of test type is governed in part by the nature of the subject matter and the type of learning to be evaluated.

In medical education, multiple choice tests, preferred for high reliability, have been used in comparative studies of teaching methods in pharmacology (Joyce & Weatherall, 1957), pharmacology and obstetrics-gynecology (Stritter et al., 1973), cardiology (Manning et al., 1970), and physiology (Miller et al., 1965).

In psychiatry, multiple choice tests have been used to measure factual knowledge. To evaluate observational skills, interviewing ability, and problem solving ability,

other rating measures have been devised (Randels et al., 1976).

Since the subject matter for this study is in the psychiatric sub-specialty area of family assessment, the investigator chose to use mostly short answer type questions to measure factual knowledge and restricted essay type questions to measure application, analysis and synthesis.

That the test is investigator devised is based on the lack of standardized tests to measure achievement in this subject area. In the field of family therapy, few instruments, either objective or subjective, are available to evaluate the effect of teaching family assessment.

Design. Each test was designed in two parts. Part A related to family assessment theory while Part B related to family problems encountered in actual clinical settings. Two modes of presentation were used to convey the family data: a written genogram and a videotape of family interaction.

The genogram, a diagram of the family constellation, was the first mode presented. Demographic information, such as name, age, sex, occupation, and length of marriage was included. The students were tested on their ability to interpret the data and make a family developmental assessment. Five edited videotapes of family interviews

were the second mode of presentation. Students were shown a tape and then expected to use this data in answering written questions. Tapes varied in length from three to seven minutes.

The choice of written and videotaped presentation modes was guided by a desire to neutralize and standardize the test situation as much as possible. Thurnblad et al. (1973) proposed that "watching a videotape equalizes the situation so that less advanced students with clinical sophistication can better compete with advanced students" (p.570). Miller and Tupin (1972) substantiate this and comment on the advantages of relevant, representative and always available data. The use of edited tapes directs students' attention to relevant material by eliminating distracting extraneous parts.

Time Limitations. The total test took approximately 75 minutes for students to complete. Part A took 15 minutes and Part B took 60 minutes of which approximately 18 minutes was taken up with watching the videotapes.

Scores. Three test scores were tabulated and analyzed: Part A, Part B and Total. Part A maximum score was 61 points; Part B maximum score was 168 points. The maximum Total score was 229 points.

Question scores were weighted by the raters prior to the beginning of the study. More weight was given to

those questions pertaining to content emphasized during the teaching sessions.

Each test and questionnaire was given a code number to reduce possible rater bias in scoring. Tests from different methods were mixed and first one question from all tests was marked, then another was marked. Responses were judged according to model answers formulated by the raters prior to the beginning of the study.

All test results were confidential and available only to the research personnel. Students, however, had the option to receive feedback on their test results.

Test Validity. Attempts were made to establish the validity of the tests. Face validity was determined by two independent experts familiar with a systems model of family assessment. One specialist was an educational psychologist and Associate Director of the Family Therapy Program, University of Calgary, Faculty of Medicine. The other expert was a family practitioner and Manager of the Unit, Family in Health and Illness.

Content validity was judged in two ways. First, test question objectives were matched against Bloom's Taxonomy of Educational Objectives (1961) to determine if there was a representative sampling.

The tests were composed of two parts, each of which was based on different goals derived from the Taxonomy. That is, knowledge of facts and principles was measured

separately from cognitive skills, such as, problem solving. Part A primarily tested students' factual knowledge of the theory of family assessment. Bloom's first educational goal, "Knowledge", involves the recall of specifics, universals, methods, processes, patterns, structures or settings. The thought process involved in responding to a knowledge item would be one of remembering facts.

Part B primarily tested students' abilities in Bloom's remaining five categories of educational objectives.

"Comprehension" is a degree of understanding which enables the individual to know what is being communicated, and to make use of the material or idea without necessarily relating it to other material or seeing its fullest implication. The thought process involved in responding to this type of item is more than "Knowledge" but less than "Application".

"Application" involves the use of abstractions in particular situations. Successful students of family assessment must have the ability to apply the concepts of family assessment to practical situations in which a procedure for assessment has not been specified or memorized.

"Analysis" emphasizes the understanding of the breakdown of material into constituent parts and the way they are organized. Students must show the ability to analyze a family's communication pattern and to distinguish

relevant from irrelevant statements.

"Synthesis" emphasizes the combining of elements or parts in such a way as to constitute a pattern or structure not clearly there before. Students must show the ability to make discoveries and generalizations.

"Evaluation" includes the process of making judgements about the value, for some purpose, of ideas, solutions, methods, works, material, etc.. Students must show the ability to distinguish between a family's adaptive and maladaptive functioning.

The distribution of test questions classified according to Bloom's Taxonomy is given in Table 4.

Table 4

Total Number of Test Questions Classified
According to Bloom's Taxonomy

Classifications	TESTS	
	Pre	Post
Part A		
Knowledge	9	7
Comprehension	1	1
Application	1	3
Part B		
Comprehension	8	4
Application	9	13
Analysis	16	15
Synthesis	2	2
Evaluation	5	4

A second method employed to determine content validity was the use of experts' opinions. The same educational psychologist and the unit manager judged the sampling adequacy of the questions in relation to unit and session objectives as well as reading material. The tests were judged valid.

Parallelism. Parallelism for comparable tests is frequently determined by administering one form of a test to a group of individuals and then administering a second form to the same group (Kolstoe, 1973; Thorndike, 1959).

In this study to estimate that the pre and posttests were parallel forms of the same instrument, a separate sub-study was undertaken. The subjects were 29 students who had been involved in family therapy training at the Family Therapy Program, University of Calgary, Faculty of Medicine. Each was contacted by telephone and invited to participate in the study. Twelve responded.

The following design was adopted:

	<u>Time 1</u>	<u>Time 2</u>
Group 1	Pretest	Posttest
Group 2	Posttest	Pretest

Two groups were administered the tests with no time interval between administrations. Test order was varied in that Group 1 took the pretest first and then the posttest. Group 2 took the post first and then the pretest. This variation was done to determine if there was an

effect due to the order in which the tests were administered. That is, did taking one test prior to the other produce an elevation in score on the second test.

Each subject was assigned a test time; the test order was randomly assigned to the time of test taking.

Descriptive characteristics of the groups are listed in Table 5.

Table 5

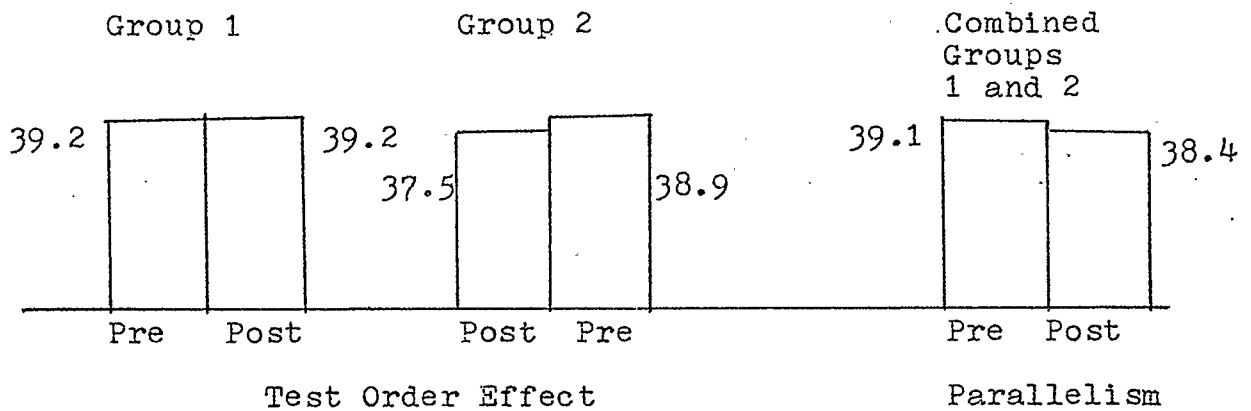
Characteristics of Parallelism Study Groups

Characteristics	Groups	
	1	2
Number of Subjects	6	6
Test Order	Pretest → Posttest	Posttest → Pretest
Sex		
Males	3	1
Females	3	5
Profession		
Clinical Clerk	2	1
Family Practice Resident	1	0
Family Practice Teaching Fellow	0	1
Family Practitioner	2	0
Nurse	0	1
School Counsellor	1	0
Educational Psychology	0	3
Graduate Students		

To establish the equivalence or parallelism of test forms, equal means and variances must be present when the forms are administered to a defined group of subjects (Wiersma, 1975). Part A, Part B, and Total pre- and posttest means are given in Figures 2, 3, and 4.

Figure 2

Part A Pre and Posttest Means

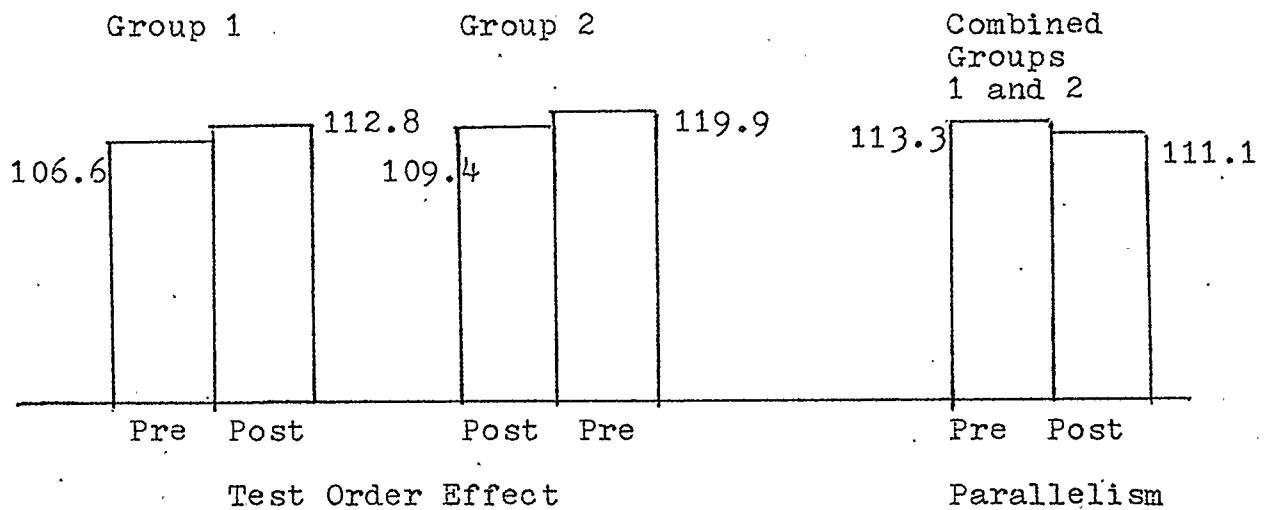


Note: Maximum Score= 61

Height of bar indicates group mean score.

Figure 3

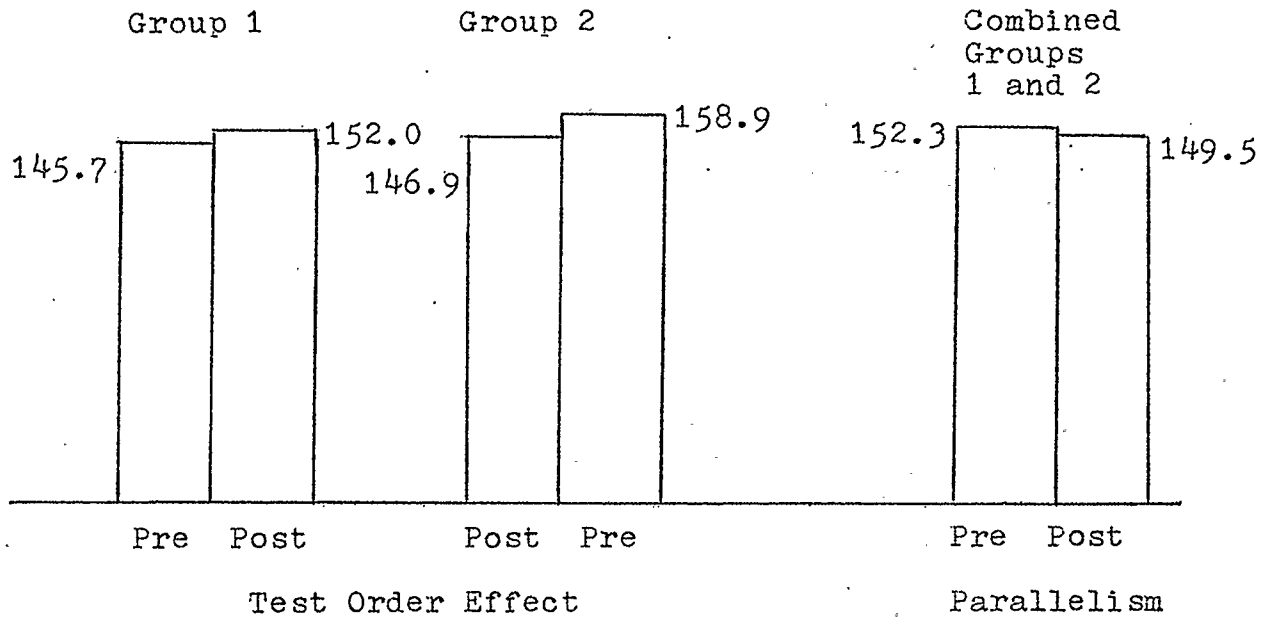
Part B Pre and Posttest Means



Note: Maximum Score= 168

Height of bar indicates group mean score.

Total Pre and Posttest Means



Note: Maximum Score= 229

Height of bar indicates group mean score.

Two-way analysis of variance with repeated measures was the statistical model used to determine if there was a significant difference between pre and posttest means. In Table 6 are given the F statistic and probability levels for the analysis.

Table 6
 Probability of No Significant Difference Between
 Tests, Test-Group Interaction and Groups

	Source of Variation	F	Probability
Part A	Tests	.104	.75
	Test-Group	.111	.75
	Groups	.191	.67
Part B	Tests	.323	.58
	Test-Group	4.802	.05*
	Groups	.176	.68
Total	Tests	.359	.56
	Test-Group	3.602	.09
	Groups	.092	.77

* $p < .05$

Results indicate that for Part A there is no significant difference between test means. For Part B there is also no significant difference between test means. There is, however, an interaction effect between groups and tests at the .05 probability level of occurrence by chance. The interaction effect can be accounted for by practice or testing effect in that the score of the second test taken is higher than the score of the first test taken. This is true for both orders of test taking. For the Total score, there is no significant difference between test means. From these analyses, it was concluded that test means were equal.

The test statistic used to determine if the variances of the dependent samples were equal was

$$F = \frac{s_1^2 - s_2^2}{\sqrt{\frac{4s_1^2 s_2^2}{n-2} (1 - r_{12}^2)}}$$

where s_1^2 and s_2^2 are the variances of the pretest and posttest samples respectively, n is the number of pairs of observations, pairing each observation of pretest with a single observation in posttest and r_{12} is the product-moment correlation coefficient calculated on the n or 12 paired observations (Glass & Stanley, 1970).

In this study variances for pre and posttests were determined without respect for test order variation. Variances are given in Table 7.

Table 7
Pre and Posttest Variances, Pearson Product-Moment Correlations and t Levels

Test Part	<u>Test Variances</u>		PPMC	t
	Pre	Post	(r_{12})	10
Part A	28.72	25.48	.04	.189
Part B	636.29	277.62	.81	2.30*
Total	753.90	391.12	.74	1.57

Note: $t_{10} = +$ or $- 2.23$ at .05 level of significance

* $p < .05$

Results indicate that the pre and posttest variances are equal for Part A and Total. Part B variances are different at the .05 level of significance.

It can be concluded from these analyses that the pre and posttest forms are equivalent for Part A and for the Total but not for Part B.

Interrater Reliability. Seven interrater reliability coefficients were estimated for each test. That is, one was determined for Part A and six for Part B. In Part B,

responses to the genogram and each of the five video selections were rated separately.

Two raters were used: a family psychiatrist designated as rater one, and the investigator designated as rater two. In practice sessions, raters were trained using the same type tests as those actually employed in the experiment. In training, each scored tests independently. Both reviewed the ratings, discussed the variations and came to a consensus with respect to the mark. The students' tests were then scored.

In estimating the interrater reliability, the number of cases and the scoring procedure used was different for several sections of the tests:

a. Part A and Genogram - 58 cases were scored separately by each rater. The average of the two measurements was then designated as the student's final score.

b. Videotape 1 - 58 cases were used in the pretest and 47 in the posttest. Again, the average of the two measurements was designated as the final score. The 11 remaining cases in the posttest were scored by the investigator alone and not included in the interrater reliability estimate.

c. Videotapes 2 to 5 - a small number of cases were selected randomly, scored by Rater 1, and compared to those scored by the investigator.

Analysis of variance with repeated measures for unadjusted data was the statistical test chosen to estimate the reliability of the ratings. Use of this test acknowledges that the variation among means from two different occasions or times is important and should be considered a source of inconsistency (Maguire & Hazlett, 1969). Analysis of variance consists of analytic process of breaking down the total sums of squares of variation from the grand mean into component parts attributable to appropriate sources, and then converting them into mean squares through division by the proper number of degrees of freedom. The mean square between people and the mean square within people were used to estimate reliability. Since the average of two measurements was used for Part A, genogram, and Videotape 1, the reliability estimate used was r_2 , the average of two measurements (Winer, 1962, 128).

$$r_2 = \frac{1 - \text{Mean Square Within People}}{\text{Mean Square Between People}}$$

For Videotapes 1 to 5, the reliability of a single measurement was calculated:

$$r_1 = \frac{\text{MS Between People} - \text{MS Within People}}{\text{MS Between People} + (K-1) \text{MS Within People}}$$

The interrater reliability coefficients for each item as well as the means, F statistic, and levels of significance between pretest means are listed in Table 8. Interrater reliability coefficients range from .76 to .99.

Table 8

Estimates of Pretest Interrater Reliability

Item/Rater	Subjects	Means	F	Signif.	r_2	r_1
Part A ₁	58	19.24	10.994	.00*	.98	
Part A ₂	58	20.34				
Part B						
Genogram ₁	58	4.49	3.643	.06	.97	
Genogram ₂	58	4.25				
Videotape One ₁	58	19.67	1.286	.26	.99	
Videotape One ₂	58	19.96				
Videotape Two ₁	12	10.08	2.725	.13		.95
Videotape Two ₂	12	10.83				
Videotape Three ₁	6	11.33	7.500	.04**		.99
Videotape Three ₂	6	12.33				
Videotape Four ₁	11	9.27	.804	.39		.90
Videotape Four ₂	11	8.50				
Videotape Five ₁	12	16.21	8.892	.01**		.76
Videotape Five ₂	12	12.38				

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

Estimates of posttest interrater reliability are given in Table 9. Interrater reliability coefficients range from .80 to .97.

Table 9

Estimates of Posttest Interrater Reliability

Item/Rater	Subjects	Means	F	Signif.	r_2	r_1
Part A ₁	58	32.91	.228	.63	.97	
Part A ₂	58	33.09				
Part B						
Genogram ₁	58	8.66	.229	.63	.93	
Genogram ₂	58	8.77				
Videotape One ₁	47	30.32	2.524	.12	.98	
Videotape One ₂	47	30.79				
Videotape Two ₁	10	13.50	.243	.63		.83
Videotape Two ₂	10	13.25				
Videotape Three ₁	5	12.40	.017	.90		.89
Videotape Three ₂	5	12.50				
Videotape Four ₁	11	21.27	1.751	.22		.80
Videotape Four ₂	11	22.18				
Videotape Five ₁	10	14.80	3.223	.11		.89
Videotape Five ₂	10	13.55				

Intrarater Reliability. Intrarater reliability estimates were obtained for the investigator. Seven reliability coefficients were estimated for each test. That is, one was determined for Part A and six for Part B. In Part B, responses to the genogram and each of the five video selections were rated separately.

Twenty-five students were randomly selected for the second rating. The time interval between ratings was approximately three weeks.

Analysis of variance with repeated measures for unadjusted data was the statistical test used to calculate the reliability estimate. The mean square between people and the mean square within people were used to estimate the reliability of a single measurement (Winer, 1962, 126):

$$r_1 = \frac{MS_{\text{Between People}} - MS_{\text{Within People}}}{MS_{\text{Between people}} + (K - 1) MS_{\text{Within people}}}$$

The intrarater reliability coefficients for each pretest item as well as the means, F statistic, and levels of significance between means are listed in Table 10. Coefficients range between .84 and .97.

Table 10
Estimates of Pretest Intrarater Reliability

Item/Time	Subjects	Means	F	Signif.	r_1
Part A ₁	25	18.22	5.515	.03**	.97
Part A ₂	25	17.24			
Part B					
Genogram ₁	25	4.99	22.770	.00*	.84
Genogram ₂	25	3.81			
Videotape One ₁	25	22.82	30.585	.00*	.94
Videotape One ₂	25	20.46			
Videotape Two ₁	25	11.50	.157	.70	.94
Videotape Two ₂	25	11.38			
Videotape Three ₁	25	11.14	1.699	.20	.92
Videotape Three ₂	25	10.78			
Videotape Four ₁	25	8.84	.695	.41	.87
Videotape Four ₂	25	9.22			
Videotape Five ₁	25	9.64	1.443	.24	.84
Videotape Five ₂	25	10.48			

* $p < .01$

** $p < .05$

Estimates of posttest intrarater reliability are given in Table 11. Coefficients range between .86 and .95.

