AN EXPERIMENTAL STUDY OF THE EFFECT OF TWO DISCUSSION TECHNIQUES ON EDUCATIONAL OUTCOMES IN A BEGINNING EDUCATIONAL PSYCHOLOGY CLASS TAUGHT BY CLOSED-CIRCUIT TELEVISION

Thesis for the Degree of Ed. D.

MICHIGAN STATE UNIVERSITY

Morton D. Dunham

1958

This is to certify that the

thesis entitled
An Experimental Study of the Effect
of Two Discussion Techniques on Educational
Outcomes in a Beginning Educational Psychology Class
Taught by Closed-Circuit Television

presented by

Morton D. Dunham

has been accepted towards fulfillment
of the requirements for

Ed.D degree in Education

William M. Vaughan
Major professor

Date October 20, 1958

OF

OUTCOMES

Submi

Dep

AN EXPERIMENTAL STUDY OF THE EFFECT OF TWO DISCUSSION TECHNIQUES ON EDUCATIONAL OUTCOMES IN A BEGINNING EDUCATIONAL PSYCHOLOGY CLASS TAUGHT BY CLOSED-CIRCUIT TELEVISION

By

Morton D. Dunham

A THESIS

Submitted to the School for Advanced Graduate Studies of
Michigan State University of Agriculture and
Applied Science in partial fulfillment of
the requirements for the degree of

DOCTOR OF EDUCATION

Department of Administrative and Educational Services

1958

Grateful ac
my major professor
course of the stud

To Dr. John

To Mr. Clare
the course, for his
worthwhile and enjo

To the studen

And especiall
typing the manuscri
study could not have
sacrificed much.

ACKNOWLEDGMENTS

Grateful acknowledgment is made to Dr. William W. Farquhar, my major professor, for his advice and assistance throughout the course of the study.

To Dr. John D. Krumboltz for his advice and suggestions.

To Mr. Clarence M. Williams, my colleague and co-worker during the course, for his kind assistance and many helps that made the year worthwhile and enjoyable.

To the students who participated in the course.

And especially to my wife, Mary, for the many hours spent in typing the manuscript and without whose understanding and patience the study could not have been completed; and my sons, Marc and Martin, who sacrificed much.

AN EXPERIMENTAL STUDY OF THE EFFECT OF TWO DISCUSSION TECHNIQUES ON EDUCATIONAL OUTCOMES IN A BEGINNING EDUCATIONAL PSYCHOLOGY CLASS TAUGHT BY CLOSED-CIRCUIT TELEVISION

By

Morton D. Dunham

AN ABSTRACT OF A THESIS

Submitted to the School for Advanced Graduate Studies of
Michigan State University of Agriculture and
Applied Science in partial fulfillment of
the requirements for the degree of

DOCTOR OF EDUCATION

Department of Administrative and Educational Services

1958

It was the

that the absence o

ference in educati

instruction under

discussion.

The experi

(TV 200) at Michig

ing room, two view

vision production.

The subject

jects, of which 4

randomly assigned

and the non-telev

jects were furthe

levels of the fou

A four was

of the classifica

two discussion te

Psychological Exa

measured by the P

The combination o

formulation of fi

cerned with main

ject or criterion

(1) achievement as

It was the purpose of the study to objectively assess the theory that the absence of maximum feedback would result in a significant difference in educational outcomes between two groups of subjects receiving instruction under conditions of television discussion and face-to-face discussion.

The experiment was conducted in an educational psychology course (FE 200) at Michigan State University. Facilities included an originating room, two viewing classrooms, and the necessary equipment for television production.

The subjects were one hundred and sixty-seven undergraduate students, of which 46 were males and 121 were females. The subjects were randomly assigned to two groups designated the television (TV) group and the non-television (NTV) group. For purposes of analysis the subjects were further divided into sixteen subgroups based on the two levels of the four independent variables.

A four was factorial design was employed with two levels in each of the classification variables. The independent variables were (1) the two discussion techniques, (2) ability level as measured by the ACE Psychological Examination, (3) preference for type of instruction as measured by the Preferred Instructor Characteristics Scale, and (4) sex. The combination of two levels for each variable made possible the formulation of fifteen testable null hypotheses of which four were concerned with main effects and eleven with interacting effects. The dependent or criterion variables used to measure educational outcomes were (1) achievement as measured by a Pre and Post Achievement Test,

(2) attitude toward teaching as measured by the Minnesota Teacher Attitude Inventory, (3) misconceptions about education as measured by the Misconceptions about Education Scale, and (4) course opinion as measured by the Course Evaluation Scale.