Table 11

Estimates of Posttest Intrarrater Reliability

Item/Time	Subjects	Means	F	Signif.	r_1
Part A ₁	25	34.16	.152	.70	.95
Part A ₂	25	33.90			
Part B					
Genogram ₁	25	8.54	16.741	.00*	.89
Genogram ₂	25	7.53			
Videotape One ₁	25	30.68	34.378	.00*	.89
Videotape One ₂	25	27.82			
Videotape Two ₁	25	14.42	2.708	.11	.89
Videotape Two ₂	25	13.78			
Videotape Three ₁	25	13.39	1.145	.30	.86
Videotape Three ₂	25	12.92			
Videotape Four ₁	25	18.58	1.793	.19	.90
Videotape Four ₂	25	17.96			
Videotape Five ₁	25	15.78	6.404	.02**	.88
Videotape Five ₂	25	14.64			

* $\underline{p} < .01$ ** $\underline{p} < .05$

Questionnaires

Two questionnaires were devised to gather data from students in three areas: background history, satisfaction with the teaching method, and time expenditure.

The choice of the questionnaire method of data collection over other alternatives was based on the desire for reliable and easily quantifiable answers. Student records were not available nor did they contain all the background history data necessary for the study. Informal "depth" interviews and sociometric scales were not chosen due to the unavailability of appropriate scales.

The questionnaire format is an eclectic one as is evidenced by the presence of both direct and indirect questions. Each specific area to which the questionnaire is addressed has a set of structured questions. Most items have Likert scale check responses. A few items have fixed probe or follow up questions which Kornhauser and Sheatsley (1959) advise for gathering specific differentiated opinions.

As several guides (Bureau of Applied Social Research, 1948; Van Dalen & Meyer, 1962) recommend, a pretest of the questionnaires was done in two phases. During Phase I, experts in the fields of family therapy, educational psychology and sociology were consulted. In Phase II, four men and four women were personally administered

the questionnaires. Although they were not medical students, three were graduate students in health related fields. In both phases, the questionnaires were revised based on the respondents' comments.

Two questionnaires were also devised to gather data from the 17 preceptors who participated in this study. Topics included background history and opinion of teaching methods. Questionnaire construction and format was similar to student's questionnaires. Seventeen pre-unit questionnaires and 15 post-unit ones were returned.

Analysis of Data

Since data in this study were from two types of instruments, each will be considered separately. Test data will be discussed first, followed by questionnaire data. (See Appendix 4 for raw data.)

Tests. Two-way analysis of variance with repeated measures was the statistical model used to determine if significant differences existed among the means. This test is used in a two factor experiment where there are repeated measures on one factor (Winer, 1962). In this experiment the two factors are: Teaching Method (1,2,3) and Pre and Posttests (Repeated Measures, A and B). This may be represented schematically as follows:

		Test		
		Pre	Post	
		A	B	
Teaching Method	1	A_1	B_1	$1_{\bar{x}}$
	2	A_2	B_2	$2_{\bar{x}}$
	3	A_3	B_3	$3_{\bar{x}}$
		$A_{\bar{x}}$	$B_{\bar{x}}$	
		Test Means		
				Teaching Method Means

Results are given for:

1. Test Effect, i.e., $A_{\bar{x}}$ vs $B_{\bar{x}}$.
2. Test-Method Interaction, i.e., $A_1 \rightarrow B_1$ vs $A_2 \rightarrow B_2$ vs $A_3 \rightarrow B_3$.
3. Method Effect, i.e., $1_{\bar{x}}$ vs $2_{\bar{x}}$ vs $3_{\bar{x}}$.

Where significant differences existed among three levels, a Newman-Keul multiple range test was used to determine where the differences lay. A .05 level of significance was adopted.

The computer program used for the two-way analysis of variance with repeated measures was ANOV23 from the University of Alberta. Statistical Package for the Social Sciences (SPSS) was used for all other analyses.

Questionnaires. Pre and post-unit questionnaire

data were of three types of classification. Background data were of nominal classification; satisfaction data were of ordinal level; time expenditure data were of a ratio scale.

Parametric statistical tests were used in analyzing questionnaire data without respect to the classification level. Anderson (1970), Burke (1969), and Gaito (1970) support this use of parametric analysis of data although Siegel (1956) is critical of its use.

To test differences between methods in general satisfaction levels reported on the post-unit questionnaire, a one-way analysis of variance was carried out.

To test that there is no significant difference between teaching methods in student time expenditure, a one-way analysis of variance was carried out on transformed data. That is, on the post-unit questionnaire, students had to indicate their time expenditures from a given range of hours such as, 1-2 hours, 3-4 hours, etc.. The mean level (eg. $1\frac{1}{2}$ hours) was chosen and used in the data analysis.

Two-way analysis of variance with repeated measures was performed to determine the influence of background variables on Part A, Part B, and Total scores. A .05 level of significance was adopted in testing all hypotheses.

Chapter 4

Results and Discussion

Introduction

The purpose of this study was to determine the effects of three teaching methods of family assessment by comparing scores of first year medical students. This was accomplished by obtaining achievement data for students in Method 1: Lecture/Demonstration, Method 2: Small Group Discussion/Demonstration, and Method 3: Small Group Discussion/Student Presentation. This chapter first presents the hypotheses and statistical findings of the study, and then discusses their relevance.

The following abbreviations are used in presenting the hypotheses.

1. Lec is the abbreviation for Method 1: Lecture/Demonstration teaching method.
2. SGD is the abbreviation for Method 2: Small Group Discussion/Preceptor Demonstration teaching method.
3. SGSP is the abbreviation for Method 3: Small Group Discussion/Student Presentation teaching method.

Testing of Hypothesis 1: Result of Student Achievement

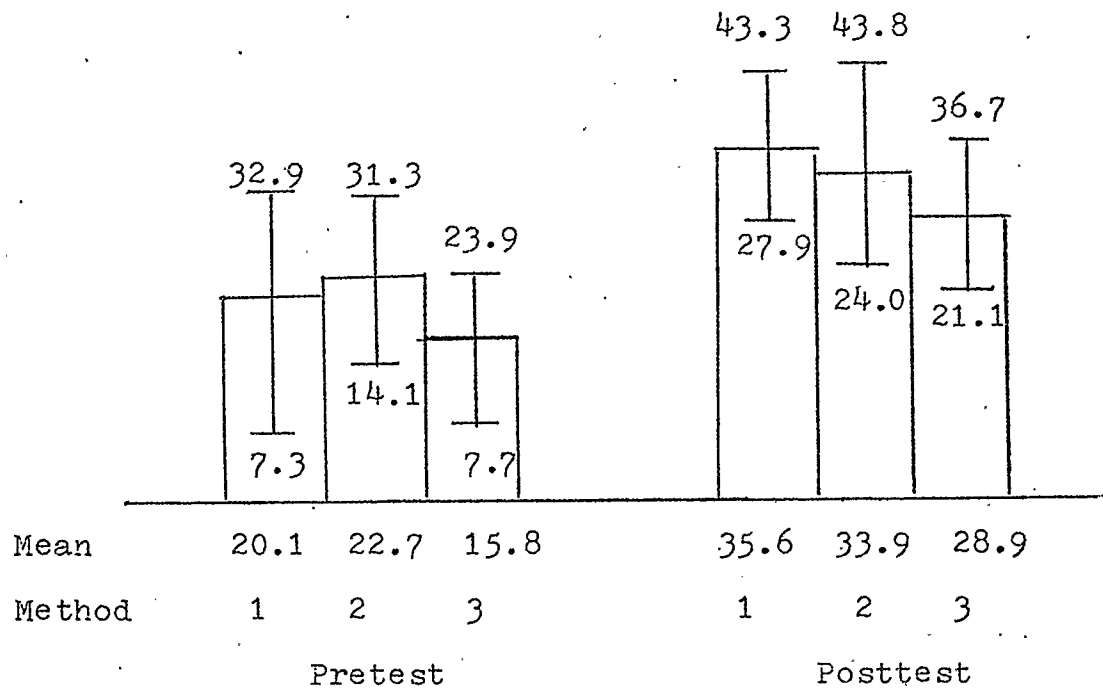
Hypothesis 1a: There is no significant difference between Methods 1, 2, 3 as measured by pretest and post-test Part A mean change scores.

$$H_0: u_1(\text{Lec}) = u_2(\text{SGD}) = u_3(\text{SGSP})$$

H_1 : The population means are not equal.

Means. Figure 5 presents Part A mean scores for the three teaching methods on the pretest and posttest.

Part A Achievement Means and Standard Deviation for
Methods 1, 2, 3



Note: Maximum Score = 61

Height of bar indicates Method mean.

One standard deviation is indicated by I.

Statistical Findings. Two-way analysis of variance with repeated measures was carried out. The F statistic and probability levels are given in Table 12.

Table 12

Probability of No Significant Difference Between
Tests Part A, Test-Method Interaction and Methods

Source of Variation	F	Probability
Tests	87.571	-.00*
Test-Method	.788	.46
Methods	3.562	.04**

* $p < .01$

** $p < .05$

Results of the analysis indicate:

1. There is a test effect. The pre and posttests are significantly different at the .00 level of probability, with the posttest being higher.

2. There is no significant interaction effect between tests and methods. That is, one teaching method is not significantly better in helping students to gain from pre to posttest.

3. There is a significant difference between method groups at the .04 level of probability. Results of a Newman-Keul multiple range test indicate Method 3 is

significantly lower than Methods 1 and 2. This most probably is accounted for by the low pretest score for Method 3.

Conclusion. Null hypothesis 1a is accepted at the .05 level of probability.

Hypothesis 1b. There is no significant difference between Methods 1, 2, 3 as measured by pretest and posttest Part B mean change scores.

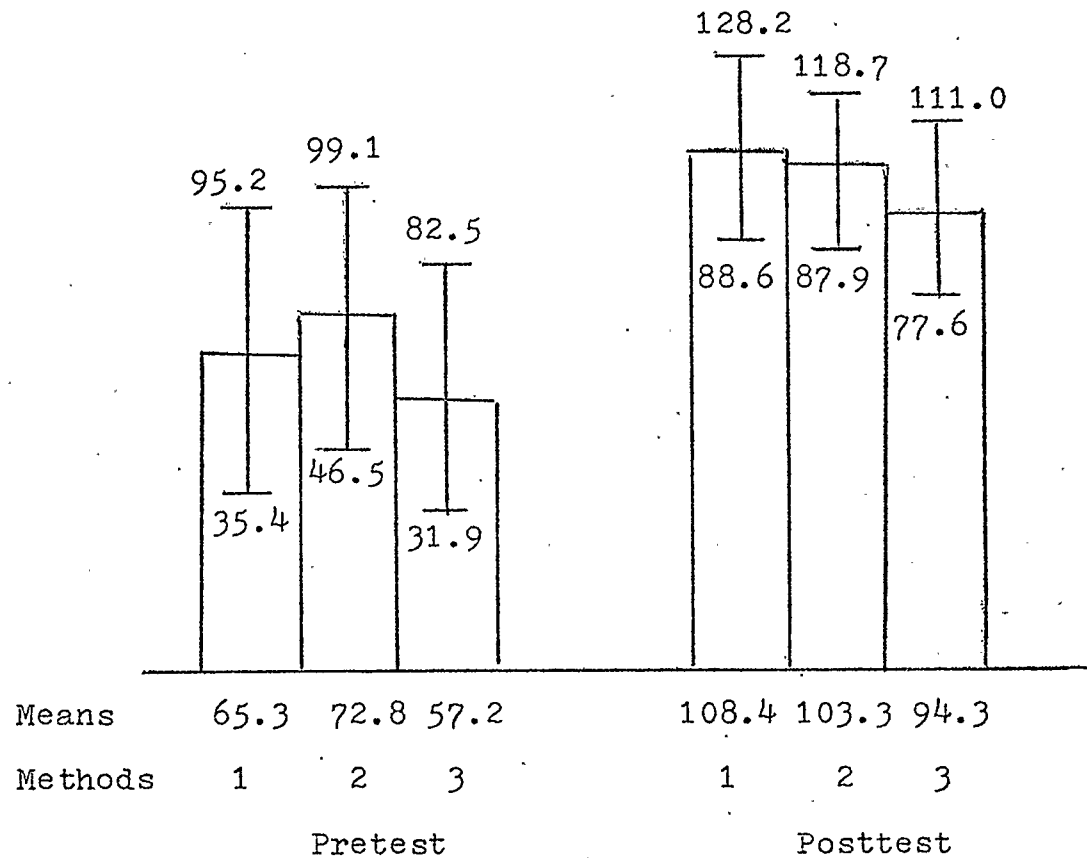
$$H_0: u_1(\text{Lec}) = u_2(\text{SGD}) = u_3(\text{SGSP})$$

H_1 : The population means are not equal.

Means. Figure 6 presents Part B mean scores for the three teaching methods on the pretest and the posttest.

Part B Achievement Means and Standard Deviation

For Methods 1, 2, 3



Note: Maximum Score = 168

Height of bar indicates method mean.

One standard deviation is indicated by I.

Statistical Findings. Two-way analysis of variance with repeated measures was carried out. The F statistic and probability levels are given in Table 13.

Table 13

Probability of No Significant Difference Between
Tests Part B, Test-Method Interaction and Methods

Source of Variation	F	Probability
Tests	126.527	-.00*
Test-Method	1.232	.30
Methods	2.383	.10

* $p < .01$

Results of the analysis indicate:

1. There is a test effect. The posttest is significantly higher than the pretest at the .00 level of probability.
2. There is no significant interaction effect.
3. There is no significant difference between method groups.

Conclusion. Null hypothesis 1b is accepted at the .05 level of significance.

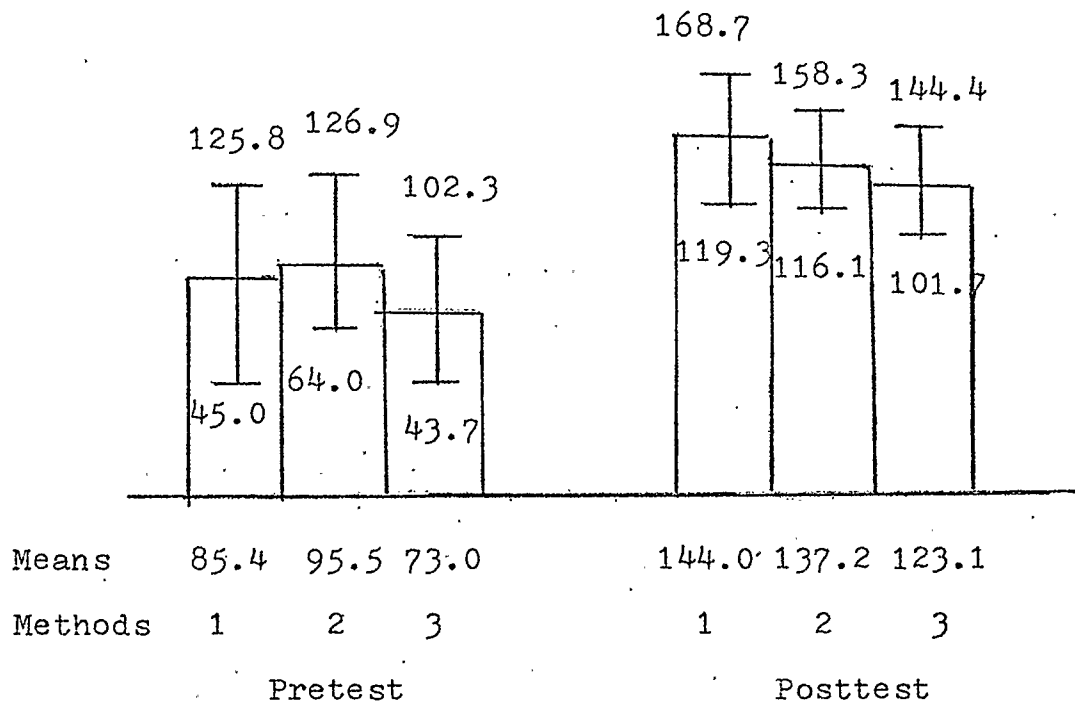
Hypothesis 1c. There is no significant difference between Methods 1, 2, 3 as measured by pretest and posttest Total mean change scores.

$$H_0: u_1(\text{Lec}) = u_2(\text{SGD}) = u_3(\text{SGSP})$$

H_1 : The population means are not equal.

Means. Figure 7 presents Total mean scores for the three teaching methods on the pretest and posttest.

Total Achievement Means and Standard Deviation
for Methods 1, 2, 3



Note: Maximum Score = 229

Height of bar indicates method mean.

One standard deviation is indicated by I.

Statistical Findings. Two-way analysis of variance with repeated measures was carried out. The F statistic and probability levels for the Total test scores are given in Table 14.

Table 14

Probability of No Significant Difference Between
Total Tests, Test-Method Interaction and Methods

Source of Variation	F	Probability
Tests	146.714	-.00*
Test-Method	1.396	.26
Methods	3.299	.04**

* $p < .01$

** $p < .05$

Results of the analysis indicate:

1. There is a test effect. The posttest is significantly higher than the pretest at the .00 probability level.

2. There is no significant interaction effect.

3. There is a significant difference between method groups at the .04 level of probability. A Newman-Keul multiple range test indicates Method 3 is significantly lower than Method 1. This can be accounted for by the low pretest score for Method 3.

Conclusion. Null hypothesis 1c is accepted at the .05 level of probability.

Testing of Hypothesis 2: Results of Satisfaction Evaluations

Hypothesis 2. There is no significant difference between Methods 1, 2, 3 in satisfaction levels as measured by post-unit questionnaire mean scores.

$$H_0: u_1(\text{Lec}) = u_2(\text{SGD}) = u_3(\text{SGSP})$$

H_1 : The population means are not equal

Means. Students were asked how satisfied they were with specific aspects of the teaching methods. Table 15 gives the mean satisfaction levels for each teaching method.

Table 15
Means for Students' Ratings of Satisfaction
With Aspects of Teaching Methods

Aspect	Method	No. of Subjects	Post Q Mean	F	Proba- bility
Method	1	20	3.1	4.457	.02*
	2	21	3.0		
	3	17	2.4		
Preceptors	1	20	3.4	.690	.50
	2	21	3.2		
	3	17	3.1		
Technical Quality of Tapes	1	20	2.2	4.468	.02*
	2	21	2.9		
	3	14	2.9		
Tape Content	1	20	3.1	3.003	.06
	2	21	3.0		
	3	13	2.5		

Note: Minimum Satisfaction = 1

Maximum Satisfaction = 4

* $p < .05$

Statistical Findings. One-way analysis of variance was carried out to determine if there was a significant difference between methods in students' satisfaction ratings. The F statistic and probability levels are given in Table 15. Where significant differences between methods were evident a Newman-Keul multiple range test was performed. Results indicate:

1. There is a significant difference between methods at the .02 level of probability in students' satisfaction with the teaching method. Method 3 is significantly less satisfactory than Methods 1 and 2.
2. There is no significant difference between methods in students' satisfaction with preceptors.
3. There is a significant difference between methods at the .02 level of probability in students' satisfaction with the technical quality of videotapes. Method 1 is less satisfied than Methods 2 and 3.
4. There is no significant difference between methods in satisfaction with tape content. The F probability of .06, however, does indicate a possible experimental effect. Method 3 is less satisfied than Methods 1 and 2.

Conclusion. Null hypothesis 2 is rejected. There is a significant difference between methods in student

satisfaction with teaching methods and videotape technical quality.

Testing of Hypothesis 3: Results of Evaluation of Time Expenditure

Time expenditure was documented in two ways: a) amount of hours students spent studying or reading, b) total time which included study time and, for Method 3 students, the time expended in videotape preparation. Each documentation will be treated separately.

Hypothesis 3a. There is no significant difference between Methods 1, 2, 3 in student study time as measured by post-unit questionnaire mean scores.

$$H_0: u_1(\text{Lec}) = u_2(\text{SGD}) = u_3(\text{SGSP})$$

H_1 : The population means are not equal.

Means. Means for student time expenditure in studying or reading are given in Table 16.

Table 16

Mean Study Hours for Students in Methods 1, 2, 3

Method	No. of Subjects	Mean
1	20	2.28
2	21	1.74
3	17	1.56

Statistical Findings. A one-way analysis of variance was carried out. The F statistic was .921 and the probability level was .40. This indicates there is no significant difference between methods in amount of hours students spent studying.

Conclusion. Null hypothesis 3a is accepted at the .05 level of significance.

Hypothesis 3b. There is no significant difference between Methods 1, 2, 3 in student total time expenditure as measured by post-unit questionnaire mean scores.

$$H_0: u_1(\text{Lec}) = u_2(\text{SGD}) = u_3(\text{SGSP})$$

H_1 : The population means are not equal.

In addition to spending time outside of class studying, students in Method 3 spent time in interviewing and preparing their videotapes of a family assessment. These hours are reported in Table 17.

Table 17

Method 3: Actual Student Time Expenditure in

Video Preparation

No. of Subjects	Hours
3	1.5
10	3.5
1	5.5
2	7.5

Since Method 3 students were given 2 hours off from class to do their interviews, this time was subtracted from the midpoint amount of hours they reported for video preparation. The time students thus expended outside of class time allotment is given in Table 18.

Table 18

Method 3: Student Time Expenditure in Video Preparation
Outside of Class Time Allotment

No. of Subjects	Hours
3	0
10	1.5
1	3.5
2	5.5

Means. This newly calculated video time expenditure was added to Method 3 study hours to account for the time Method 3 students spent outside of class. Means for total time expenditure are given in Table 19.

Table 19
Mean Hours of Student Time Expenditure Outside of
Class

Method	Mean Study Hours	Mean Video Preparation Hours (Adjusted)	Total Mean Time
1	2.28	—	2.28
2	1.74	—	1.74
3	1.56	1.84	3.29

Statistical Findings. A one-way analysis of variance was carried out to determine if there was a significant difference between methods in student time expenditure outside of class. The F ratio was 3.495 and the probability level was .04. A Newman-Keul multiple range test was carried out at the .05 level to see where the differences lay between methods. Method 3 students spent significantly more time outside of class than did Method 2 students.

Conclusion. Null hypothesis 3b is rejected.

Testing of Hypothesis 4: Results of Influence of Sex
on Achievement Scores

Hypothesis 4a. There is no significant difference between males and females as measured by pretest and posttest Part A, Part B, and Total mean scores.

$$H_0: u_1(\text{males}) = u_2(\text{females})$$

H_1 : The population means are not equal.