Discussion was conducted over the closed-circuit television system for the TV group and in a face-to-face classroom situation for the NTV group. Discussion was conducted in small groups of six students, in panel groups, and in the entire class situation. Discussion periods were fifty-minutes in length. All discussion periods were preceeded by fifty-minute lectures which were televised to both TV and NTV groups.

Evaluation of results was accomplished by means of the Chi-square test, "t" test, and analysis of variance and covariance. No significant differences were found for the achievement and misconceptions about education criteria. A significant interaction effect in terms of the attitude toward teaching criteria was found for the interaction of the discussion techniques and preference for instruction variables. A significant difference was obtained in terms of the opinion criteria for the discussion techniques effect.

The conclusions based upon the study:

1. There is no evidence to indicate the superiority of either the face-to-face discussion methods or the television discussion methods for producing changes in educational outcomes in terms of achievement and misconceptions about education criteria.

2. The combination of face-to-face discussion and students with a "cognitive" preference for instruction is more effective in producing

changes in attitude toward teaching then is the combination of face-to-face discussion and "affective" students.

3. Television discussion is about equally effective for "cognitive" and "affective" students in attitude change, but less so than face-to-face discussion with "cognitive" students.

4. Face-to-face discussion produces more favorable course opinions than television discussion.

5. Face-to-face discussion results in fewer course criticisms than television discussion.

William W. Farquhar

Dr. William W. Farquhar

TABLE OF CONTENTS

CHAPTER PAGE

PAGE

LIST OF TABLES

TABLE PAGE

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

The r

rapidly in t

duced in the

where it can

an attempt t

of instructi

method.

The a

views and cl

instrument o

problems of

adjunct to s

Dunha

structional

Final

that tel

single d

can be b

educatio

and reso

Presi

1952 at State

CHAPTER I

DEFINITION OF THE PROBLEM

The role of instructional television in education has increased rapidly in the past few years. In many instances it has been introduced in the classroom setting with little knowledge of how, when and where it can be used most effectively. The present investigation is an attempt to evaluate the effectiveness of this relatively new medium of instruction as an aid in teaching large classes by the discussion method.

I. THE PROBLEM

The appearance of television has brought forth many divergent views and claims about the potentialities of this new medium as an instrument of education. Some view it as a virtual panacea for the problems of education while others more modestly see television as an adjunct to sound educational practices.

Dunham and Lowdermilk (16) in concluding a brief review of instructional television research state that:

> Finally, experimental studies have tended, generally, to show that television can serve as an instrumentality whereby every single device, technique and process known to the art of teaching can be brought to bear in group instruction--in short, that its educational potential is limited only by the creative imagination and resourcefulness of those who undertake to use it.

President Millis of Western Reserve University in an address in 1952 at State College, Pennsylvania stated that television was "one of

the most magnificent instruments for raising the quality of teaching whether it be college teaching or elementary teaching."[1]

On the other hand Telford Taylor, former general counsel for the Federal Communication Commission, told the New York commission on educational television:

> Television is not a substitute for the teacher, but a new tool for the teacher to use...Its proper use will not contract, but will greatly expand, the scope and opportunity for question and answer between the teacher and the student.[2]

And Milton Eisenhower has observed that television will "supplement rather than replace the classroom."[3]

The present study was undertaken in the hope of initiating as least some clarification of such diverse views as those expressed above, on one aspect of instructional television, the use of the discussion method.

The discussion method is used at all levels and for a wide variety of subject areas as an instructional method. There was no attempt in the present study to determine the effectiveness of the discussion method as contrasted with the lecture method or other teaching methods not using discussion techniques. It was an underlying assumption of the study that benefits are to be derived from the use of the discussion method. Although this assumption was not tested in the

---

[1]As quoted by Martin Packman, "Educational Television," Editorial Research Reports, May 18, 1954, Vol. 1, No. 10, p. 376.

[2]Ibid, p. 374.

[3]Ibid, p. 374.

present experiment, a number of research studies indicate that the discussion method is more effective than the lecture method in the learning of certain kinds of materials, and a brief review of this research will be made here.

Studies of the acquisition of information where the criterion has been the ability to recall or recognize factual material on objective type tests have shown the lecture and discussion methods to be equal in effectiveness in investigations by Bane (4), Carlson (11), Eglash (18), Gerberich and Warner (23), Hudelson (24), and Johnson and Smith (26). On the other hand in the same type of investigations Hudelson (24), Ruja (50) and Spence (53) have reported the lecture to be more effective than discussion methods.