Means. Mean achievement scores for males and females are given in Table 20.

Table 20
Pre and Posttest Achievement Means for Males and
Females

No. of Subjects		Pretest	Posttest
Part A	39 males	17.27	31.15
	19 females	24.97	36.80
Part B	39 males	63.59	99.53
	19 females	69.83	108.30
Total	39 males	80.86	130.70
	19 females	94.79	145.10

Data from Table 20 indicate women's scores are consistently higher.

Statistical Findings. A two-way analysis of variance with repeated measures was carried out to determine if there was a significant difference between males and

females on both tests. The F statistic and probability levels are given in Table 21.

Table 21
Probability of No Significant Difference Between
Males and Females on Pre and Posttests

	F	Probability
Part A	10.371	.002*
Part B	1.878	.176
Total	4.251	.043**

* $p < .01$

** $p < .05$

Data from Table 21 indicate women are significantly higher than men on Part A and Total scores.

Conclusion. Null hypothesis 4a is rejected at the .05 level of probability for Part A and Total.

Hypothesis 4b. There is no significant difference between males and females as measured by pretest and posttest Part A, Part B, and Total mean change scores.

$$H_0: u_1(\text{males}) = u_2(\text{females})$$

H_1 : The population means are not equal.

Means. Mean achievement scores for males and females are given in Table 20.

Statistical Findings. A two-way analysis of variance with repeated measures was carried out to determine:

1. if there was a significant difference between the male group and the female group in their gain from pre to posttest.

2. if there was a significant difference between tests.

The F statistic and probability levels are given in Table 22.

Table 22

Probability of No Significant Difference Between
Tests and Test - Group Interaction

	Source of Variation	F	Probability
Part A	Tests	73.149	.00*
	Test-Group	.462	.50
Part B	Tests	111.303	.00*
	Test-Group	.125	.72
Total	Tests	125.546	.00*
	Test-Group	.003	.96

* $p < .01$

Results from Table 22 indicate:

1. There is a test effect. The posttest is significantly higher than the pretest.

2. There is no significant interaction effect.

Women do not gain more from pre to posttest than do men.

Conclusion. Null hypothesis 4b is accepted at the .05 level of probability.

Testing of Hypothesis 5: Results of Influence of Marital Status on Achievement Scores

Hypothesis 5. There is no significant difference between married and unmarried students as measured by pretest and posttest Part A, Part B, and Total mean change scores.

$$H_0: u_1(\text{married}) = u_2(\text{unmarried})$$

H_1 : The population means are not equal.

Means. Mean achievement scores for married and unmarried students are given in Table 23.

Table 23

Pre and Posttest Achievement Means for Married
and Unmarried Students

	No. of Subjects	Pretest	Posttest
Part A	18 married	20.98	33.76
	39 unmarried	19.34	32.47
Part B	18 married	65.72	100.90
	39 unmarried	67.25	102.60
Total	18 married	86.69	134.70
	39 Unmarried	86.59	135.10

Statistical Findings. A two-way analysis of variance with repeated measures was carried out. The F statistic and probability levels are given in Table 24.

Table 24

Probability of No Significant Difference Between
Tests, Test-Group Interaction, and Married and
Unmarried Groups

	Source of Variation	F	Probability
Part A	Tests	71.064	-.00*
	Test-Group	.012	.91
	Groups	.404	.53
Part B	Tests	117.319	-.00*
	Test-Group	.001	.98
	Groups	.082	.78
Total	Tests	130.602	-.00*
	Test-Group	.003	.95
	Groups	.000	.98

* $p < .01$

Data from Table 24 indicate a significant difference between tests with the posttest being higher.

Conclusion. Null hypothesis 5 is accepted at the .05 level of probability.

Testing of Hypothesis 6: Results of Influence of Work Experience on Achievement Scores

Hypothesis 6. There is no significant difference between students who have worked in the social service field and students who have not worked in the social service field as measured by pretest and posttest Part A, Part B, and Total mean change scores.

$$H_0: u_1(\text{worked}) = u_2(\text{not worked})$$

H_1 : The population means are not equal.

Means. Pre and posttest means for students who have worked in the social service field and for those who have not worked in the field are given in Table 25.

Table 25

Pre and Posttest Achievement Means for Students Who Have Worked And for Students Who Have Not Worked In The Social Service Field

	No. of Subjects	Pretest	Posttest
Part A	26 Worked	20.92	34.79
	32 Not Worked	18.87	31.55
Part B	26 Worked	64.21	104.80
	32 Not Worked	66.79	100.50
Total	26 Worked	85.12	139.50
	32 Not Worked	85.67	132.00

Statistical Findings. A two-way analysis of variance with repeated measures was carried out. The F statistic and probability levels are given in Table 26.

Table 26

Probability of No Significant Difference Between
Tests, Test-Group Interaction and Work Experience
Groups

	Source of Variation	F	Probability
Part A	Tests	87.067	.00*
	Test-Group	.177	.68
	Groups	1.583	.21
Part B	Tests	126.594	.00*
	Test-Group	1.084	.30
	Groups	.026	.87
Total	Tests	145.248	.00*
	Test-Group	.931	.34
	Groups	.271	.60

* $p < .01$

Data from Table 26 indicate a significant difference between tests with the posttest being higher.

Conclusion. Null hypothesis 6 is accepted at the .05 level of probability.

Discussion

To compare different teaching methods is a complex task. Variables are difficult to define and caution should be exercised to avoid overinterpretation of soft data. The purpose of this section is to discuss the study's results, relate them to the research reported in Chapter 1, present possible explanations for why certain results were obtained, and state any implications stemming from the results.

Comparison of Teaching Methods Using Achievement Scores

An analysis of the achievement score data revealed no significant difference between methods in effecting student gain from pre to posttest.

These results for Part A were not supported by the research literature which indicated lecture to be more effective in teaching factual knowledge and discussion more effective for teaching intellectual abilities and skills (Bloom, 1953). Since Part A concentrated on factual knowledge, those methods which were most similar to lecture should have been favored, i.e. Methods 1 and 2. One possible explanation for the nonsignificant findings may be found in the nature of the test. Part A was composed of a mixture of factual knowledge and comprehension/application questions. There were nine knowledge questions and two comprehension/application type questions on the pretest while on the posttest there were only seven

knowledge questions and four comprehension/application ones. Had there been just factual knowledge questions, the lecture method might have gained more significantly than the other methods.

Ruja (1954) also combined factual knowledge with cognitive skills on his achievement tests. He too found no significant difference when he compared discussion and lecture methods for psychology and philosophy courses. In psychology the lecture achieved more than the discussion method while in philosophy both methods achieved approximately the same. He concluded that perhaps the difference in subject matter could be important.

From this study of comparing methods of teaching family assessment, a more valid conclusion might be that achievement tests which are composed of a mixture of factual knowledge and cognitive skills type questions confound the comparison.

Another possible explanation for the nonsignificant findings relates to the use of written handout material. Students in all three methods received handouts which outlined family assessment theory. Perhaps the answers given to Part A questions were obtained not from knowledge acquired during class sessions but rather from reading written handouts. Costin (1972) warns of this problem which was one reason why Part A scores were analyzed separately from the Total score.

For Part B, a possible explanation for the nonsignificant findings between methods may be found in the nature of the test. The questions focussed on assessing students' conceptual and perceptual skills in clinical situations. For example, students were asked to comment on a family's basic verbal interaction pattern, i.e. who speaks to whom, frequency of interruptions, etc. There are two problems with these types of questions:

1. Students are cued by the question vocabulary to identify the clinical problem. Steele (1975) discussed this drawback when he advocated asking students open-ended questions.
2. Much of the vocabulary involved in describing family interaction is not unique to family assessment theory. For example, the terms "attachment" or "behavioral controls" are readily understood by the general population. Thus, no students in one teaching method were at a greater advantage than students in another method when asked to give evidence for attachment of family members to one another.

For the Total score, a possible explanation for the nonsignificant findings may be found in the nature of the teaching methods. Methods 1, 2, and 3 all focussed on teaching conceptual and perceptual skills in family assessment.

In addition, however, Method 3 had to grapple with executive skills or the actual interviewing techniques required to elicit family interaction. This was based on the hypothesis that increased student involvement with the material, i.e. family interview, would lead to increased conceptual and perceptual skills (Liddle & Halpin, Note 2). The data of this study do not support this premise.

Perhaps, had further analysis been done to distinguish high ability students from low ability ones, the high ability project students might have demonstrated greater gain than the lecture students of high ability. Ward (1956) was able to demonstrate this finding in his comparison of the lecture-demonstration method with a group study method in physical science instruction. The question, however, of which instrument to use to identify "high ability" students is a complex one.

Another possible explanation for the lack of a significantly greater gain by Method 3 students is that their preceptors had less experience with family interviewing than did the preceptors in Methods 1 and 2.

Preceptor attitude may also have influenced the results in that only 2 of 17 preceptors polled on the pre-unit questionnaire thought that students should make their own tapes during the first six weeks of the course. All thought students should prepare tapes after the first six weeks when basic principles had been taught.

Comparison of Teaching Methods Using Satisfaction Evaluations

An analysis of the satisfaction evaluations revealed significant differences between teaching methods. Method 3 students were significantly less satisfied than Method 1 and 2 students with the teaching method. Possible explanations for this dissatisfaction include: a lack of having a well outlined assessment interview structure prior to interviewing a family, the increased time involvement necessary for the interview, the difficulty of getting a family to videotape, the lack of social science background of some students, and the lack of guidance in the small group. These explanations were offered by Method 3 students to the investigator in response to a statement of interest in their subjective evaluation of the teaching method to which they were exposed and its effect on their learning. They are consistent with DiVesta's report (1953) of students' opinions of the relative productiveness of different teaching methods. He found that students evaluated lectures and case conferences as most beneficial.

A possible explanation for student dissatisfaction with the lack of guidance in the small groups can be provided by the G.A.P. report (1958) statement that discussion groups generate several hypotheses to explain what has been observed. Each group member thus comes upon new questions and alternatives. This fact of no "closure" may be interpreted by some students as a lack of guidance

by the preceptors. On the other hand, some small groups in Method 3 did have preceptor pairs who had never previously taught in the unit, Family in Health and Illness. This may possibly account for a lack of guidance.

Another factor which Method 3 students and preceptors had to contend with was the variability in their family assessment videotapes. At the preceptor meeting each week prior to the class session, the family psychiatrist would outline the core concepts to be taught that week. He would then show videotape clips to illustrate these concepts. Methods 1 and 2 used this prepared video material in their class sessions. Method 3 preceptors had to rely on student prepared video material to illustrate the concepts. Due to the variability of the student presentations, some were more germane to the topics than were others. Students may have been dissatisfied with this lack of exact harmony between tape and discussion topic.

Significant differences between teaching methods were also noted with respect to student satisfaction with the technical quality of videotapes. Method 1 students were significantly less satisfied than were Method 2 and 3 students. One possible explanation for this is that Method 1 students expected the preceptor to have polished, professional quality tapes and were less tolerant of audio-visual problems than were Method 2 and 3 students. An alternative explanation is that Method 1 actually

experienced significantly more technical problems than did the other methods. The latter explanation seems most likely based on the verbal reports from Method 1 to the investigator.

There were no significant differences between teaching methods in student satisfaction with preceptors. Students did however report what they liked and disliked about preceptors in various methods. Method 1 students appreciated the consistency of having only one preceptor in their method and therefore there was uniformity of exposure. Some students in Method 2 reported, "preceptors added a lot to the group both by their own participation and experience and also by their ability to extract the information from the students. Discussions were relaxed and yet had a lot to offer". Cohen et al.(1976) support this need for students of family therapy to participate freely in the learning setting. Goldenberg et al.(1975) also comment on the advantage of the preceptor exposing his/her own experience and abilities so that students may learn directly from this sharing. Negative comments about preceptors in Method 2 centered on their inability to stimulate discussion. One student wrote, "waiting for students to stumble on the answer the preceptor wants is a total bore. Only those opinions coinciding with the preceptor's preconceived notions were given any thought or discussion". Students in Method 3 liked the fact that

the preceptors gave them "plenty of opportunity for clarification". However one student felt that "sessions were disorganized and I felt at times we did the work and the leading".

The results of this study seemed to indicate that students were satisfied with their preceptors in that they functioned according to the role prescribed by the respective teaching method. That is, students liked what are reported in the literature as the advantages of each method. They liked that the lecturer was consistently present and they did not experience variation in preceptor. Small group discussion preceptors were liked for their ability to generate interest, provide guidance, and give constructive feedback. They were not liked if they did not generate discussion or provide opportunity for clarification, facets which are advantages of the small group discussion method.

Comparison of Teaching Methods Using Time Expenditure

Evaluations

An analysis of the time expenditure data revealed a significant difference between teaching methods in total time spent but not in study hours. Method 3 students had the greatest total time involvement. These results are consistent with student reports that making tapes is "too time consuming" and that there is "too little payoff for time involved".

It was expected there might be a difference between methods in amount of study time. That is, that Method 3 students might study significantly less because of their involvement in the practical experience of making a family assessment. The data indicate that Method 3 spent 32% less time than Method 1 and 10% less time than Method 2. This is somewhat consistent with Miller, Allender, and Wolf's findings (1965) that lecture method students report the most study time. However, the difference between the methods in this study is not significant. This may be accounted for by the unreliability of the measuring instrument. Joyce and Weatherall (1957) contend that students are highly inaccurate in reporting study time expenditure.

Relationship of Background Variables to Achievement

Influence of Sex on Achievement. An analysis of the influence of gender on student achievement scores revealed significant differences between males and females. Women were higher than men on Part A and Total scores. This can be accounted for by women's high pretest scores, which is consistent with the literature (Bardwick, 1971) which postulates that women are more expressive or concerned with affect than are men. Also, women tend to be more involved than men in child care and the helping professions. Therefore, because of their increased contact with family interaction, one could hypothesize that this accounts for their high entry level of achievement in

family assessment theory. That women are not significantly different than men on Part B is supported by Maccoby and Jacklin (1974) who report little difference between the sexes in their ability to interpret, or be more sensitive to, social cues. One possible explanation for why there is no significant difference between men and women in their gain from pre to posttest is that both groups were equally exposed to the same new material.

A further implication of the influence of sex differences in this study is that the comparison between teaching methods may be influenced by this variable. That is, there were only three women in Method 3 while there were six and ten in Methods 1 and 2 respectively. Perhaps the variable distribution of sex accounts for the low pretest score for Method 3 and the high pretest score for Method 2.

Influence of Marital Status on Achievement. An analysis of the influence of marital status on achievement scores revealed no significant difference between married and unmarried students. The premise underlying the testing of this hypothesis was that married students might be more familiar with intense interpersonal relationships and thus would perform better on an achievement test of family assessment. The data did not support this premise.

Influence of Work Experience on Achievement. An analysis of the influence of work experience on achievement scores revealed no significant difference between those students who had worked and those who had not worked in the social service field. One possible explanation for this lack of difference is that the achievement test was insensitive to recognizing students' previous work experience. It rather concentrated on a highly specific systems model of family assessment which is not in general use in the social service field.

Another possible explanation is that the question used to categorize these students was too non-specific. For example, students who volunteered in a hospital for two months were combined with those who had taught school for five years. Due to the variety of time intervals in employment, no distinction (for purposes of statistical analyses) was made between those who had worked for a short time and those who had worked for longer times.

That few significant differences emerged from the study of the influence of background variables on achievement in family assessment can be explained by the general lack of knowledge and research in this field. Bloom (1963) points out that the basic problem of dealing with specialized talents becomes one of determining which abilities are more relevant to the educational process and which are less relevant. This is a particularly difficult task

in the field of family assessment where the literature documenting the training process is sparse.

Implications

After considering the results of this study, three significant findings emerge. There is no significant difference between teaching methods in effecting student achievement gain from pre to posttest. Students in the Small Group Discussion/Student Presentation method are significantly less satisfied than other students and spend significantly more time outside of class than do Small Group Discussion/Demonstration students.

A few minor findings also emerged from this study. Lecture/Demonstration required the least amount of faculty time. Women score significantly higher than men on factual knowledge and total achievement but do not demonstrate a significantly different gain from pre to posttest.

Limitations

The following limitations are acknowledged as possibly influencing the results of this study. They will be discussed under the sub-headings:

- i. Limitations Related to the Research Problem Definition
- ii. Limitations Related to Design
- iii. Limitations Related to Methodology
- iv. Limitations Related to Instruments
- v. Limitations Related to Data Analysis

Limitations Related to Research Problem Definition.

Conclusions about the effectiveness of the three teaching methods must be made with caution. One reason for this is that the operational definition of the methods contains several variables which are difficult to quantify or measure accurately. For example, preceptors in Methods 2 and 3 were instructed to encourage critical thinking among the students. To operationally define critical thinking is very difficult. Another problem in defining the difference between methods is to clearly delineate the nature of classroom activities for different groups. In Method 1, the preceptor was to determine the activities while in Methods 2 and 3, they were to be determined with group consent. How preceptors actually achieved group consent in determining the activities themselves is difficult to explicate.

Limitations Related to Design. There are several limitations inherent in the use of this research design:

- a. No control group was used due to the ethical question involved in depriving students of teaching input. Had a control group been used, the results might have shed some light on the validity of the instruments, especially Part B.
- b. The same written handouts were given to all students. This may have equalized family assessment theory input and in fact superceded the input

derived from the teaching methods.

- c. Preceptors of varying background and expertise differed in all three teaching methods. A senior, experienced family therapist was designated, not randomly assigned, to teach Method 1 students.
- d. The methods differed in amount of faculty-student contact time which may confound the results.
- e. Limited assurance can be given that the exact teaching method specified was carried out in the small group discussion sessions. Student and preceptor questionnaire responses, however, indicate that this was the case.
- f. Contamination or between method student discussion about their teaching method's content may be a limiting factor in interpreting the results. Student questionnaire data, however, does not support the contamination issue.

Limitations Related to Methodology. There are two limitations related to methodology:

- a. The sample for the experiment was not particularly large, consisting of less than 22 subjects in each of the three teaching methods, providing greater risk for statistical inference.
- b. Not all students took the pre and posttest at the same time and this may influence the test results. However, student verbal and written statements do

not support the consideration of this limitation as a significant one. (See Appendix 3.)

Limitation Related to Instruments. Several limitations should be considered:

- a. The visual and audio quality as well as the content of the videotapes used in Part B differed from selection one to selection five. That is, some selections were longer than others; some depicted two people while others showed five family members. Some involved noisy children while others showed more subdued interaction. Some students may have been distracted by these variances in quality and content.
- b. The types of questions used in the tests may have cued students' answers.
- c. The sample for the parallelism study was especially small, consisting of only six subjects in each group in which test order was varied.
- d. The subjects for the parallelism study were not randomly assigned to the study but rather agreed to participate when contacted. Thus, there might be some selective factor that systematically differentiates this group from the normal population of students who participate in family therapy training.
- e. Parallelism could not be established for Part B.

Limitations Related to Data Analysis. There are several limitations related to data analysis which may influence conclusions about the effectiveness of the teaching methods.

- a. The variance of student test scores was especially large on the pretest.

(See Appendix 5 for frequency distributions of scores.)

Bloom (1963) asserts that the investigator must recognize that "his test may not yield an equally reliable and valid sample for all students in his study. He may be most confident of the results...when the groups approach the 50 per cent level as the mean score" (p.38). He further advocates regarding the extreme scores as less dependable than the scores nearer the mean.

That a few students did not seem to try on the pretest was evident from the limited amount of writing on their tests. One student quit taking the test midway through and wrote a note on his paper stating this fact. It was difficult to actually be sure of this with other students, however, because there were a few answers scattered throughout their tests.

- b. Some achievement scores used in the data analysis were gathered from students who did not attend every class. All students who took both the pre and posttest were included in the sample without respect to their class attendance. Campbell and Stanley (1963) support this procedure to avoid a sampling bias. One student in Method 3 did not attend any classes at all and yet achieved well. The question of the contribution of written handouts and discussion between students thus presents itself.
- c. The influence of the variable, sex difference, on each method's achievement was not analyzed. Attempts were made to carry out a three-way analysis of variance with repeated measures to control for the unequal distribution of women in each method. However, due to the small sample involved, the interpretation of the results would not have been meaningful.

Chapter 5

Summary and Conclusions

The Problem

One of the functions of the physician is to recognize the influence of the family in maintaining emotional problems in patients. It is therefore important that medical students be educated in the knowledge and skills that will enable them to assess family interaction and its influence on emotional complaints. Recently, there has been a trend in medical education to increase the psychosocial content in the medical curriculum and to teach students to assess family interaction.

If family assessment is to be taught in medical school courses, two questions should be considered:

1. how do medical educators teach students a family assessment model,
2. how do educators evaluate the effectiveness of this teaching of medical students.

At the University of Calgary, Faculty of Medicine, the first question has been approached in various ways. Several methods of instruction have been employed in the past four years. These methods have included lectures, small group discussion, role playing, and videotape demonstrations by skilled interviewers. Students have assessed families and presented a written evaluation or videotaped interview. Several of these methods are

expensive in terms of faculty time involvement, student time expenditure and audio-visual resource use. Both faculty and students have expressed varying opinions with respect to the relative efficacy of each of the teaching methods.

The attempt of this study was to answer the second question regarding evaluation of training programs. The purpose of the study was to determine the effects of three teaching methods of family assessment by comparing scores of groups of medical students. This was accomplished by obtaining achievement tests and questionnaire data for students in Method 1: Lecture/Demonstration, Method 2: Small Group Discussion/Demonstration, and Method 3: Small Group Discussion/Student Presentation.

Approval for the study was secured from the Ethics Committee, Continuity Course Committee, and the Associate Dean (Education), Faculty of Medicine.

Summary of Methodology

Fifty-eight first year University of Calgary medical students were the subjects of this study. Each student participated in six weeks of an undergraduate medical school course, Family in Health and Illness, and was randomly assigned to one of three teaching methods.