Studies by Bane (4) and Rickard (47) of the retention of learned material have shown that knowledge acquired in discussion classes was retained significantly better than that resulting from reading and lectures without discussion.

Investigations by Bloom (6), Brinkley (10), Edmitson and Braddock (17), Ruja (50), and Ward (57) in which the objective was concerned with aiding students in the ability to evaluate, draw inferences, synthesize, perceive relationships, and make application of learned material it has been found that discussion is significantly superior to the lecture.

It has also been shown that discussion is superior to the lecture method in affecting changes in attitudes and behavior in investigations by Bond (7), Levine and Butler (33), Lewin (34), Maier (40), and Ruja (50).

Although there is no agreement about the relationship between learning and classroom morale, it has been demonstrated that an important relationship exists between the emotional adjustment of students and between teacher-pupil relationships and classroom morale. Studies by Asch (3), Bovard (9), Faw (22), Lewin (35), and Ruja (50) indicate that the discussion method is more conducive to promoting high morale and good interpersonal relationships between the teacher and the pupil than the lecture method.

The assumption made previously that benefits are to be derived from the use of the discussion technique would seem to be supported on the basis of the above research.

An explanation for the findings that in certain kinds of learning the discussion method is superior to the lecture method may be found by reference to current learning theory. In simplified terms, according to the association theorists such as Thorndike, Guthrie and Hull learning occurs by the association of cues (stimuli) and responses. The association or connection of cue and response is enhanced by various factors such as exercise (practice), effect (reward or non-reward), the strength of the cues, the distinctiveness of the cues, the relevancy of the cues, and others.

Although relatively little is known about higher reasoning processes in human learning, it may be theorized that this type of learning is reducible to simple components of a stimulus-response nature. Thus the discussion method may be more effective than the lecture method

because there is

and responses.

presented to the

the learner to t

are presented to

hand under lectu

number of cues,

specific cues, a

to simple trial

method allows fo

results, whereas

trials with litt

It will b

studies of the l

found to be at l

discussion metho

the learning of

applying learned

the nature of ro

highly abstract

and response is

actually interfe

ambiguous or irr

The result

Leavitt and Muell

because there is a greater opportunity for the association of the cues and responses. In the discussion method a greater number of cues are presented to the learner for appraisal, more opportunity is afforded the learner to try-out and select specific cues, and more occasions are presented to the learner to practice the material. On the other hand under lecture conditions the learner is presented a limited number of cues, he is given less opportunity to try-out and select specific cues, and there are fewer opportunities to practice. Reduced to simple trial and error terms, it may be said that the discussion method allows for a greater number of trials plus the knowledge of results, whereas the lecture method allows for a limited number of trials with little or no knowledge of results.

It will be recalled from the research cited above that in the studies of the learning of factual information the lecture method was found to be at least equal to and in some experiments superior to the discussion method. The learning of factual information as opposed to the learning of problem-solving abilities such as drawing inferences, applying learned information, synthesizing, and the like, is more in the nature of rote memorization. Where the information is not of a highly abstract nature it may be theorized that the association of cue and response is easily made by the learner and that the discussion may actually interfere with learning due to the presentation of too many ambiguous or irrelevant cues.

The results of a related study in communication theory by Leavitt and Mueller (32) is of interest. The authors reasoned that

according to the information theory of cyberneticists and the trial and error theories of psychologists in order for A to hit some target B it is necessary for A to be constantly informed of A's own progress. Therefore if A attempts to hit B with some information, A will be more successful if B provides A with some cues which A cannot obtain directly. In other words where communication between A and B is the goal, feedback in the form of expressive or verbal language should make for greater effectiveness, especially if the material being communicated is new or relatively abstract.

In the experiment by Leavitt and Mueller a series of rectangular figures of equal area but of varying shapes were communicated to a group of students under four degrees of feedback. The students were required to produce the size and the shape of the patterns. The four conditions of feedback were: (1) zero feedback in which the experimenter described the figures to the students, but the subjects could not see the experimenter nor communicate with him in any way, (2) partial feedback in which the experimenter described the figure to the students and the students could communicate to the experimenter by facial expressions, (3) partial feedback in which the experimenter described the figures to the subjects and the experimenter could answer questions of the subjects by a "yes" or "no", and (4) free feedback in which the experimenter and the subjects could freely communicate with one another. Of course, under all conditions the geometric figure was never visible to the subjects. The subjects were assigned to groups corresponding to the four conditions of feedback and all groups experi-

enced each of the feedback conditions in varying orders.

It was found that the mean accuracy score for the subjects in reproducing the patterns increased steadily from the conditions of zero to free feedback. It was also found that the subjects displayed more hostility when the condition of zero feedback followed the other patterns. The authors report the following conclusions:

(1) Feedback increases the accuracy with which information is transmitted.

(2) Feedback increases sender and receiver confidence in what they have accomplished.

(3) Increased feedback results in more time spent in explanation.

(4) Feedback experience improves subsequent zero feedback trials considerably.

(5) Sender experience contributes more than receiver experience to improved accuracy of communication.

(6) Zero feedback engenders hostility in the receiver.

(7) Zero feedback engenders doubt in the sender.

Insofar as teaching is the communication of information from the instructor to the student, it would seem on the basis of the research and arguments presented above, that learning will be at a maximum where feedback from the student to the instructor is also at a maximum. The above principle would seem to be particularly applicable to those situations where the material to be learned is of an abstract nature or where an attempt is being made to teach students the ability to make

applications of

In situations w

tions between s

in attitude and

tally at least,

a maximum to in

Under co

assume that fee

the most ideal

and technical p

microphones so

plete visual an

practical purpo

ations where th

factual informa

absence of feedi

situations wher

tained; that is

are of a more al

techniques, the

or the condition

developed here,

student. It is

would result in

crease in the ho

applications of learned material, to draw inferences, and the like. In situations where high classroom morale and good interpersonal relations between student and teacher are deemed important or where changes in attitude and behavior are objectives it would also appear, theoretically at least, that feedback between student and teacher should be at a maximum to insure maximum effectiveness.

Under conditions of television instruction it is reasonable to assume that feedback from the student to the instructor, under even the most ideal conditions, would never be at a maximum. The physical and technical problems involved in arranging cameras, monitors and microphones so that the instructor and the students would be in complete visual and audio communication at all times are such, that for practical purposes, ideal conditions could not be obtained. In situations where the communication is principally one-way; that is, where factual information is being conveyed by means of the lecture, the absence of feedback is probably not a serious handicap. However, in situations where it is important that two-way communication be maintained; that is, where the subject-matter or the objectives to be met are of a more abstract nature such as the learning of problem-solving techniques, the ability to apply principles, the absence of feedback or the condition of partial feedback would, according to the argument developed here, prove a handicap to both the instructor and the student. It is also theorized that the absence of maximum feedback would result in the lowering of class morale and a corresponding increase in the hostility and insecurity displayed by the students.

Can a discussion conducted by closed-circuit television between an instructor and students in two different rooms produce as effective educational outcomes as those produced by conventional face-to-face discussion?

It was the purpose of the study to ascertain the relative effectiveness of two discussion techniques, one employing conventional face-to-face methods and the other using closed-circuit television, on educational outcomes as measured by four criteria: (1) achievement of course content, (2) attitudes toward teaching, (3) misconceptions about education principles and practices, and (4) student opinion about the course.

A complete statement of the specific null hypotheses to be tested will be found in section two of Chapter III, "The Design of the Experiment."

II. IMPORTANCE OF THE STUDY

The recent and predicted increases in the number of students at all educational levels in proportion to the present and predicted supply of available teachers poses serious problems. It is important that more efficient teaching techniques be identified and used. Inasmuch as the number of students is increasing more rapidly than the number of teachers, it is also important that methods be devised for providing instruction for classes of larger size.

The advent of television in recent years has provided a means of increasing markedly the number of students that can be reached by a

single instructor.
communicate with
rangement. The m
vision would list

(1) One i
nication with a l
viding a solution

(2) The
better qualified
of the savings.

(3) The
be expected from

(4) A be
specialists in p

(5) The
visual aids and

(6) The
a resulting savi

The above
based on the bel
effectiveness to
all of the prese
television as th
research in rece
instruction by t

single instructor. There are a number of advantages in being able to communicate with large groups by a television and loud-speaker arrangement. The more enthusiastic advocates of instructional television would list the following as important:

(1) One instructor is provided with a means of two-way communication with a large number of students simultaneously, thereby providing a solution to the problem of teacher shortages.

(2) The number of instructors required can be reduced and better qualified and trained instructors can be employed as a result of the savings.