Method 1: Lecture/Demonstration was taught by one family therapist. Method 2: Small Group Discussion/Demonstra-

tion was taught by two preceptors, a family practitioner and a social scientist/family therapist, in each small group. Method 3: Small Group Discussion/Student Presentation was taught by two co-preceptors, a family physician and a social scientist/family therapist, in each small group. Methods 2 and 3 had four groups of six students. Each method utilized videotapes in a different manner. In Method 1, the preceptor demonstrated family assessment by use of faculty prepared edited tapes. In the second method, small group discussion, faculty prepared edited videotapes were also used. Preceptors actively encouraged students to observe and comment on the verbal and non-verbal evidence within the tape as well as to discuss their emotional reactions to family assessment. In the third method, students engaged in small group discussion but did not use faculty prepared tapes. Working in pairs, students interviewed a family, made a videotape of the assessment, and presented it to the discussion group.

Medical students were evaluated by pre and posttests. Three scores were reported for each test. The first score was for Part A which primarily measured knowledge. The second score was for Part B which measured intellectual skills and abilities. A Total score which combined Part A and Part B scores was also reported. Additional evaluations of satisfaction with the teaching method as

well as accounts of time expenditure, were documented on a post-unit questionnaire. Background variables, such as, sex, marital status, and previous social service work experience were also examined to determine if they influenced student achievement.

Two instruments, a pretest and posttest were devised by the investigator. Each was composed of short answer and essay restricted-type questions. Part A was related to family assessment factual knowledge while Part B was related to family problems encountered in actual clinical settings. Five edited videotape clips of family interaction formed the basis for Part B questions.

Face and content validity for the pre and posttest was established by the use of experts' opinions.

An attempt was made to establish parallelism of pre-and posttest forms. Equivalence was established for Part A and the Total test but not for Part B.

Two types of reliability were measured: 1. inter-rater reliability: coefficients ranged from .76 to .99, 2. intrarater reliability: coefficients ranged from .84 to .97.

Statistical analysis included use of two-way analysis of variance with repeated measures to test the effect of the teaching method on the dependent variable, achievement scores. One-way analysis of variance was used to

test the effect of the teaching method on satisfaction and time expenditure evaluations.

Additional data collected included two questionnaires filled out by the 17 behavioral scientists or family practitioners who were the preceptors in this study. A questionnaire was filled out prior to the beginning of the unit and at the completion of the study.

Findings

An analysis of the statistical data relative to the achievement level of students as measured by pre-to posttest gains revealed no significant difference between teaching methods.

An analysis of the data relative to the satisfaction level of students, as measured by the post-unit questionnaire, revealed significant differences between methods. Method 3 students were significantly less satisfied than were Method 1 and 2 students, with the teaching method. Method 1 students, however, were less satisfied than were Method 2 and 3 students with the technical quality of videotapes.

An analysis of the data relative to the expenditure of time by students, as measured by the post-unit questionnaire, revealed Method 3 students had the greatest total time involvement. However, no significant differences were noted between methods in accounts of student

study time.

An analysis of the statistical data relative to the influence of background variables on achievement scores, as measured by pre and posttest, revealed no significant differences between married and unmarried students, and students who had work experience in the field of social service as compared to those who did not have this experience. Significant differences did exist though between males and females. Women scored higher than men on Part A and Total achievement scores but did not gain significantly more than men from the pretest to the posttest.

Other Significant Findings

In addition to testing the stated hypotheses, the relationship between attendance and achievement was also examined. Achievement scores of medical students who attended all class sessions and those who attended 75% or less of class sessions were compared. This was irrespective of method. Data are given in Appendix 6. Results indicate students who attended all classes in their respective methods gained more on Part A and Total achievement scores. Several possible explanations for this phenomenon are possible. It may reflect the advantage of increased faculty-student contact time or it may be explained by the compulsivity and academic striving of those who attend all classes.

Other findings of some import were that 1. evaluation research on teaching family assessment in a medical school was possible, 80% of the class participated, and 2. the students in all three teaching methods reported favorably on the value of seeing videotaped family assessments. They preferred faculty prepared edited tapes rather than their own or another student's tape of a family assessment.

Conclusions

After considering the results of this study, two conclusions emerge. There is no significant difference between teaching methods in effecting student achievement gain from pre to posttest. Students in the Small Group Discussion/Student Presentation method are significantly less satisfied than other students and spend significantly more time in this method.

Recommendations

Recommendations for future direction can be considered for three categories: further studies based on this research, general suggestions for work on family assessment teaching, and practical recommendations based on the results of this study.

Further Studies. Based on this study, several suggestions can be given for further work in the area of comparing methods of teaching family assessment.

These recommendations can be made for work on research problem definition, design, methodology and instruments. Each will be considered separately.

Research Problem Definition. Method differences should be more clearly delineated. Perhaps a self-instruction group should be used.

Design. Preceptor variability could be controlled by comparing a lecture method, a small group discussion method (with the same preceptor), and a control group. However, the problem of intervening history (Campbell & Stanley, 1963) would need to be assessed. That is, how much of the preceptor's teaching of one group would be based on questions generated by the previous group. Also, the contamination problem of student cross-discussion would need to be considered. Students would have to be assured of the availability of competent remedial help.

Methodology. The evaluation of actual preceptor input could be monitored in future studies by the use of either observation by way of one-way mirrors, videotaping, or audiotaping (Bloom 1963).

Instruments. Several suggestions for instrument refinement can be given:

- a. Test validity should be examined more closely. Several, instead of two, independent judges should be consulted. If test validity is to be judged by concurrence with class session

objectives, then such objectives should be clearly included in the total unit objectives.

- b. All questions on Part A should be of the knowledge type rather than the application/comprehension type.
- c. Some test questions need to be reworded to ensure more accurate understanding by examinees. Ebel (1959) suggests this may be a problem with investigator devised tests.
- d. More strict marking criteria need to be established for the raters.
- e. One of the videotape selections from Part B could be excluded when the test is being revised. That is, questions associated with both Tapes 3 and 4 assess students' ability to evaluate basic family interaction. This is redundant and time consuming. A t-test was carried out to compare the pretest without Videotape 3 scores with the pretest without Videotape 4 scores and no significant difference was found between the means. The same procedure was carried out for the post-test and the results were similar. This would perhaps indicate that either tape selection could be eliminated without substantially altering the nature of the test results. Validity and differences in test variances would have to be assessed first, however, before this revision could be carried out.

- f. The major weakness of this study is the questionable parallelism of the measuring instruments. Whether the pre and posttests are equivalent forms of the same instrument, should be evaluated with a larger sample than that which was used in this study. The sample should not be composed of volunteers but rather should be randomly selected and assigned to specific test orders, i.e. Pre → Posttest or Post → Pretest.

Tests should be administered at the same time of day and under the same conditions.

- g. To begin to evaluate the predictive validity of the tests, i.e. their ability to identify student performance in actual family interviews, the following procedure could be carried out:

- i. Give the pretest to all students starting a block placement elective in the Family Therapy Program. Faculty of Medicine, University of Calgary.
- ii. Identify the student's strengths and weaknesses according to specific categories derived from the test. (For example, does a student know the developmental stages of the family life cycle? Does he/she answer this item correctly on Part A and Part B, the application part?)
- iii. Have the student's family therapy supervisor identify the student's strengths and weaknesses in actual clinical interviews. Use the same categories, eg. family developmental stages, as was used with the written tests.

- iv. Attempt to correlate the supervisor's evaluation with the student's test result.
- v. Administer the posttest to the student at the time of his/her completion of the elective.

General Suggestions. The area of comparing methods of teaching family assessment is a complex one. Perhaps more future time should be spent in grappling with specific issues of program definition and evaluation rather than in comparing different teaching programs.

As a result of this study, some recommendations can be given for further work in the area of family assessment teaching:

- a. Define more clearly the specific objectives and content of the material being taught.
- b. Differentiate into several levels the material being taught. That is, the following should be considered:
 - i. The entry level of the student. Psychiatric residents, for example, should not receive the same level of family assessment material as first year medical students.
 - ii. The goal attainment level of the student. Family practitioners will assess and treat families in a different manner than will family therapists. Thus, the teaching of family assessment should be geared to these differences in future practice.
 - iii. The amount of student time available to learn family assessment differs. Thus, hierarchical levels of training should be established for first and second year medical students, clinical clerks, residents, etc.

It is with the issues of conceptual framework, program design, and evaluation instruments that future study should be concerned.

Practical Recommendations. Two practical recommendations can be made as a result of this study of comparing methods of teaching family assessment:

- a. If student satisfaction and time expenditure is important, then caution should be exercised in the promotion of family assessment interviewing and videotaping by students during the first six weeks of the unit, Family in Health and Illness.
- b. If only student achievement on this family assessment test is important, then it does not matter which teaching method is employed by the Faculty of Medicine. However, if in addition to student achievement on a test, the desire is to inculcate in students an appreciation of the values associated with family medicine as well as an opportunity to expose their own biases about families, then small group discussion would seem to provide this opportunity.

Reference Notes

1. Continuity Course Objectives. Document, Faculty of Medicine, University of Calgary, Calgary, Alberta.
2. Liddle, H., & Halpin, R. Family therapy training and supervision literature: a comparative review. Manuscript submitted for publication, 1976.
3. Phillips, C. Innovations in training of marriage and family counselors. Paper presented at the annual meeting of the National Council of Family Relations, Salt Lake City, August 22, 1975.

References

- Alger, I. Integrating immediate videoplayback in family therapy. In P. Guerin (Ed.), Family therapy: theory and practice. New York: Gardner Press, 1976.
- Anderson, N. Scales and statistics: parametric and non-parametric. In E. Heermann & L. Braskamp (Eds.), Readings in statistics for the behavioral sciences. Englewood Cliffs: Prentice-Hall, 1970.
- Bardwick, J. Psychology of women. New York: Harper & Row, 1971.
- Balint, M. The structure of the training-cum-research seminars: its applications for medicine. Journal of the Royal College of General Practitioners, 1969, 17, 201-211.
- Beal, E. Current trends in the training of family therapists. American Journal of Psychiatry, 1976, 133, 137-141.
- Bloom, B. Thought processes in lectures and discussions. Journal of General Education, 1953, 7, 160-169.
- Bloom, B. Taxonomy of educational objectives. New York: Longmans, Green and Co., 1961.
- Bloom, B. Testing cognitive ability and achievement. In N. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.
- Bodin, A. Family therapy training literature: a brief guide. Family Process, 1969, 8, 272-279. (a)
- Bodin, A. Videotape applications in training family therapists. The Journal of Nervous and Mental Disease, 1969, 148, 251-261. (b)
- Bureau of Applied Social Research. Training guide on constructing questionnaires and interview schedules. Columbia University, 1948.
- Burke, C. Additive scales and statistics. In W. Gephart & R. Ingle (Eds.), Educational research. Columbus: C. E. Merrill, 1969.
- Campbell, D., & Stanley, J. Experimental and quasi-experimental designs for research on teaching. In N. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.

- Carter, G., Bandler, B. & Bakst, H. Integration of psychiatry. Journal of Medical Education, 1953, 28, 21-25.
- Cassie, J., Collins, G., & Daggett, C. The use of videotapes to improve clinical teaching. Journal of Medical Education, 1977, 52, 353-354.
- Cleghorn, J., & Levin, S. Training family therapists by setting learning objectives. American Journal of Orthopsychiatry, 1973, 43, 439-446.
- Cline, D. & Garrard, J. Evaluation of the SAID teaching program. American Journal of Psychiatry, 1973, 130, 582-585.
- Cohen, M., Gross, S., & Turner, M.. A note on a developmental model for training family therapists through group supervision. Journal of Marriage and Family Counseling, 1976, 2, 48.
- Costin, F. Lecturing versus other methods of teaching: a review of research. British Journal of Educational Technology, 1972, 3, 4-31.
- Diemer, R., & Mazzocco, D. A comparison between lecture and independent study methods of instruction in dental radiology with provision for individual differences. 1976. (ERIC Document Reproduction Service No. ED 101 130).
- DiVesta, F. Evaluation of several teaching methods by adult students. Journal of Educational Research, 1953, 46, 659-671.
- Dubin, R. & Taveggia, T. The teaching-learning paradox. Eugene, Oregon: Center for the Advanced Study of Educational Administration, 1968.
- Ebel, R. Writing the test item. In E. Lindquist (Ed.), Educational measurement. Menasha, Wisconsin: George Banta, 1959.
- Elder, S., Meckstroth, G., Nice, C., & Meyers, P. Comparison of a linear program in radiation protection with a traditional lecture presentation. Journal of Medical Education, 1964, 39, 1078-1082.
- Epstein, N., & Levin, S. Training for family therapy within a faculty of medicine. Canadian Psychiatric Association Journal, 1973, 18, 203-207.

- Ferber, A., & Mendelsohn, M. Training for family therapy. Family Process, 1969, 8, 25-32.
- Flint, A., & Rioch, M. An experiment in teaching family dynamics. American Journal of Psychiatry, 1963, 119, 940-944.
- Flomenhaft, K., & Carter, R.. Family therapy training: a statewide program for mental health centers. Hospital and Community Psychiatry, 1974, 25, 789-791.
- Forgotson, J., & Sweeney, T. Use of videotapes to teach child psychiatry to medical students. Journal of Medical Education, 1977, 52, 351-353.
- Frantz, D. A comparison of two methods of instruction with respect to instructional preferences 1976 (ERIC Document Reproduction Service No. ED 116 906).
- Gaito, J. Scale classification and statistics. In E. Heermann & L. Braskamp (Eds.), Readings in statistics for the behavioral sciences. Englewood Cliffs: Prentice-Hall, 1970.
- Glass, G., & Stanley, J. Statistical methods in education and psychology. Englewood Cliffs: Prentice Hall, 1970.
- Goldenberg, I., Stier, S., & Preston, T. The use of multiple family marathon as a teaching device. Journal of Marriage and Family Counseling, 1975, 1, 343-350.
- Group for the Advancement of Psychiatry. Small group teaching in psychiatry for medical students. New York: Publications Office, 1958.
- Haley, J. Why a mental health clinic should avoid family therapy. Journal of Marriage and Family Counseling, 1975, 1, 3-13.
- Holt, J. Comparison of lecture-demonstration with self-instruction. 1975. (ERIC Document Reproduction Service No. ED 096 277).
- James, N. Personal preference for method as a factor in learning. Journal of Educational Psychology, 1962, 53, 43-47.
- Joyce, C., & Weatherall, M. Controlled experiments in teaching. Lancet, 1957, 273, 402-407.
- Kark, S., Family and community practice in the medical

- curriculum: a clinical teaching program in social medicine.. Journal of Medical Education, 1959, 34, 905-910.
- Kolstoe, R. Introduction to statistics for the behavioral sciences. Homewood, Illinois: The Dorsey Press, 1973.
- Leton, D. An evaluation of course methods in teaching child development. Journal of Educational Research, 1961, 55, 118-121.
- Levie, W., & Dickie, K. The analysis and application of media. In R. Travers (Ed.), Second handbook of research in teaching. Chicago: Rand McNally, 1973.
- Lumsdaine, A. Commentaries on research in teaching. In C. Lee (Ed.), Improving college teaching. Washington, D. C.: American Council on Education, 1967.
- Maccoby, E., & Jacklin, C. Psychology of sex differences, Stanford: Stanford University Press, 1974.
- Maguire, T., & Hazlett, C. Reliability for the researcher. Alberta Journal of Educational Research, 1969, XV, 117-126.
- Manning, P., Abrahamson, S., & Dennis, D. Comparison of four teaching techniques: programmed text, textbook, lecture-demonstration and lecture-workshop. Journal of Medical Education, 1968, 43, 356-359.
- McKeachie, W. The improvement of instruction. Review of Educational Research, 1960, 30, 351-360.
- Messner, E., & Schmidt, D. Videotape in the training of medical students in psychiatric aspects of family medicine. International Journal of Psychiatry in Medicine, 1974, 5, 269-273.
- Metzner, R., & Bittker, T. Videotape production by medical educators: some practical considerations. Journal of Medical Education, 1973, 48, 743-751.
- Miller, G., Allender, J., & Wolf, A. Differential achievement with programmed text, teaching machine, and conventional instruction in physiology. Journal of Medical Education, 1965, 40, 817-831.
- Miller, P. & Tupin, J. Multimedia teaching of introductory psychiatry. American Journal of Psychiatry, 1972, 128, 39-43.

- Morris, V. A positive approach to the utilization of student feedback in medical education. Journal of Medical Education, 1976, 51, 541-545.
- Novak, J. An experimental comparison of a conventional and a project centered method of teaching a college general botony course. Journal of Experimental Education, 1958, 26, 217-230.
- O'Hare, C., Heinrich, A., Kirschner, N., Oberstone, A., & Ritz, M. Group training in family therapy - the students' perspective. Journal of Marriage and Family Counseling, 1975, 1, 157-162.
- Perlberg, H., Peri, J., Weinreb, M., Nitzam, E., & Shimron, J. Microteaching and videotape recordings: a new approach to improving teaching. Journal of Medical Education, 1972, 47, 43-50.
- Perlmutter, M., Loeb, D., Gumpert, G., O'Hara, F., & Higbie, I. Family diagnosis and therapy using videotape playback. American Journal of Orthopsychiatry, 1967, 37, 900-905.
- Randels, P., Kilpatrick, D., McCurdy, L., & Saunders, P. Comparison of the psychiatry learning system and traditional teaching of psychiatry. Journal of Medical Education, 1976, 51, 751-757.
- Ruja, H. Outcomes of lecture and discussion procedures in three college courses. Journal of Experimental Education, 1954, 22, 385-394.
- Sanborn, D., Pyke, H., & Sanborn, C. Videotape playback and psychotherapy: a review. Psychotherapy: Theory, Research and Practice, 1975, 12, 179-186.
- Schopler, E., Fox, R., & Cochrane, C. Teaching family dynamics to medical students. American Journal of Orthopsychiatry, 1967, 37, 906-911.
- Siegel, S. Non-parametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.
- Steele, T. Evaluation of first year medical students' ability to recognize suicidal potential. Journal of Medical Education, 1975, 50, 203-205.
- Stier, S., & Goldenberg, I. Training issues in family therapy. Journal of Marriage and Family Counseling, 1975, 1, 63-68.

- Stritter, F., Burford, H., Johnson, S., & Talbert, L. Documentation of the effectiveness of self-instructional materials. Journal of Medical Education, 1973, 48, 1129-1132.
- Thorndike, R. Reliability. In E. Lindquist (Ed.), Educational Measurement. Menasha, Wisconsin: George Banta, 1959.
- Thurnblad, R., Muslin, H., & Loesch, J. A test of clinical learning by medical students: American Journal of Psychiatry, 1973, 130, 568-570.
- Tomm, K. A family approach to emotional problems of children. Canadian Family Physician, 1973, 19, 1, 51-54.
- Van Dalen, D., & Meyer, W. Understanding Educational research: an introduction. New York: McGraw-Hill Book Co., 1962.
- Ward, J. Group-study versus lecture/demonstration method in physical science instruction for general education college students. Journal of Experimental Education, 1956, 24, 197-210.
- Welser, J., Lewis, R., & Stockton, J. Audio-tutorial vs. lecture-recitation: a comparative trial in the teaching of canine radiographic anatomy. British Journal of Medical Education, 1970, 4, 316-322.
- Weston, W. Emotional problems - a family affair. Canadian Family Physician, 1972, 18, 60-64.
- Wiersma, W. Research methods in education. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1975.
- Winer, B. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Appendix 1
Pre and Posttests

PART A: Factual Knowledge

1. According to systems theory, the whole is greater than the sum of the parts. In what way is the whole family greater than the sum of its members?
2. With respect to difference, contrast the characteristics of a symmetrical and a complementary relationship.
3. Give three techniques that you could use in interviewing a family which differ from techniques you could use in interviewing an individual.
 - i.
 - ii.
 - iii.
4. Choose one stage of the family life cycle and describe two tasks with which families normally struggle at this stage.

Stage

Task

i.

ii.

5. List the last four of Erikson's eight stages of psychosocial crises of individual development.

Stage 5:	Stage 7:
Stage 6:	Stage 8:
6. Differentiate between internal and external family structure.

7. List five basic affects. Indicate which are emergency and which are welfare emotions.

	<u>Emergency</u>	<u>Welfare</u>
i.		
ii.		
iii.		
iv.		
v.		

8. Give an example of 2 formal and 2 idiosyncratic or informal family roles.

<u>Formal</u>	<u>Informal</u>
i.	i.
ii.	ii.

9. Distinguish between instrumental, corporal, and psychological controls used in child rearing.

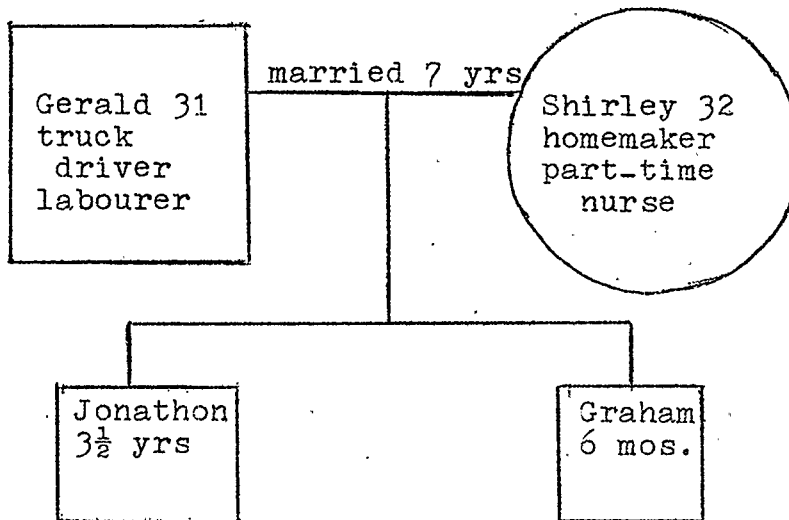
10. Define and give an example of a family myth.

PART B: Application of Assessment Concepts

The amount of data (by report or observation) that any practitioner has on a family is always limited. Try your best to answer the following questions with the limited amount of information provided.

Family #1 (5 min.)

- I. Answer the following questions (1 - 4) based on the data available from this genogram.



1. At what stage in the family life cycle is this family?
2. What family developmental tasks arise at this stage?
 - a.
 - b.
3. Using Erikson's framework, state the individual developmental stage for each family member.
4. How may the father's individual developmental task affect the accomplishment of one of the family developmental tasks at this stage?

II. Five minutes of videotaped interaction of the family diagrammed above (Family #1) will be shown once. Answer the following 5 questions based on the data available from the tape. (9 min. + tape time)

1. Note one idiosyncratic feature of this family's functioning in each of the following categories of basic interaction.
 - a. Proxemics
 - b. Kinesics
 - c. Initiators/reactors
 - d. Eye contact
 - e. Talk time
 - f. Interruptions
2. List 3 behaviors the mother used in trying to stop Jonathon from leaving the room and comment on the relative effectiveness of each.
 - i.
 - ii.
 - iii.
3. Although these family members function within a normal range of interpersonal attachment (bonding) there are differences in each dyad. Estimate the overall degree of attachment in each dyadic relationship relative to the marital dyad (which has been assigned an arbitrary "5"). Cite evidence from the tape to back up your ratings.

Dyad	Relative degree of Attachment		Evidence
	Less	More	
a. Husband-Wife	1 2 3 4 (5) 6 7 8 9		
b. Father-Jonathon	1 2 3 4 5 6 7 8 9		
c. Mother-Jonathon	1 2 3 4 5 6 7 8 9		
d. Mother-Graham	1 2 3 4 5 6 7 8 9		
e. Father-Graham	1 2 3 4 5 6 7 8 9		
f. Jonathon-Graham	1 2 3 4 5 6 7 8 9		

4. In what way did Jonathon's medical problem (left hydronephrosis diagnosed at age 6 months) affect the following relationships?
 - a. Husband-Wife
 - b. Mother-Jonathon
 - c. Father-Jonathon
 - d. Family-Medical Profession
5. If Jonathon's medical problems flared up again, what three potential behavioral or relationship problems could you anticipate in this family?
 - i.
 - ii.
 - iii.

Family #2 (6 min. + tape time)

Read the following questions (1 - 7). A videotape of family interaction will then be shown once.

- III. This family was asked to "plan something that (they) could do together". The 3 minute segment of tape shows the family discussing the task after the interviewer left the room, and serves as the basis for the next questions.
 1. Describe specific behavioral events to illustrate this family's pattern of expressive functioning in the following categories of verbal communication.
 - a. who speaks to whom
 - b. topic continuity
 - c. agreement/disagreement
 - d. commitment/evasion
 - e. content congruent with affect and behavior

- f. talk time
 - g. clear vs masked statements
2. List two affects that were conveyed during the interaction and provide supporting evidence.
 3. List two affects that were not communicated.
 4. What informal roles do the son and daughter play in this family?
 5. What kind of sibling relationship would you expect?
 6. What kind of marital relationship would you expect?
 7. Describe how the childrens' roles in the family (question 4) could contribute to the stability of the marital relationship (question 6).

Family #3 (6 min. + tape time)

IV. Another family with a similar constellation was given the same task of "plan something that (you) could do together". Answer the next 5 questions with data from the tape.

1. Describe specific behaviors to illustrate this family's functioning in the following categories.

- a. who speaks to whom
- b. relative talk time
- c. topic continuity
- d. instrumental vs affective solutions
- e. agreement/disagreement
- f. consensus sensitivity
- g. commitment/evasion

2. List two affects that were conveyed during the interaction and provide behavioral evidence to support your answer.

3. List two affects that were not communicated.

4. What kind of sibling relationship would you expect?

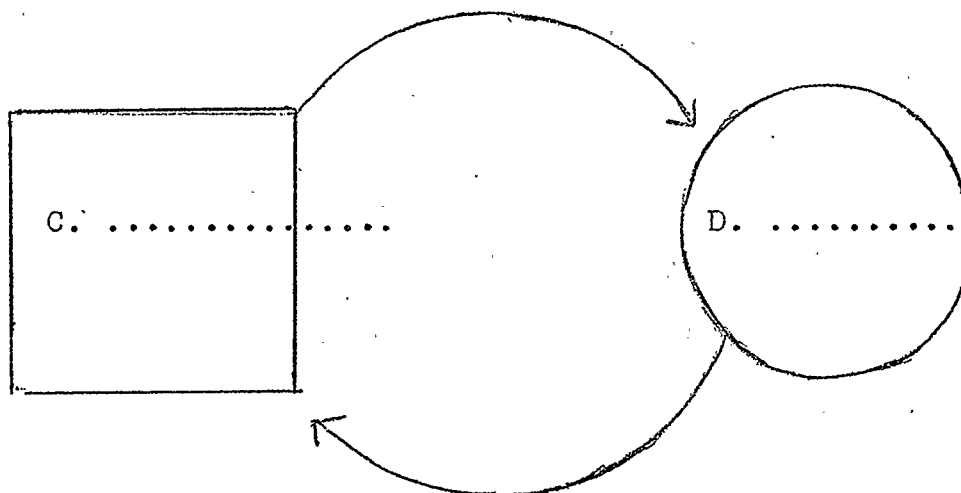
5. How would the childrens' roles in this family contribute to the stability of this marital relationship?

Family #4 (5 min. + tape time)

V. The following 5 questions are based on a two minute videotape of a therapist interviewing a married couple. The first voice heard is that of the therapist.

1. Conceptualize a possible maladaptive circular pattern of interaction between this husband and wife. Use the diagram below to document your perceived pattern by filling in the blanks A, B, C, D with single words or phrases that most aptly describe the pattern.

A.....



B.....

2. Give specific verbal and non-verbal evidence from the tape to support what you entered at A..... in the pattern above. (Do not include behaviors that would not support the specific pattern you have conceptualized)

3. Give specific verbal and non-verbal evidence from the tape to support what you entered at B..... in the pattern above. (Do not include behaviors that would not support the specific pattern you have conceptualized.)

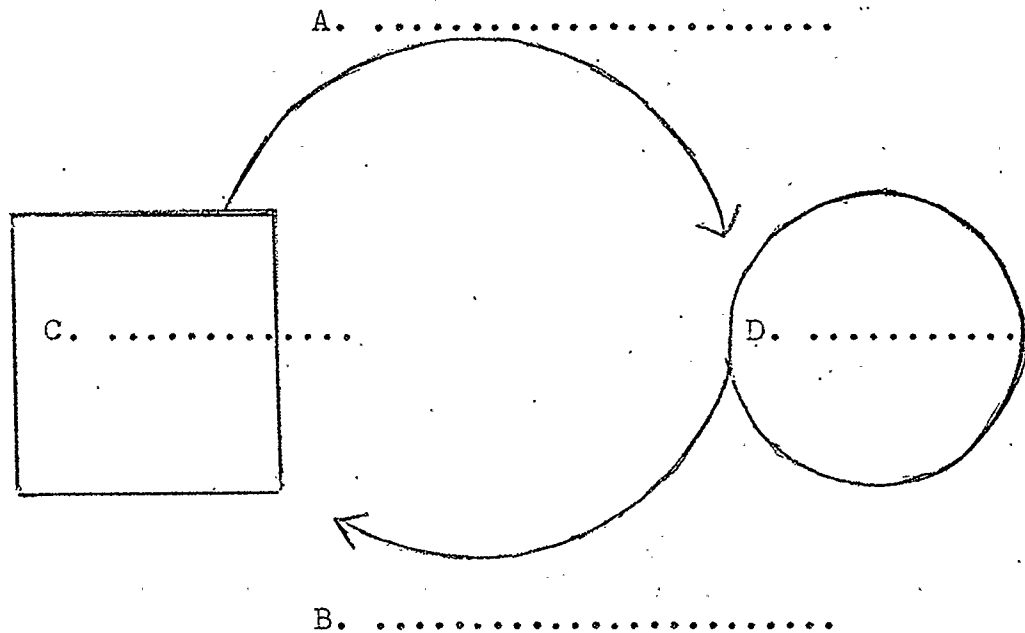
4. Give two reasons for selecting the specific word or phrase in C..... for this particular pattern.

5. Give two reasons for selecting the specific word or phrase in D..... for this particular pattern.

Family #5 (5 min. + tape time)

- VI. Read the following questions (1 - 6). A two minute videotape of a spontaneous discussion will be shown once. Answer the questions based on your observation of the tape. The tape will then be shown a second time.
1. Describe the symmetrical aspect of their interaction apparent in the content of their speech.

2. Conceptualize a possible underlying complementary pattern of interaction that is maladaptive. To document your perceived pattern, fill in the blanks A, B, C, D with single words or phrases that most aptly describe the pattern.



3. Give specific verbal and non-verbal evidence from the tape to support what you entered at A. in the pattern above. (Do not include behaviors that would not support the specific pattern you have conceptualized.)

4. Give specific verbal and non-verbal evidence from the tape to support what you entered at B. in the pattern above. (Do not include behaviors that would not support the specific pattern you have conceptualized.)
5. Give one reason for selecting the specific word or phrase in C.
6. Give one reason for selecting the specific word or phrase in D.

P O S T T E S TPART A: Factual Knowledge

1. According to systems theory, the whole is greater than the sum of the parts. In what way is the whole family greater than the sum of its members?
2. With respect to change, what is the difference between a positive and negative feedback loop? Give an example of a positive feedback loop in a family system.
3. Name two medical illnesses which may be typically aggravated/maintained by a complementary relationship.
 - i.
 - ii.
4. Choose one stage of the family life cycle and describe two tasks with which families normally struggle at this stage.

<u>Stage</u>	<u>Tasks</u>
	i.
	ii.
5. List the first four of Erikson's eight stages of psychosocial crises of individual development.

Stage 1:	Stage 3:
Stage 2:	Stage 4:
6. Define social network as it refers to family assessment.

7. Explain the difference between direct vs displaced communication in a family.

8. Give an example of 2 instrumental and 2 expressive family problems.

Instrumental

i.

ii.

Expressive

i.

ii.

9. Define the term catastrophic expectation as it refers to family functioning.

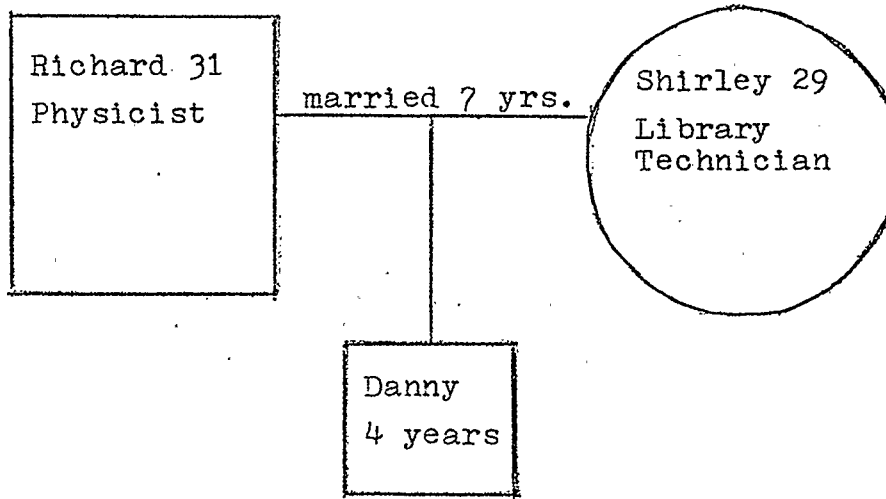
10. Differentiate between alignment and split as they refer to family functioning.

PART B: Application of Assessment Concepts

The amount of data (by report or observation) that any practitioner has on a family is always limited. Try your best to answer the following questions with the limited amount of information provided.

Family #1 (5 mins.)

- I. Answer the following questions (1 - 4) based on the data available from this genogram.



1. At what stage in the family life cycle is this family?
2. What family developmental tasks arise at this stage?
 - a.
 - b.
3. Using Erikson's framework, state the individual developmental stage for each family member.
4. How may the mother's individual developmental task affect the accomplishment of one of the family developmental tasks at this stage?

II. Six minutes of videotaped interaction of the family diagrammed above (Family #1) will be shown once. Answer the following 6 questions based on the data available from the tape. (9 min. + tape time)

1. Note one idiosyncratic feature of this family's functioning in each of the following categories of basic interaction.
 - a. Proxemics
 - b. Kinesics
 - c. Initiators/reactors
 - d. Eye contact
 - e. Talk time
 - f. Tonality
2. Give 2 examples of behavioral events (evident on tape) where family member exerts control/influence on another family member.
 - i.
 - ii.
3. Describe the difference between the husband's and wife's report of problems in terms of the following categories:
 - a. Affective problems:
 - b. Instrumental problems:
4. Although these family members function within a normal range of interpersonal attachment (bonding) there are differences in each dyad. Estimate the overall degree of attachment in each dyadic relationship relative to the marital dyad (which has been assigned an arbitrary "5"). Cite evidence from the tape to back up your ratings.

Dyad	Relative Degree of Attachment									Evidence
	Less				More					
a. Husband-- Wife	1	2	3	4	5	6	7	8	9	
b. Father-- Danny	1	2	3	4	5	6	7	8	9	
c. Mother-- Danny	1	2	3	4	5	6	7	8	9	

5. In what way did Danny's medical problem (Cleft palate) affect the following relationships?
 - a. Husband-Wife
 - b. Mother-Danny
 - c. Father-Danny
 - d. Family-Medical Profession
6. If Danny developed another medical problem, what 3 potential behavioral or relationship problems could you anticipate in this family?
 - i.
 - ii.
 - iii.

Family #2 (6 min. + tape time)

III. This family was asked to "plan something that (they) could do together". The 4 minute segment of tape showing the family discussing the task after the interviewer left the room, serves as the basis for the next questions.

1. Describe specific behavioral events to illustrate this family's pattern of expressive functioning in the following categories of verbal communication.
 - a. who speaks to whom
 - b. topic continuity
 - c. agreement/disagreement
 - d. commitment/evasion
 - e. content congruent with affect and behavior

- f. talk time
- g. interruptions

2. List two affects that were conveyed during the interaction and provide supporting evidence.
3. List two affects that were not communicated.
4. What informal role does each child play in this family?
 - a. Serena (girl):
 - b. Alex (boy seated next to girl):
 - c. boy (seated between mom and dad):
5. Describe the behavioral controls used by the mother.
6. Describe the behavioral controls used by the father.

Family #3 (6 min. + tape time)

- IV. Another family with a similar constellation was given the same task of "plan something that (you) could do together". Answer the next 5 questions with data from the tape.
1. Describe specific behaviors to illustrate this family's functioning in the following categories.
 - a. who speaks to whom
 - b. relative talk time
 - c. topic continuity
 - d. agreement/disagreement
 - e. consensus sensitivity
 - f. commitment/evasion
 2. List two affects that were conveyed during the interaction and provide behavioral evidence to support your answer.
 3. List two affects that were not communicated.
 4. What kind of sibling relationship would you expect?

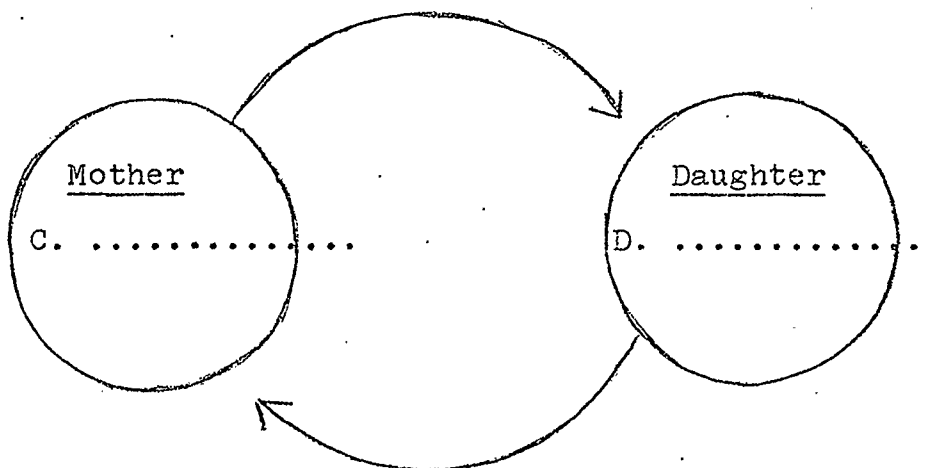
5. How would the childrens' roles in this family contribute to the stability of this marital relationship?

Family #4 (5 min. + tape time)

V. The following 5 questions are based on a two minute videotape of a therapist interviewing a mother and daughter. The first voice heard is that of the therapist.

1. Conceptualize a possible maladaptive circular pattern of interaction between this mother and daughter. Use the diagram below to document your perceived pattern by filling in the blanks A, B, C, D with single words or phrases that most aptly describe the pattern.

A.....



B.....

2. Give specific verbal and non-verbal evidence from the tape to support what you entered at A..... in the pattern above. (Do not include behaviors that would not support the specific pattern you have conceptualized)

3. Give specific non-verbal evidence from the tape to support what you entered at B. in the pattern above. (Do not include behaviors that would not support the specific pattern you have conceptualized.)

4. Give two reasons for selecting the specific word or phrase in C. for this particular pattern.

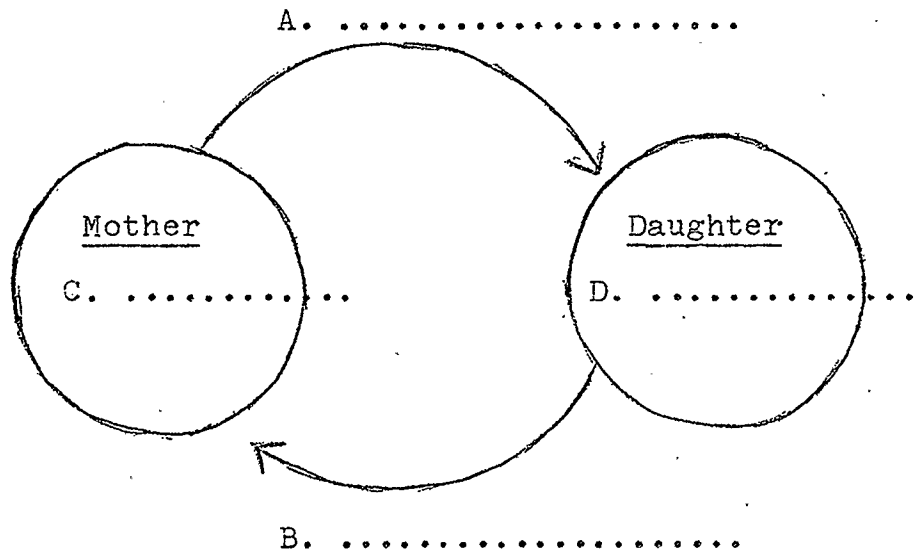
5. Give two reasons for selecting the specific word or phrase in D. for this particular pattern.

Family #5 (5 min. + tape time)

VI. A two minute videotape of a discussion between daughter, mother, stepfather, and therapist will be shown once. Answer the questions based on your observation of the tape.

1. Describe the symmetrical aspect of the interaction between mother and daughter.

2. Conceptualize the complementary pattern of interaction that is maladaptive. To document your perceived pattern, fill in the blanks A, B, C, D with single words or phrases that most aptly describe the pattern.



3. Give specific verbal and non-verbal evidence from the tape to support what you entered at A. in the pattern above. (Do not include behaviors that would not support the specific pattern you have conceptualized.)

4. Give specific verbal and non-verbal evidence from the tape to support what you entered at B. in the pattern above. (Do not include behaviors that would not support the specific pattern you have conceptualized.)

5. Give one reason for selecting the specific word or phrase in C.

6. Give one reason for selecting the specific word or phrase in D.

Appendix 2
Student Questionnaires

STUDENT NAME: _____

CODE #: (Leave blank) _____

DATE: _____

EVALUATION OF TEACHING METHODSSTUDENT PRE-UNIT QUESTIONNAIRE

The purpose of this questionnaire is to gather data on two areas:

I. Student Opinions

II. Student Background

The information provided by your answers will be treated as confidential and available only to the research personnel. Your name on the front page is to allow for coding and for my contacting you if some misunderstanding should show up on the completed form.

Information from this study will only be used in a summary form. Individual data will not be reported.

Since this data is being collected for my thesis, I would greatly appreciate your cooperation in filling out this questionnaire as completely and carefully as possible. Thank you very much.

Maureen Leahey
Graduate Student

Code #(Leave blank) _____

DATE: _____

EVALUATION OF TEACHING METHODSSTUDENT PRE-UNIT QUESTIONNAIREI. Opinions:

1. To what extent do you think intervention in the family is an important adjunct in management of the following medical problems.

Disease	No extent	Slight extent	Moderate extent	Great extent
1. 35 yr. old male with peptic ulcers	1	2	3	4
2. 14 yr. old female with Juvenile Diabetes	1	2	3	4
3. 30 yr. old female with Multiple Sclerosis	1	2	3	4
4. 40 yr. old alcoholic male	1	2	3	4
5. 55 yr. old female with depression	1	2	3	4
6. 5 yr. old hyperactive boy	1	2	3	4

2. How confident do you now feel in your ability to assess developmental, interpersonal, and emotional problems in a family?
1. _____ Not confident
 2. _____ A little confident
 3. _____ Somewhat confident
 4. _____ Very confident

3. How relevant do you think a knowledge of family dynamics will be to your future career in medicine?
1. _____ Not relevant
 2. _____ A little relevant
 3. _____ Somewhat relevant
 4. _____ Very relevant
 5. _____ I don't know what my future career is.
4. In what area of medicine do you think you may eventually practice?
1. _____ Family Practice
 2. _____ Pediatrics
 3. _____ Internal Medicine
 4. _____ Other (Please specify) _____
 5. _____ I don't know.
5. The following three teaching methods have been designated for this unit. Given a choice, by which method would you prefer to be taught family assessment? (Check one only).
1. _____ Lecture (24 students)/instructor demonstrates a family assessment by means of a prepared videotape.
 2. _____ Co-preceptors lead small group discussion (6 students)/preceptors demonstrate a family assessment by means of a prepared videotape.
 3. _____ Co-preceptors lead small group discussion (6 students)/students make and present their own videotapes of a family assessment.
 4. _____ No preference.
6. Why would you prefer to be taught family assessment by means of the method you designated in question #5?
- _____
- _____
- _____
- _____
7. When you are being taught family assessment in a small group discussion session, how much student verbal participation do you prefer to have as compared to the verbal participation by the preceptors?
1. _____ Mostly student verbal participation
 2. _____ Slightly more student than preceptor verbal participation
 3. _____ About equal verbal participation
 4. _____ Slightly more preceptor than student verbal participation
 5. _____ Mostly preceptor verbal participation

8. How would you rate your usual attendance at Continuity Course since January? (including all classes and practical labs)
1. _____ Go to less than 25% of the sessions
 2. _____ Go to 25 - 50% of the sessions
 3. _____ Go to 50 - 75% of the sessions
 4. _____ Go to 75 - 90% of the sessions
 5. _____ Go to 90 - 100% of the sessions
9. During the first six weeks of this unit, while basic principles are being taught, to what extent do you feel that the following factors will help to increase your ability to assess developmental, interpersonal, and emotional problems in a family?

	No Extent	Slight Extent	Moderate Extent	Great Extent	Uncertain
1. Readings about family assessment	1	2	3	4	5
2. Lecture	1	2	3	4	5
3. Small group Discussion	1	2	3	4	5
4. Seeing a video of a family assess- ment done by a family therapist	1	2	3	4	5
5. Seeing a video of a family assess- ment done by a medical student from a previous class	1	2	3	4	5
6. Doing a family interview myself	1	2	3	4	5

II. Background Data

It has been suggested that a person's past experience is a determinant in his/her ability to understand family functioning. I would like to examine this premise and therefore would appreciate your sharing some background data.

1. What was your age at your last birthday?
_____ years.
2. Are you:
1. _____ Male
2. _____ Female
3. Were you an adopted child?
1. _____ Yes
2. _____ No
4. Were you a foster child?
1. _____ Yes
2. _____ No
5. What is your position in your family of origin or family in which you were raised?
1. _____ Eldest child
2. _____ Second child
3. _____ Third child
4. _____ Fourth child
5. _____ Fifth child
6. _____ Other (Please specify) _____
6. What is the sex of your siblings? Indicate according to age order and include yourself.
1. Eldest: 1. _____ male 2. _____ female
2. Second child: 1. _____ male 2. _____ female
3. Third child: 1. _____ male 2. _____ female
4. Fourth child: 1. _____ male 2. _____ female
5. Fifth child: 1. _____ male 2. _____ female
More than 5 children (please continue to specify number and sex).
7. What is your present marital status?
1. _____ married
2. _____ divorced, separated
3. _____ never married
4. _____ other (Please specify) _____
8. Do you have any children?
1. _____ Yes-----How many? _____
2. _____ No

9. How many children are currently living in your household? _____
10. How many of these children fall within the following age ranges?
- | | |
|----------------------|--------------------------|
| 1. _____ 0 - 1 yr. | 5. _____ 13 - 15 yrs. |
| 2. _____ 2 - 3 yrs. | 6. _____ 16 - 18 yrs. |
| 3. _____ 4 - 6 yrs. | 7. _____ Not applicable. |
| 4. _____ 7 - 12 yrs. | |

The next questions deal with your personal background history.

1. List the number of social science courses which you have taken at university?

- | | |
|----------|------------------------------|
| 1. _____ | Psychology |
| 2. _____ | Sociology |
| 3. _____ | Anthropology |
| 4. _____ | Social Welfare |
| 5. _____ | Other (Please specify) _____ |

2. Have you ever held a paid or volunteer job involving service to people in need of help, e.g. social work, psychology, nursing, C.U.S.O., etc.?

1. ___ Yes---For how many months/years were you engaged in this social service field?
- | | |
|----------|--------|
| 1. _____ | months |
| 2. _____ | years. |
2. ___ No

Thank you very much for your cooperation. If you have any questions about the questionnaire or the research of which it is a part, I would be happy to attempt to answer them. Thank you.

Maureen Leahey
Graduate Student

Student Name: _____

Code #(leave blank) _____

Date: _____

EVALUATION OF TEACHING METHODSSTUDENT POST-UNIT QUESTIONNAIRE

The purpose of this questionnaire is to gather data on your opinion of the unit you have just finished.

The information provided by your answers will be treated as confidential and available only to the research personnel. Your name on the front page is to allow for coding and for my contacting you if some misunderstanding should show up on the completed form.

Information from this study will only be used in summary form. Individual data will not be reported.

Since this data is being collected for my thesis, I would greatly appreciate your cooperation in filling out this questionnaire as completely and carefully as possible.

Thank you very much.

Maureen Leahey
Graduate Student

Code #(leave blank) _____

Date _____

EVALUATION OF TEACHING METHODSSTUDENT POST-UNIT QUESTIONNAIREI. Opinions

1. How confident do you now feel in your ability to assess developmental, interpersonal, and emotional problems in a family?
 1. _____ Not confident at all
 2. _____ A little confident
 3. _____ Somewhat confident
 4. _____ Very confident
2. Compared to the beginning of the unit, has your confidence in making a family assessment
 1. _____ Increased
 2. _____ Decreased
 3. _____ Remained the same (If you check #3, please go to question 5)
3. If your confidence level has increased or decreased, to what extent has it changed?
 1. _____ Slight extent
 2. _____ Moderate extent
 3. _____ Great extent
4. To what do you attribute this change in confidence in your ability to make a family assessment?

5. How relevant do you think a knowledge of family dynamics will be to your future career in medicine?
 1. _____ Not relevant
 2. _____ A little relevant
 3. _____ Somewhat relevant
 4. _____ Very relevant
 5. _____ I don't know what my future career is.

6. In what area of medicine do you think you may eventually practice?
1. _____ Family Practice
 2. _____ Pediatrics
 3. _____ Internal Medicine
 4. _____ Other (Please specify) _____
 5. _____ I don't know.
7. How interested would you be in further study about the family? (excluding this unit)
1. _____ Not interested
 2. _____ A little interested
 3. _____ Somewhat interested
 4. _____ Very interested
8. To what extent do you feel that the following factors helped to increase your ability to assess developmental, interpersonal, and emotional problems in a family?

	No oppor- tunity	No Extent	Slight Extent	Moderate Extent	Great Extent	Uncer- tain
1. Readings about family assess- ment		1	2	3	4	5
2. Lecture		1	2	3	4	5
3. Small Group Discus- sion		1	2	3	4	5
4. Seeing a video of a family assess- ment done by a fam- ily ther- apist		1	2	3	4	5
5. Seeing a video of a family assess- ment done by a med- ical student from a previous class		1	2	3	4	5
6. Doing a family interview myself		1	2	3	4	5

9. How much time outside of class did you spend getting help or discussing the unit with your preceptor?
1. _____ No time
 2. _____ 0 - 15 mins.
 3. _____ 16 - 30 mins.
 4. _____ 31 - 60 mins.
 5. _____ 1 - 2 hours
 6. _____ 3 - 4 hours
 7. _____ over 4 hours
10. In a study of this kind, there is usually a fair bit of discussion between students from different teaching methods. How much time would you say you spent discussing unit content with students that were exposed to different teaching methods than your own?
1. _____ less than 30 min.
 2. _____ 30 - 60 mins.
 3. _____ 1 - 2 hours
 4. _____ 3 - 4 hours
 5. _____ over 4 hours
 6. _____ no time
11. To what extent do you think intervention in the family is an important adjunct in management of the following medical problems?

Disease	No Extent	Slight Extent	Moderate Extent	Great Extent
1. 35 yr. old male with peptic ulcers	1	2	3	4
2. 14 yr. old female with Juvenile Diabetes	1	2	3	4
3. 30 yr. old female with Multiple Sclerosis	1	2	3	4
4. 40 yr. old alcoholic male	1	2	3	4
5. 55 yr. old female with depression	1	2	3	4
6. 5 yr. old hyperactive boy	1	2	3	4

12. How satisfied are you with the following factors associated with the teaching of this unit in the first six weeks?

Factors	Not Satisfied	Slightly Satisfied	Satisfied	Very Satisfied	Comments
1. Teaching method you had i.e. lecture or discussion	1	2	3	4	
2. Preceptors	1	2	3	4	
3. Videotape technical quality	1	2	3	4	
4. Videotape content	1	2	3	4	
5. Classroom Facilities i.e. size, chairs, etc.	1	2	3	4	
6. Readings/ Handouts					
a. Family Assessment Model	1	2	3	4	
b. Expressive Family Functioning	1	2	3	4	
c. Family Structure Guide Notes	1	2	3	4	
d. Family Record Forms	1	2	3	4	
e. Attachment	1	2	3	4	
f. Circular Pattern Diagramming	1	2	3	4	

13. If you elected not to attend some of the class sessions, could you please share your reasons for this choice?

14. How would you rate the amount of time the preceptor(s) verbally participated in your class versus the amount of time the students verbally participated?
1. ☐ Mostly student verbal participation
 2. ☐ More student than preceptor verbal participation.
 3. ☐ About equal participation
 4. ☐ More preceptor than student verbal participation
 5. ☐ Mostly preceptor verbal participation

15. The next question is for students in teaching methods 1 and 2 only.

Approximately how much time did you spend studying or reading outside of class?

1. ☐ less than 1 hour
2. ☐ 1 - 2 hours
3. ☐ 3 - 4 hours
4. ☐ 5 - 6 hours
5. ☐ 7 - 8 hours
6. ☐ Other (please specify) _____

Please go to question #19

Questions 16, 17 and 18 are for those students who made their own tapes.

16. How much total time did you spend making your videotape? (Include preparation of equipment, notes, editing, and preparing presentation)
1. ☐ less than 1 hour
 2. ☐ 1 - 2 hours
 3. ☐ 3 - 4 hours
 4. ☐ 5 - 6 hours
 5. ☐ 7 - 8 hours
 6. ☐ Other (Please specify) _____
17. How much time did you spend studying or reading outside of class and outside of the time spent preparing your video presentation?
1. ☐ less than 1 hour
 2. ☐ 1 - 2 hours
 3. ☐ 3 - 4 hours
 4. ☐ 5 - 6 hours
 5. ☐ 7 - 8 hours
 6. ☐ Other (Please specify) _____

18. Do you think that you learned more about family assessment from interviewing a family/making the videotape or more from working with another student in interviewing/preparing the tape?

I learned more about family assessment from:

1. ☐ working with another student
2. ☐ Interviewing a family/making the tape
3. ☐ I learned equally from both working with another student and interviewing/making a tape.

II. Preceptor Evaluation

Circle the appropriate number to indicate whether you agree or disagree with the following items about each of your preceptors.

Name of Preceptor A _____

Name of Preceptor B _____

The preceptor		strongly disagree	disagree	agree	Strongly agree	Un-Certain
		1	2	3	4	5
19.	gave me constructive feedback on my understanding of concepts	A 1	2	3	4	5
20.	was sufficiently familiar with the unit content.	A 1	2	3	4	5
21.	explained concepts rather than repeated textbook material	B 1	2	3	4	5
22.	was a valuable learning aid.	A 1	2	3	4	5
23.	clearly stated the objectives for each session.	A 1	2	3	4	5
		B 1	2	3	4	5

III. The next question refers to your educational background.

24. List the number of half year social science courses which you have taken at university.

1. ☐ Psychology
2. ☐ Sociology
3. ☐ Anthropology
4. ☐ Other (please specify) _____

25. Up to now, this questionnaire has elicited your opinions in a structured manner. I would be interested in your subjective evaluation of the teaching method to which you were exposed and its effect on your learning.

Thank you very much for your cooperation. If you have any questions about this questionnaire, or the research of which it is a part, I would be happy to attempt to answer them. Thank you.

Maureen Leahey

Appendix 3
Test Data Collection

Test Data Collection

For various reasons (Hangover, vacation, sickness, etc.) subjects did not all take the tests at the same time. Rather, the pretest was given seven times and the posttest five times. The breakdown of number of students taking the tests on various dates and times is as follows:

Pretest

<u>Date</u>	<u>Hour</u>	<u>No. of Subjects</u>	<u>Mean Scores</u>		
			<u>Part A</u>	<u>Part B</u>	<u>Total</u>
2/23	0830	31	19.68	59.49	79.17
2/23	1430	18	20.54	68.65	89.18
2/25	0100	1	35.00	84.50	119.50
2/28	1830	1	26.40	73.00	99.40
3/1	1400	4	16.61	90.63	107.24
3/1	1930	2	20.13	72.44	92.56
3/3	1000	1	0	62.00	62.00

Posttest

4/5	1330	41	32.21	100.25	132.46
4/5	1700	5	41.15	118.50	159.65
4/5	1900	3	34.58	101.92	136.50
4/6	1045	2	37.25	120.56	157.81
4/11	1900	7	29.89	98.49	128.39

One-way analysis of variance was used to determine if there was a significant difference between the subjects who took the tests at various times. For homogeneity of variance, pretest subjects were divided into three groups of 31, 18 and 9 subjects who took the test at times 1, 2 or 3 respectively. Posttest subjects were divided into two groups of 41 and 17 students who took the test at times 1 or 2 respectively. There is no significant difference between the subjects.

Group/Time means, F levels, and probability statistics for students who took the tests at different times are listed in Table 1.

TABLE I

Probability of No Significant Difference Between
Achievement Means of Students who Took Tests at
Times 1, 2 or 3.

	Group/ Time	No. of Subjects	Mean	F	Probability
<u>Pretest</u>					
Part A	1	31	19.68	.098	.907
	2	18	20.54		
	3	9	18.68		
Part B	1	31	59.49	2.333	.107
	2	18	68.65		
	3	9	80.76		
Total	1	31	79.17	1.347	.269
	2	18	89.18		
	3	9	99.44		
<u>Posttest</u>					
Part A	1	41	32.21	1.104	.298
	2	17	34.90		
Part B	1	41	100.25	2.017	.161
	2	17	107.58		
Total	1	41	132.46	2.201	.144
	2	17	142.48		

Appendix 4

Raw Data

Background Data and Student Scores

Column	Key	1	2	3	4	5	6	7	8	9	10	11	12
		CASE-NO	ID	SEX	MARRIAGE	WORKEXP	HRATTEND	PREA	PREB	PRETOT	POSTA	POSTB	POSTTOT
1	Case #	1	2.	1.	1.	1.	6.	37.	113.	150.	35.	115.	150.
		2	7.	1.	3.	1.	6.	42.	116.	157.	37.	128.	165.
		3	34.	1.	1.	1.	6.	5.	5.	9.	34.	64.	98.
		4	46.	1.	3.	1.	6.	7.	48.	55.	32.	120.	152.
		5	47.	1.	1.	2.	6.	19.	48.	67.	38.	75.	113.
		6	60.	1.	3.	2.	6.	8.	0	8.	21.	91.	112.
2	Subject ID	7	62.	1.	1.	1.	6.	0	62.	62.	21.	97.	118.
		8	71.	1.	3.	1.	6.	26.	73.	99.	21.	94.	114.
		9	25.	1.	3.	1.	6.	13.	71.	84.	30.	111.	141.
		10	37.	1.	1.	2.	6.	9.	66.	75.	33.	80.	113.
3	Sex*	11	38.	1.	3.	2.	6.	14.	64.	77.	41.	108.	149.
		12	59.	1.	3.	2.	6.	17.	86.	103.	39.	121.	160.
		13	63.	1.	3.	1.	6.	26.	77.	103.	41.	124.	165.
		14	70.	1.	1.	1.	6.	16.	49.	64.	46.	122.	168.
4	Marital Status**	15	8.	2.	1.	2.	6.	24.	85.	109.	41.	122.	163.
		16	30.	2.	3.	2.	6.	38.	94.	132.	43.	120.	163.
		17	31.	2.	1.	1.	6.	43.	86.	128.	45.	122.	167.
5	Previous Social	18	49.	2.	3.	1.	6.	30.	79.	109.	44.	140.	184.
	Service Work Exper-	19	61.	2.	1.	1.	6.	11.	40.	51.	37.	118.	155.
	ience***	20	69.	2.	3.	1.	6.	18.	48.	65.	36.	98.	133.
		21	57.	1.	3.	2.	6.	21.	69.	89.	34.	99.	133.
		22	10.	1.	3.	2.	6.	25.	86.	111.	30.	91.	121.
		23	54.	1.	3.	2.	6.	22.	67.	89.	26.	106.	132.
		24	24.	1.	3.	2.	6.	11.	70.	81.	6.	74.	80.
6	Hours Students	25	64.	1.	3.	1.	6.	24.	61.	85.	44.	121.	165.
	Attended Class	26	65.	1.	3.	2.	6.	34.	102.	136.	33.	122.	155.
		27	23.	1.	3.	2.	6.	18.	80.	97.	30.	119.	149.
		28	64.	1.	3.	2.	6.	9.	92.	101.	40.	106.	146.
		29	50.	1.	3.	2.	6.	13.	47.	60.	36.	117.	153.
	Test Scores****	30	4.	1.	3.	1.	6.	19.	79.	97.	32.	86.	118.
		31	13.	1.	3.	1.	6.	17.	39.	56.	27.	81.	128.
		32	9.	2.	3.	1.	6.	24.	100.	124.	26.	112.	137.
		33	32.	2.	1.	1.	6.	35.	85.	120.	22.	87.	109.
7	Pretest Part A	34	53.	2.	3.	2.	6.	19.	36.	55.	30.	81.	111.
		35	17.	2.	1.	2.	6.	29.	69.	98.	50.	94.	144.
		36	29.	2.	3.	1.	6.	35.	99.	134.	47.	108.	156.
		37	39.	2.	3.	2.	6.	37.	96.	133.	37.	103.	140.
8	Pretest Part B	38	40.	2.	3.	2.	6.	13.	88.	101.	33.	115.	148.
		39	3.	2.	4.	2.	6.	16.	1.	17.	40.	120.	160.
		40	26.	2.	1.	2.	6.	36.	108.	142.	40.	126.	165.
		41	45.	2.	3.	1.	6.	24.	56.	80.	31.	102.	132.
9	Pretest Total	42	27.	1.	3.	1.	6.	6.	13.	19.	26.	92.	118.
		43	48.	1.	1.	1.	6.	21.	20.	41.	35.	81.	115.
		44	51.	1.	1.	2.	6.	20.	81.	100.	29.	116.	145.
		45	6.	1.	1.	2.	6.	18.	55.	73.	14.	105.	119.
10	Posttest Part A	46	35.	1.	3.	2.	6.	16.	54.	70.	34.	93.	127.
		47	74.	1.	1.	2.	6.	16.	81.	97.	38.	103.	140.
		48	36.	1.	3.	2.	6.	3.	23.	26.	18.	62.	80.
11	Posttest Part B	49	55.	1.	3.	2.	6.	33.	76.	109.	38.	119.	157.
		50	12.	1.	1.	2.	6.	25.	84.	110.	22.	91.	113.
		51	43.	1.	1.	1.	6.	17.	49.	65.	33.	99.	132.
12	Posttest Total	52	67.	1.	3.	1.	6.	11.	110.	121.	36.	115.	151.
		53	19.	1.	3.	2.	6.	23.	65.	88.	29.	83.	112.
		54	44.	1.	3.	2.	6.	13.	57.	70.	18.	66.	84.
		55	5.	1.	1.	2.	6.	5.	46.	50.	23.	89.	112.
		56	48.	2.	3.	2.	6.	6.	64.	70.	29.	102.	131.
		57	22.	2.	3.	1.	6.	16.	35.	41.	36.	109.	145.
		58	66.	2.	3.	1.	6.	23.	60.	83.	36.	79.	115.

*1=Male,2=Female

**1=Married,3=Unmarried,4=Other

***1=Yes,2=No

****Scores carried to 3 decimal places.

Time of Test Taking

		<u>Key</u>				
		1	2	3	4	5
		CASE-NO	CODENUMB	METHOD	TOOKTEST	TOOKPOST
<u>Column</u>	<u>Item</u>					
1	Case #	1	2.	1.	1.	4.
		2	7.	1.	5.	1.
		3	34.	1.	1.	1.
		4	46.	1.	1.	1.
		5	47.	1.	2.	1.
2	Subject ID	6	60.	1.	1.	5.
		7	62.	1.	7.	3.
		8	71.	1.	4.	1.
		9	25.	1.	1.	1.
		10	37.	1.	1.	1.
3	Method	11	38.	1.	1.	2.
		12	59.	1.	2.	1.
		13	53.	1.	1.	2.
		14	70.	1.	1.	2.
		15	8.	1.	1.	1.
4	Time Subject Took Pretest*	16	30.	1.	1.	1.
		17	31.	1.	1.	1.
		18	49.	1.	2.	1.
		19	61.	1.	1.	1.
		20	69.	1.	1.	1.
5	Time Subject Took Posttest*	21	57.	2.	2.	1.
		22	10.	2.	2.	5.
		23	54.	2.	1.	1.
		24	24.	2.	2.	1.
		25	68.	2.	1.	1.
		26	65.	2.	2.	5.
		27	23.	2.	0.	1.
		28	64.	2.	5.	1.
		29	50.	2.	1.	5.
		30	4.	2.	1.	1.
		31	33.	2.	1.	5.
		32	9.	2.	2.	1.
		33	32.	2.	3.	1.
		34	53.	2.	2.	1.
		35	17.	2.	1.	3.
		36	29.	2.	2.	1.
		37	39.	2.	2.	1.
		38	40.	2.	2.	3.
		39	3.	2.	2.	2.
		40	26.	2.	1.	4.
		41	45.	2.	1.	1.
		42	27.	3.	1.	1.
		43	48.	3.	2.	1.
		44	51.	3.	1.	1.
		45	6.	3.	1.	5.
		46	35.	3.	2.	1.
		47	74.	3.	2.	1.
		48	36.	3.	1.	1.
		49	55.	3.	1.	2.
		50	12.	3.	1.	1.
		51	43.	3.	1.	1.
		52	67.	3.	5.	1.
		53	19.	3.	6.	5.
		54	44.	3.	2.	1.
		55	5.	3.	5.	1.
		56	58.	3.	2.	1.
		57	22.	3.	1.	1.
		58	66.	3.	1.	1.

* Actual dates and times that subjects
took the tests are given in Appendix 3.

Satisfaction Evaluations and Time Expenditure

Column	Key Item	1 CASE-NO	2 ID	3 METHOD	4 METHODSAT	5 PRECSAT	6 TAPE TECH	7 TAPECONT	8 STUDY	9 VIDEO	10 THIRSTUDY
1	Case #	1	2.	1.	3.	3.	1.	3.	6.	-0	-0
		2	7.	1.	3.	3.	2.	2.	3.	-0	-0
		3	14.	1.	3.	3.	1.	3.	2.	-0	-0
		4	46.	1.	4.	4.	3.	4.	3.	-0	-0
		5	47.	1.	2.	2.	2.	2.	2.	-0	-0
		6	60.	1.	3.	2.	3.	4.	1.	-0	-0
2	Subject ID	7	62.	1.	1.	3.	1.	2.	2.	-0	-0
		8	71.	1.	3.	4.	3.	3.	1.	-0	-0
		9	25.	1.	4.	3.	3.	3.	3.	-0	-0
		10	37.	1.	2.	3.	2.	3.	3.	-0	-0
3	Method	11	38.	1.	4.	4.	2.	3.	2.	-0	-0
		12	54.	1.	3.	3.	2.	3.	5.	-0	-0
		13	63.	1.	3.	4.	3.	3.	2.	-0	-0
	Satisfaction Evaluations*	14	70.	1.	2.	3.	1.	3.	3.	-0	-0
		15	8.	1.	4.	4.	3.	4.	1.	-0	-0
		16	30.	1.	4.	4.	2.	3.	2.	-0	-0
		17	31.	1.	4.	4.	1.	3.	2.	-0	-0
4	Method	18	49.	1.	3.	4.	4.	4.	3.	-0	-0
		19	61.	1.	3.	3.	1.	3.	2.	-0	-0
		20	69.	1.	4.	4.	3.	3.	3.	-0	-0
5	Preceptors	21	57.	2.	3.	3.	3.	1.	1.	-0	-0
		22	10.	2.	3.	3.	2.	3.	1.	-0	-0
		23	54.	2.	3.	4.	4.	4.	2.	-0	-0
		24	24.	2.	3.	3.	4.	4.	1.	-0	-0
6	Tape Technical Quality	25	68.	2.	3.	3.	2.	3.	3.	-0	-0
		26	65.	2.	3.	3.	3.	3.	3.	-0	-0
		27	23.	2.	3.	4.	4.	4.	5.	-0	-0
		28	64.	2.	4.	4.	3.	3.	2.	-0	-0
7	Tape Content	29	50.	2.	3.	4.	2.	3.	3.	-0	-0
		30	4.	2.	3.	4.	3.	3.	1.	-0	-0
		31	33.	2.	4.	3.	4.	3.	2.	-0	-0
8	Study Time**	32	9.	2.	1.	2.	3.	1.	1.	-0	-0
		33	32.	2.	4.	4.	3.	4.	1.	-0	-0
		34	53.	2.	3.	2.	3.	3.	1.	-0	-0
		35	17.	2.	3.	4.	2.	3.	3.	-0	-0
9	Video Preparation Time**	36	29.	2.	3.	3.	1.	2.	1.	-0	-0
		37	39.	2.	2.	2.	1.	2.	1.	-0	-0
		38	40.	2.	4.	4.	3.	3.	2.	-0	-0
		39	3.	2.	3.	4.	4.	3.	3.	-0	-0
10	Study Time**	40	26.	2.	2.	2.	4.	3.	2.	-0	-0
		41	45.	2.	4.	3.	3.	3.	2.	-0	-0
		42	27.	3.	3.	3.	3.	3.	3.	-0	-0
		43	48.	3.	3.	4.	4.	4.	4.	-0	-0
		44	51.	3.	3.	4.	3.	3.	3.	-0	-0
		45	6.	3.	3.	3.	2.	2.	2.	-0	-0
		46	35.	3.	3.	3.	4.	3.	3.	-0	-0
		47	74.	3.	2.	3.	4.	3.	3.	-0	-0
		48	36.	3.	1.	3.	2.	-0	-0	-0	-0
		49	55.	3.	4.	4.	3.	3.	3.	-0	-0
		50	12.	3.	2.	3.	3.	2.	2.	-0	-0
		51	43.	3.	1.	2.	-0	-0	3.	-0	-0
		52	67.	3.	2.	2.	-0	-0	-0	-0	-0
		53	19.	3.	3.	3.	3.	3.	2.	-0	-0
		54	44.	3.	3.	3.	-0	-0	3.	-0	-0
		55	5.	3.	1.	3.	3.	3.	1.	-0	-0
		56	58.	3.	1.	1.	1.	1.	2.	-0	-0
		57	22.	3.	3.	4.	3.	3.	3.	-0	-0
		58	66.	3.	2.	4.	3.	2.	5.	-0	-0

*1=Not Satisfied,2=Slightly Satisfied

3=Satisfied,4=Very Satisfied

**1= 1Hr.,2=1-2Hrs.,3=3-4Hrs.,

5=7-8Hrs.,6=Other

Parallelism Study Scores *

1	2	3	4	5	6	7	8	9
CASE-NO	ID	ORDER	PREA	PREB	PRETOT	POSTA	POSTB	POSTTOT
1	1.	1.	50.	126.	176.	47.	132.	179.
2	2.	1.	38.	136.	174.	37.	124.	161.
3	3.	1.	40.	55.	94.	32.	87.	118.
4	4.	1.	36.	110.	146.	46.	119.	164.
5	5.	1.	29.	85.	113.	41.	94.	135.
6	6.	1.	42.	129.	171.	34.	121.	155.
7	7.	2.	32.	106.	138.	42.	115.	156.
8	8.	2.	39.	141.	180.	39.	126.	165.
9	9.	2.	38.	136.	174.	34.	119.	153.
10	10.	2.	42.	124.	166.	42.	107.	149.
11	11.	2.	43.	118.	160.	37.	112.	149.
12	12.	2.	40.	95.	135.	32.	78.	110.

Key

Column

Item

1	Case #
2	Subject ID
3	Test Order 1 = Pre→Post, 2 = Post→Pre
4	Pretest Part A
5	Pretest Part B
6	Pretest Total
7	Posttest Part A
8	Posttest Part B
9	Posttest Total

* Scores used in actual data analysis were carried to three decimals.

Interrater Reliability

Key		1	2	3	4	5	6	7	8	9
Column	Item	CASE-#	CODENUM	METHOD	KSPREA	HSPREA	KSPREGEN	HSPREGEN	KSPREONE	HSPREONE
1	Case #	1	2.	1.	36.	37.	4.	8.	31.	30.
		2	7.	1.	43.	41.	4.	9.	36.	33.
		3	34.	1.	5.	5.	4.	4.	2.	0
		4	46.	1.	7.	7.	0	0	12.	12.
		5	47.	1.	16.	19.	0	0	12.	15.
		6	50.	1.	8.	8.	0	0	0	0
		7	52.	1.	0	0	0.	4.	29.	29.
2	Subject ID	8	71.	1.	25.	24.	0.	7.	19.	19.
		9	25.	1.	14.	12.	4.	3.	25.	25.
		10	37.	1.	4.	4.	3.	2.	17.	16.
3	Method	11	34.	1.	14.	14.	2.	2.	17.	19.
		12	59.	1.	15.	14.	7.	4.	23.	28.
		13	63.	1.	20.	27.	2.	2.	28.	26.
		14	70.	1.	15.	17.	2.	3.	15.	14.
	Pretest Scores	15	8.	1.	25.	23.	7.	7.	22.	21.
		16	30.	1.	35.	41.	0.	8.	34.	33.
		17	31.	1.	45.	41.	4.	6.	29.	32.
4	Rater 1,Part A	18	49.	1.	29.	31.	10.	10.	24.	28.
		19	61.	1.	12.	11.	0	0	0	0
		20	69.	1.	16.	19.	0.	0.	13.	16.
		21	57.	2.	14.	23.	0.	5.	11.	9.
5	Rater 2,PartA	22	10.	2.	24.	26.	7.	0.	30.	28.
		23	54.	2.	22.	23.	4.	4.	23.	21.
		24	24.	2.	8.	13.	2.	2.	39.	39.
		25	58.	2.	23.	25.	0	0	23.	24.
6	Rater 1,Genogr	26	65.	2.	32.	37.	10.	11.	25.	26.
		27	23.	2.	14.	17.	0.	7.	19.	19.
		28	64.	2.	10.	8.	7.	7.	23.	23.
		29	50.	2.	13.	13.	5.	4.	12.	13.
7	Rater 2,Genogr	30	4.	2.	18.	19.	4.	4.	14.	13.
		31	33.	2.	14.	20.	3.	3.	15.	14.
		32	9.	2.	23.	26.	4.	2.	31.	33.
8	Rater 1,Tape 1	33	32.	2.	34.	36.	0	0	26.	24.
		34	53.	2.	20.	19.	4.	4.	14.	11.
		35	17.	2.	28.	31.	5.	5.	24.	25.
		36	24.	2.	35.	35.	12.	10.	23.	23.
9	Rater 2,Tape 1	37	39.	2.	34.	39.	4.	10.	31.	31.
		38	40.	2.	14.	12.	10.	11.	31.	31.
		39	3.	2.	17.	15.	0	0	2.	0
		40	25.	2.	32.	37.	0.	7.	29.	29.
		41	45.	2.	25.	23.	2.	2.	26.	25.
		42	27.	3.	3.	9.	2.	2.	8.	8.
		43	48.	3.	21.	20.	1.	1.	0	0
		44	51.	3.	21.	18.	5.	4.	24.	26.
		45	6.	3.	10.	18.	0.	0.	18.	20.
		46	35.	3.	15.	14.	5.	4.	16.	15.
		47	74.	3.	14.	17.	7.	5.	25.	29.
		48	36.	3.	3.	3.	0	0	10.	10.
		49	55.	3.	29.	36.	7.	7.	28.	27.
		50	12.	3.	24.	27.	5.	4.	35.	39.
		51	43.	3.	17.	17.	1.	0	17.	17.
		52	67.	3.	11.	11.	4.	3.	29.	33.
		53	19.	3.	23.	23.	7.	0.	20.	19.
		54	44.	2.	11.	15.	5.	0.	17.	18.
		55	5.	3.	5.	5.	1.	3.	6.	9.
		56	58.	3.	6.	6.	4.	4.	19.	20.
		57	22.	3.	16.	15.	2.	2.	0	0
		58	66.	3.	22.	25.	1.	1.	20.	21.

* Scores used in actual data analysis were carried to three decimals.

Interrater Reliability

Key		1	2	3	4	5	6	7	8	9	10	11
Column	Item	CASE-NO	CODE-NO	METHOD	PRETEST SCORE	MSORFIDR	MSORFIDR	MSORFIDR	MSORFIDR	KMRPRETW	KMRPRETH	KMRPREFO
1	Case #	1	2.	1.	23.	20.	10.	15.	-0	20.	-0	15.
		2	7.	1.	20.	21.	19.	14.	17.	20.	-0	18.
		3	34.	1.	0	0	0	0	-0	-0	-0	-0
		4	46.	1.	11.	13.	11.	2.	12.	-0	11.	-0
		5	47.	1.	9.	13.	7.	7.	-0	-0	9.	-0
		6	60.	1.	0	0	0	0	0	-0	-0	-0
		7	62.	1.	4.	7.	10.	3.	-0	-0	-0	-0
2	Subject ID	8	71.	1.	17.	16.	13.	2.	-0	-0	-0	-0
		9	25.	1.	12.	13.	8.	10.	-0	-0	9.	15.
		10	37.	1.	12.	14.	9.	13.	-0	-0	-0	-0
		11	38.	1.	11.	13.	9.	15.	-0	-0	-0	-0
3	Method	12	54.	1.	14.	15.	14.	13.	-0	-0	-0	-0
		13	53.	1.	12.	17.	7.	8.	-0	-0	-0	-0
		14	70.	1.	12.	13.	6.	2.	-0	-0	-0	-0
	Pretest Scores*	15	8.	1.	11.	11.	14.	22.	11.	-0	-0	-0
	Rater	16	30.	1.	13.	16.	13.	13.	-0	-0	16.	-0
	Tape	17	31.	1.	21.	9.	11.	8.	-0	-0	-0	-0
		18	49.	1.	20.	13.	5.	7.	-0	-0	8.	-0
		19	61.	1.	9.	9.	8.	14.	-0	-0	-0	-0
		20	64.	1.	4.	12.	5.	2.	8.	-0	-0	-0
4	2	21	57.	2.	16.	10.	14.	14.	-0	-0	-0	-0
		22	10.	2.	11.	11.	9.	19.	-0	-0	-0	-0
		23	54.	2.	9.	19.	12.	1.	-0	-0	-0	-0
		24	24.	2.	13.	7.	7.	3.	-0	-0	6.	11.
5	2	25	68.	2.	13.	16.	2.	7.	-0	-0	-0	-0
		26	65.	2.	12.	16.	10.	22.	11.	-0	-0	-0
		27	23.	2.	18.	11.	9.	17.	-0	-0	-0	15.
		28	64.	2.	22.	18.	13.	10.	-0	-0	-0	-0
6	2	29	50.	2.	10.	7.	2.	12.	-0	-0	7.	-0
		30	4.	2.	17.	16.	15.	14.	-0	14.	-0	19.
		31	33.	2.	4.	5.	13.	0	-0	-0	-0	-0
7	2	32	9.	2.	13.	15.	14.	18.	16.	-0	-0	21.
		33	32.	2.	21.	19.	14.	7.	-0	-0	-0	-0
		34	53.	2.	9.	6.	5.	1.	-0	-0	-0	-0
		35	17.	2.	14.	14.	4.	8.	-0	-0	-0	-0
8	1	36	29.	2.	19.	14.	19.	14.	-0	-0	14.	-0
		37	39.	2.	17.	10.	11.	19.	-0	-0	-0	-0
		38	40.	2.	18.	15.	10.	4.	-0	-0	-0	-0
		39	3.	2.	0	0	0	0	-0	0	-0	0.
9	1	40	26.	2.	19.	15.	22.	17.	17.	-0	20.	23.
		41	45.	2.	12.	7.	9.	1.	-0	-0	-0	-0
		42	27.	3.	0	2.	1.	1.	-0	-0	0	-0
10	1	43	48.	1.	9.	10.	1.	0	-0	-0	2.	-0
		44	51.	3.	15.	21.	14.	2.	-0	-0	-0	-0
		45	6.	1.	4.	7.	5.	12.	2.	6.	-0	10.
		46	35.	3.	11.	9.	2.	12.	9.	-0	-0	-0
11	1	47	74.	3.	9.	10.	13.	17.	-0	-0	-0	-0
		48	36.	3.	0	8.	0	5.	-0	-0	-0	-0
		49	55.	3.	12.	13.	4.	14.	-0	-0	-0	-0
		50	12.	1.	9.	10.	14.	11.	8.	-0	-0	-0
		51	43.	3.	7.	10.	10.	6.	-0	-0	-0	-0
		52	67.	1.	22.	17.	14.	23.	-0	-0	-0	-0
		53	19.	3.	13.	14.	6.	8.	-0	-0	-0	-0
		54	44.	3.	14.	11.	0	11.	12.	-0	-0	-0
		55	5.	3.	7.	11.	7.	10.	-0	9.	-0	16.
		56	54.	3.	7.	7.	12.	15.	-0	-0	-0	-0
		57	22.	3.	2.	3.	10.	19.	-0	-0	-0	20.
		58	66.	1.	9.	4.	13.	10.	-0	-0	-0	-0

* Scores used in actual data analysis were carried to three decimal places.

Interrater Reliability

Key		1	2	3	4	5	6	7	8	9
		CASE-NO	CODENUM	METHOD	KSPPOSTA	MSPPOSTA	KSPPOSTEN	MSPPOSTEN	KSPPOSTONE	MSPPOSTONE
Column	Item	1	2	1	35	36	8	12	32	30
1	Case #	2	7	1	34	37	13	13	43	43
		3	34	1	32	36	10	10	17	19
		4	46	1	32	32	4	8	20	23
		5	47	1	37	39	0	4	25	28
		6	60	1	21	22	5	5	25	27
2	Subject ID	7	62	1	21	20	5	3	-0	28
		8	71	1	21	20	5	8	-0	26
		9	25	1	29	31	4	8	30	30
		10	37	1	35	31	10	10	23	23
		11	38	1	43	59	12	12	37	32
3	Method	12	59	1	46	34	12	12	37	34
		13	73	1	40	43	13	11	-0	37
		14	70	1	47	44	12	12	-0	36
		15	4	1	39	42	4	9	37	38
		16	30	1	43	42	7	6	40	37
4	Rater 1,Part A	17	31	1	46	44	13	14	36	38
		18	49	1	46	42	11	13	36	39
		19	61	1	34	44	11	11	34	35
		20	69	1	37	34	4	9	-0	23
		21	57	2	30	38	10	8	32	32
5	Rater 2,Part A	22	10	2	36	31	0	10	22	24
		23	54	2	27	25	4	7	31	29
		24	24	2	4	4	0	0	36	36
		25	48	2	40	43	7	7	-0	40
		26	65	2	32	35	13	12	-0	37
6	Rater 1,Genogram	27	23	2	30	30	11	12	35	36
		28	64	2	43	38	7	7	-0	26
		29	50	2	36	34	11	10	35	38
		30	4	2	32	32	5	4	25	26
		31	33	2	45	44	7	9	14	19
7	Rater 2,Genogram	32	9	2	25	25	12	12	29	31
		33	32	2	22	21	7	7	26	27
		34	53	2	31	30	4	3	23	23
		35	17	2	50	51	12	12	33	34
		36	29	2	47	44	4	10	39	39
8	Rater 1,Tape 1	37	39	2	36	39	11	13	33	31
		38	40	2	35	32	7	10	38	37
		39	3	2	40	40	10	10	38	38
		40	26	2	42	37	13	13	30	30
		41	45	2	27	34	11	6	29	28
9	Rater 2,Tape 2	42	27	3	27	24	12	11	19	17
		43	43	3	34	35	11	14	17	20
		44	51	2	27	31	11	10	40	43
		45	6	3	15	12	2	3	37	37
		46	35	3	35	34	11	11	35	35
		47	74	3	37	34	13	12	-0	38
		48	36	3	14	14	4	4	36	37
		49	55	3	37	39	14	13	35	38
		50	12	3	23	20	3	0	24	27
		51	43	3	33	32	4	8	28	27
		52	67	3	38	35	4	4	-0	35
		53	19	3	28	31	4	3	28	29
		54	44	3	14	14	4	5	23	24
		55	5	3	22	23	4	10	27	26
		56	53	3	30	24	7	4	25	22
		57	22	3	37	36	13	14	31	31
		58	77	3	35	37	7	10	-0	20

Interrater Reliability

Key			1	2	3	4	5	6	7	8	9	10	11
Column	Item		CASE-NO	CODENUM	METHOD	POSTTEST	POSTTEST	POSTTEST	POSTTEST	KMRPOTW	KMRPOTW	KMRPOTW	KMRPOTW
1	Case #		1	2.	1.	18.	18.	22.	17.	23.	17.	-0	16.
2	Subject ID		2	7.	1.	22.	18.	19.	15.	-0	-0	-0	-0
3	Method		3	34.	1.	9.	1.	22.	5.	-0	-0	7.	-0
4	2	2	4	46.	1.	18.	21.	20.	26.	-0	-0	-0	-0
5	2	3	5	47.	1.	9.	11.	12.	12.	-0	-0	-0	-0
6	2	4	6	60.	1.	12.	12.	14.	17.	-0	-0	-0	-0
7	2	5	7	62.	1.	14.	4.	22.	21.	-0	-0	-0	-0
8	1	4	8	71.	1.	13.	18.	17.	14.	-0	-0	-0	-0
9	1	5	9	25.	1.	14.	15.	25.	22.	-0	-0	17.	-0
10	1	2	10	37.	1.	11.	10.	18.	8.	-0	-0	-0	-0
11	1	3	11	38.	1.	18.	14.	14.	13.	-0	-0	-0	-0
12	Posttest Score		12	59.	1.	16.	16.	21.	19.	-0	-0	-0	-0
13	Rater		13	63.	1.	18.	18.	21.	21.	20.	-0	-0	-0
14	Tape		14	70.	1.	22.	15.	17.	21.	-0	-0	-0	-0
15	2	2	15	8.	1.	13.	14.	20.	23.	-0	-0	15.	-0
16	2	3	16	30.	1.	17.	12.	25.	22.	-0	23.	19.	-0
17	2	4	17	31.	1.	14.	15.	24.	15.	-0	21.	-0	-0
18	2	5	18	49.	1.	18.	22.	28.	24.	-0	-0	-0	-0
19	2	6	19	61.	1.	18.	16.	19.	20.	-0	-0	-0	-0
20	2	7	20	69.	1.	14.	16.	20.	18.	-0	-0	-0	-0
21	2	8	21	57.	2.	15.	9.	16.	18.	-0	-0	-0	-0
22	2	9	22	10.	2.	14.	11.	15.	20.	-0	-0	13.	-0
23	2	10	23	54.	2.	11.	13.	20.	18.	23.	-0	-0	-0
24	2	11	24	24.	2.	11.	4.	7.	16.	-0	-0	-0	-0
25	2	12	25	68.	2.	21.	14.	24.	16.	-0	-0	-0	-0
26	2	13	26	65.	2.	17.	17.	19.	19.	-0	-0	-0	-0
27	2	14	27	23.	2.	13.	18.	21.	21.	-0	-0	-0	-0
28	2	15	28	64.	2.	20.	17.	18.	19.	17.	-0	-0	-0
29	2	16	29	50.	2.	18.	18.	18.	18.	-0	-0	-0	-0
30	2	17	30	4.	2.	16.	7.	25.	8.	20.	9.	-0	9.
31	2	18	31	33.	2.	10.	11.	19.	20.	-0	-0	-0	-0
32	2	19	32	9.	2.	14.	19.	22.	15.	-0	-0	-0	-0
33	2	20	33	32.	2.	14.	15.	15.	11.	18.	12.	13.	-0
34	2	21	34	53.	2.	11.	14.	12.	18.	-0	-0	-0	-0
35	2	22	35	17.	2.	17.	17.	8.	7.	-0	-0	-0	-0
36	2	23	36	29.	2.	14.	9.	23.	14.	-0	15.	-0	-0
37	2	24	37	39.	2.	17.	11.	23.	9.	-0	-0	-0	-0
38	2	25	38	40.	2.	18.	13.	22.	16.	-0	-0	-0	-0
39	2	26	39	3.	2.	13.	12.	24.	19.	28.	19.	12.	12.
40	2	27	40	26.	2.	24.	18.	23.	18.	24.	-0	-0	-0
41	2	28	41	45.	2.	11.	14.	20.	20.	20.	-0	-0	-0
42	2	29	42	27.	3.	12.	15.	22.	14.	-0	12.	13.	-0
43	2	30	43	48.	3.	7.	15.	13.	15.	-0	-0	-0	-0
44	2	31	44	51.	3.	15.	17.	23.	11.	-0	-0	-0	-0
45	2	32	45	6.	3.	15.	12.	20.	12.	25.	15.	-0	10.
46	2	33	46	35.	3.	14.	13.	9.	11.	-0	-0	-0	-0
47	2	34	47	74.	3.	12.	13.	17.	10.	-0	-0	-0	-0
48	2	35	48	36.	3.	13.	8.	0	0	-0	-0	-0	-0
49	2	36	49	55.	3.	17.	20.	17.	16.	-0	-0	-0	-0
50	2	37	50	12.	3.	12.	14.	17.	18.	-0	-0	-0	-0
51	2	38	51	43.	3.	10.	13.	21.	20.	-0	-0	-0	-0
52	2	39	52	67.	3.	14.	17.	23.	22.	-0	-0	-0	-0
53	2	40	53	19.	3.	14.	17.	12.	8.	-0	-0	14.	-0
54	2	41	54	44.	3.	11.	8.	13.	4.	-0	-0	-0	-0
55	2	42	55	5.	3.	15.	14.	21.	5.	17.	7.	14.	15.
56	2	43	56	58.	3.	16.	13.	24.	21.	-0	-0	-0	-0
57	2	44	57	72.	3.	11.	15.	21.	18.	-0	-0	-0	-0
58	2	45	58	60.	3.	13.	10.	15.	13.	-0	-0	-0	-0

* Scores used in actual data analysis were carried out to three decimals.

Intrarater Reliability

		Key									
		1	2	3	4	5	6	7	8	9	10
Column	Item	CASE 1-10	CODE 1-10	CODE 1-10	CODE 1-10	CODE 1-10	CODE 1-10	CODE 1-10	CODE 1-10	CODE 1-10	CODE 1-10
1	Case #	1	2.	1.	-0	0.	24.	-0	-0	-0	-0
		2	7.	1.	-0	-0	-0	-0	-0	-0	-0
		3	34.	1.	-0	-0	-0	-0	-0	-0	-0
2	Subject ID	4	46.	1.	4.	0	-0	12.	11.	12.	12.
		5	47.	1.	-0	-0	-0	-0	-0	-0	-0
		6	50.	1.	-0	-0	-0	-0	-0	-0	-0
3	Method	7	52.	1.	0	-0	26.	9.	0.	11.	4.
		8	71.	1.	-0	5.	17.	-0	-0	-0	-0
		9	25.	1.	13.	3.	24.	13.	13.	8.	12.
	Intrarater Reliability Pretest Time 2 Scores*	10	37.	1.	4.	2.	-0	13.	10.	10.	10.
		11	38.	1.	13.	-0	-0	-0	-0	-0	-0
		12	54.	1.	-0	-0	-0	-0	-0	-0	-0
		13	63.	1.	-0	2.	23.	17.	17.	5.	9.
		14	70.	1.	-0	-0	8.	11.	12.	5.	2.
		15	8.	1.	22.	7.	20.	15.	11.	15.	23.
		16	30.	1.	-0	4.	32.	-0	-0	-0	-0
		17	31.	1.	-0	-0	-0	-0	-0	-0	-0
		18	49.	1.	-0	4.	27.	21.	13.	10.	15.
		19	51.	1.	8.	-0	0	9.	11.	7.	14.
4	Part A	20	69.	1.	18.	-0	-0	-0	-0	-0	-0
		21	57.	2.	22.	-0	6.	13.	10.	13.	12.
		22	10.	2.	-0	5.	24.	-0	-0	-0	-0
5	Genogram	23	54.	2.	22.	-0	-0	-0	-0	-0	-0
		24	24.	2.	12.	0	34.	12.	6.	6.	8.
		25	68.	2.	-0	-0	23.	13.	16.	3.	2.
6	Videotape 1	26	65.	2.	-0	-0	24.	11.	16.	16.	17.
		27	23.	2.	21.	0.	19.	18.	10.	8.	18.
		28	64.	2.	-0	-0	-0	-0	-0	-0	-0
7	Videotape 2	29	50.	2.	-0	3.	-0	10.	6.	4.	13.
		30	4.	2.	15.	3.	11.	13.	16.	15.	14.
		31	33.	2.	19.	3.	14.	3.	5.	17.	0
8	Videotape 3	32	9.	2.	23.	-0	-0	-0	-0	-0	-0
		33	32.	2.	-0	-0	-0	-0	-0	-0	-0
		34	53.	2.	-0	-0	-0	-0	-0	-0	-0
9	Videotape 4	35	17.	2.	30.	4.	21.	16.	14.	3.	8.
		36	29.	2.	-0	-0	-0	-0	-0	-0	-0
		37	39.	2.	34.	4.	30.	-0	-0	-0	-0
10	Videotape 5	38	40.	2.	-0	10.	-0	17.	16.	14.	0
		39	3.	2.	-0	-0	-0	-0	-0	-0	-0
		40	26.	2.	-0	-0	-0	-0	-0	-0	-0
		41	45.	2.	20.	-0	-0	-0	-0	-0	-0
		42	27.	3.	-0	-0	-0	-0	-0	-0	-0
		43	48.	3.	-0	1.	-0	9.	10.	1.	4.
		44	51.	3.	20.	-0	-0	-0	-0	-0	-0
		45	6.	3.	12.	4.	17.	3.	5.	8.	12.
		46	35.	3.	-0	4.	7.	-0	-0	-0	-0
		47	74.	3.	-0	3.	27.	8.	7.	10.	19.
		48	36.	3.	1.	-0	-0	-0	-0	-0	-0
		49	55.	3.	32.	-0	-0	-0	-0	-0	-0
		50	12.	3.	24.	2.	33.	9.	11.	16.	11.
		51	43.	3.	-0	0	-0	7.	8.	9.	6.
		52	67.	3.	-0	-0	-0	-0	-0	-0	-0
		53	19.	3.	21.	-0	-0	-0	-0	-0	-0
		54	44.	3.	-0	-0	-0	-0	-0	-0	-0
		55	5.	3.	5.	-0	-0	-0	-0	-0	-0
		56	54.	3.	-0	4.	17.	7.	9.	8.	17.
		57	22.	3.	-0	-0	-0	-0	-0	-0	-0
		58	56.	3.	-0	-0	-0	-0	-0	-0	-0

*Scores used in actual data analysis were carried out to three decimals.

Intrarater Reliability

		1	2	3	4	5	6	7	8	9	10
	Key	CASE-NO	CODENUM	METHOD	INPOSTA	INPOSTGE	INPOSTUN	INPOSTTW	INPOSTTH	INPOSTFO	INPOSTFI
Column	Item	1	2.	1.	35.	4.	25.	18.	16.	20.	15.
		2	7.	1.	41.	13.	43.	23.	18.	20.	15.
		3	34.	1.	32.	9.	14.	8.	1.	24.	5.
		4	45.	1.	-0	-0	-0	-0	-0	-0	-0
1	Case #	5	47.	1.	42.	-0	23.	8.	13.	11.	12.
		6	50.	1.	21.	-0	-0	-0	-0	-0	-0
		7	62.	1.	-0	-0	-0	-0	-0	-0	-0
		8	71.	1.	-0	-0	-0	-0	-0	-0	-0
2	Subject ID	9	25.	1.	34.	-0	-0	-0	-0	-0	-0
		10	37.	1.	34.	-0	15.	12.	8.	14.	6.
		11	34.	1.	-0	-0	-0	-0	-0	-0	-0
3	Method	12	59.	1.	-0	11.	35.	-0	-0	-0	-0
		13	53.	1.	-0	6.	31.	14.	16.	24.	21.
		14	70.	1.	-0	-0	32.	17.	11.	23.	20.
		15	8.	1.	45.	10.	-0	13.	14.	23.	21.
	Intrarater	16	30.	1.	-0	5.	30.	-0	-0	-0	-0
	Reliability	17	31.	1.	-0	-0	-0	-0	-0	-0	-0
	Posttest	18	44.	1.	-0	-0	-0	-0	-0	-0	-0
	Time 2	19	61.	1.	35.	10.	34.	18.	10.	19.	20.
	Scores*	20	69.	1.	-0	9.	21.	16.	14.	20.	17.
		21	57.	2.	-0	6.	24.	11.	10.	17.	17.
		22	10.	2.	-0	-0	-0	-0	-0	-0	-0
		23	54.	2.	-0	7.	24.	-0	-0	-0	-0
		24	24.	2.	5.	0	-0	9.	9.	9.	15.
		25	63.	2.	45.	7.	34.	21.	15.	24.	17.
4	Part A	26	65.	2.	-0	-0	-0	-0	-0	-0	-0
		27	23.	2.	-0	-0	-0	-0	-0	-0	-0
		28	64.	2.	-0	5.	20.	18.	15.	18.	19.
		29	50.	2.	-0	-0	-0	-0	-0	-0	-0
5	Genogram	30	4.	2.	34.	-0	-0	-0	-0	-0	-0
		31	33.	2.	41.	-0	14.	11.	13.	14.	14.
		32	9.	2.	25.	11.	-0	16.	19.	17.	15.
6	Videotape 1	33	32.	2.	24.	-0	-0	-0	-0	-0	-0
		34	53.	2.	-0	3.	21.	11.	12.	10.	17.
		35	17.	2.	40.	12.	-0	14.	17.	8.	8.
		36	29.	2.	47.	8.	-0	17.	11.	21.	15.
7	Videotape 2	37	39.	2.	-0	-0	-0	-0	-0	-0	-0
		38	40.	2.	-0	-0	-0	-0	-0	-0	-0
		39	3.	2.	40.	10.	37.	10.	12.	24.	18.
		40	26.	2.	44.	-0	-0	-0	-0	-0	-0
8	Videotape 3	41	45.	2.	34.	-0	24.	11.	11.	18.	13.
		42	27.	2.	-0	-0	-0	-0	-0	-0	-0
		43	44.	2.	33.	-0	-0	-0	-0	-0	-0
9	Videotape 4	44	51.	2.	28.	-0	40.	14.	19.	24.	9.
		45	6.	2.	-0	-0	-0	-0	-0	-0	-0
		46	35.	2.	-0	4.	35.	-0	-0	-0	-0
		47	74.	2.	-0	-0	-0	-0	-0	-0	-0
10	Videotape 5	48	36.	2.	-0	-0	-0	-0	-0	-0	-0
		49	55.	2.	-0	12.	32.	-0	-0	-0	-0
		50	12.	2.	20.	4.	-0	11.	15.	14.	19.
		51	43.	2.	-0	-0	-0	-0	-0	-0	-0
		52	67.	2.	-0	-0	-0	-0	-0	-0	-0
		53	19.	2.	27.	3.	27.	11.	16.	10.	2.
		54	44.	2.	-0	-0	-0	-0	-0	-0	-0
		55	5.	2.	-0	-0	-0	-0	-0	-0	-0
		56	54.	2.	-0	4.	21.	16.	14.	24.	22.
		57	22.	2.	37.	-0	-0	-0	-0	-0	-0
		58	60.	2.	-0	7.	15.	-0	-0	-0	-0

*Scores used in actual

data analysis were

carried out to three

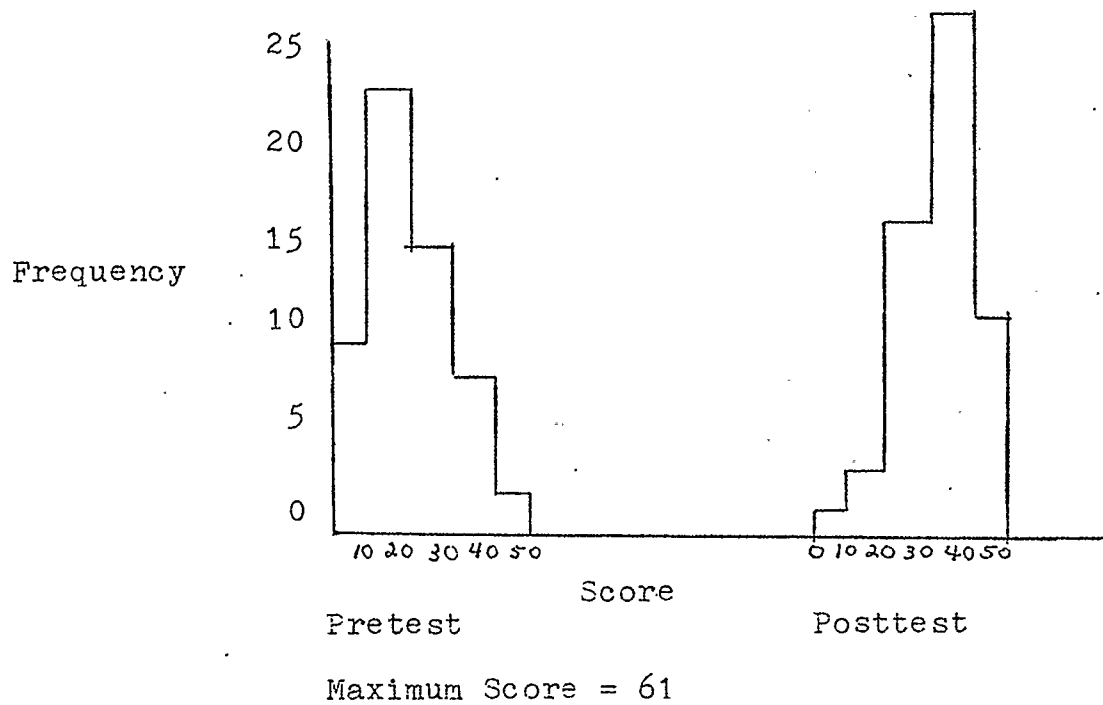
decimal places.

Appendix 5
Distribution of Achievement Scores

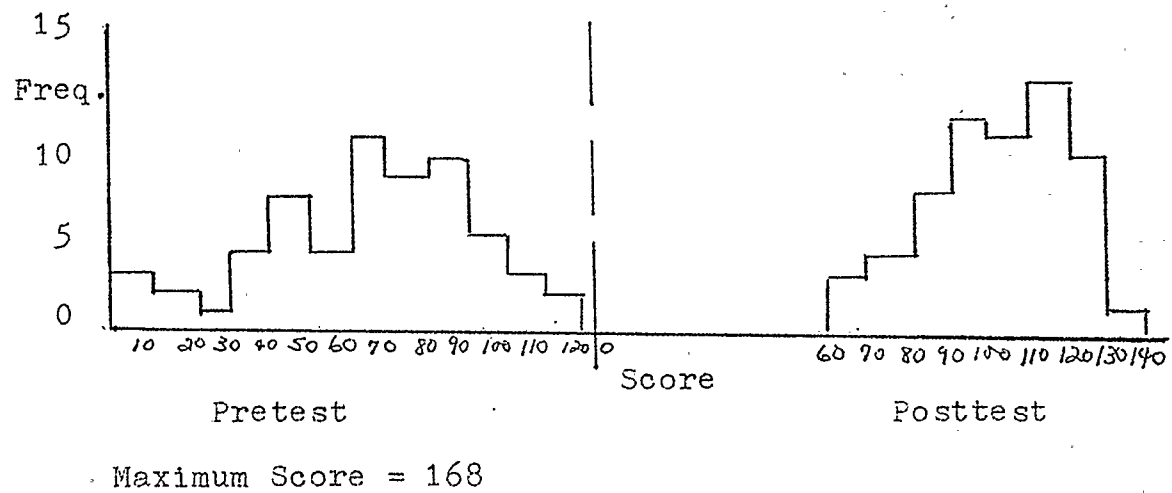
Distribution of Achievement Scores

178

Part A

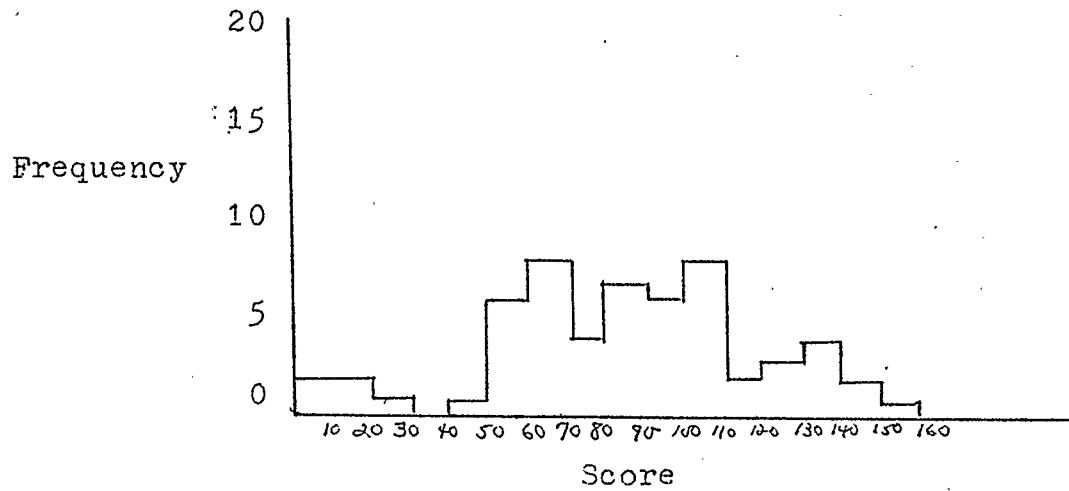


Part B



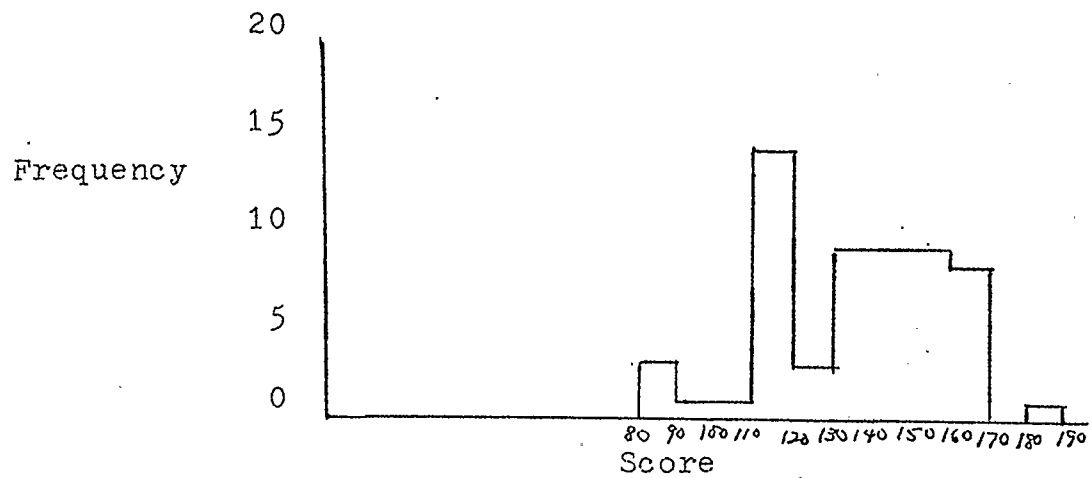
Pretest Total

179



Maximum = 229

Posttest Total



Maximum = 229

Appendix 6

Effect of Attendance on Achievement

Effect of Attendance on Achievement

39 students missed no classes. 19 students missed one or more classes. The attendance figures for each method are given:

Attendance	Method		
	1	2	3
Missed 0 Classes	12	13	14
Missed 1 Class	8	7	2
Missed 2 Classes	0	0	0
Missed 3 Classes	0	1	1

A two-way analysis of variance with repeated measures was used to test the effect of attendance on achievement. The F statistic and probability levels are given in Table 1.

Table 1
Probability of No Significant Difference Between
Tests, Test-Group Interaction, and Attendance

Groups			
	Source of Variation	F	Probability
Part A	Tests	69.94	.00*
	Test-Group	5.88	.02**
	Groups	3.35	.07
Part B	Tests	101.95	.00*
	Test-Group	2.41	.13
	Groups	2.77	.10
Total	Tests	118.77	.00*
	Test-Group	4.61	.04**
	Groups	3.23	.08

* $p < .01$

** $p < .05$

Results from Table 1 indicate for:

1. Part A: There is a significant test effect with the posttest being higher than the pretest. There is a significant interaction effect with those who missed class gaining 8.53 points and those who attended all classes gaining 15.50.
2. Part B: There is a significant test effect.
3. Total: There is a significant test effect and

a significant interaction effect. Those who attended all classes gained 56.02 points while those who missed classes gained 37.55 points.