(3) The quality of instruction will be more uniform than can be expected from several instructors.

(4) A better quality of instruction will result by the use of specialists in particular area of instruction.

(5) The amount of supplementary equipment needed such as audio-visual aids and demonstration material can be reduced.

(6) The number of small classrooms needed can be reduced with a resulting savings in construction costs.

The above advantages for the use of television instruction are based on the belief that television instruction is at least equal in effectiveness to the more conventional teaching procedures and that all of the present techniques can be performed as effectually over television as they can be in the face-to-face situation. Much of the research in recent years in instructional television has claimed that instruction by the medium of television is at least equal and in many

instances superior to regular instruction, and some of the more pertinent of these studies will be reviewed in Chapter Two. Much of this research does not support the claims that are made for it, and many of the claims are based on hasty generalisations from inadequate research. It is important for practical purposes that a clearer understanding of the strengths and weaknesses of this new medium be investigated, and the present experiment was designed in an attempt to answer a practical and important question concerning the use of instructional television.

III. UNIQUE ASPECTS OF THE STUDY

The design of the present study was a unique feature, in that it made use of the more recent statistical techniques. The design was a four-factor classification design with four classification variables and two levels or categories in each of the variables. Such a design allows the experimenter to simultaneously measure the main and interaction effects of several control variables under different treatment conditions without resorting to the traditional procedures of several independent experiments of the single variable type. Such a design increases the precision of the experiment because the subjects are randomized within the treatment groups and it permits a separate study of the treatment effects at different levels of the control variable. By adapting this design to an analysis of covariance technique it was possible to statistically control for initial differences in subjects on the control variables. A complete description of the de-

sign is given in Chapter Three.[3]

Another unique aspect of the experiment was the attempt to investigate the problem in a realistic setting. Inasmuch as one of the problems confronting educators at the present time is that of providing instruction to large numbers of students, the present study was conducted in a large class setting. Although other investigators have included the use of large classes in the framework of their experimentation, so far as is known to the present writer, there have been no investigations in which the television and discussion techniques have been used simultaneously in the large-class setting.

A review of the more pertinent background studies of instructional television and of studies closely related to the present study will be presented in Chapter Two. Chapter Three will contain a description of the experimental design, a statement of the null hypotheses, and a description of the statistical procedures used. The experimental methods and procedures will be reviewed in Chapter Four, and Chapter Five will contain a description of the evaluation instruments. The results of the analysis of the data will be discussed in Chapter Six.

---

[3]For a detailed discussion of modern experimental designs see E. F. Lindquist, Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin Co., 1953.

CHAPTER II

REVIEW OF THE LITERATURE

Much has been written on television communication of which relatively little is concerned with research. Prior to 1950 most of the research was devoted to the technological and scientific aspects of television, but from 1950 to the present an increasing amount of experimentation has been related to the sociological and educational problems of television. These latter studies may be classified into four areas: (1) studies of the general effects of television in the everyday lives of people, (2) studies of the content of television programs, (3) studies of the technical aspects of using television, and (4) studies of the effectiveness of television used as a means of instruction. The literature appropriate to the scope of the present study concerns experimentation with educational television and will be restricted specifically to instructional television.

In the review that follows studies will be classified under two main headings: (1) general background studies and (2) specific studies closely related to the present investigation. The former classification will contain studies that provide background and information which will aid in the understanding of the present study, and the latter will comprise studies which are directly related to the present study. The reviewed studies in each of the above classifications will be further divided on the basis of the purpose of the experiment.

I. EVALUATION CRITERIA

The diversity of purposes, experimental designs, experimental methods, analyses of results, and interpretations of findings make it necessary for the reviewer of research to employ an objective standard for evaluating studies. Such a standard is necessary in order that subjective factors on the part of the reviewer may be eliminated as much as possible, and to communicate to the reader the evaluation standards employed. Farquhar and Krumbolts (21) have devised a useful check list for this purpose, and it is the one used in the present review. It is reproduced below.

A Check List for Evaluating Experimental Research in Psychology and Education

Satisfactory	Unsatisfactory	Questions
		A. The Problem
		1. Was the problem clearly defined?
		2. Was the problem framed in the form of hypotheses which were experimentally verifiable?
		B. The Design
		1. Was the selected statistical design appropriate to the particular experimental methods, conditions, subjects, and hypotheses under test?
		2. Was the population from which the sample was drawn clearly specified?
		3. Was the method of drawing the sample from the population clearly specified?
		4. Was the control group chosen in the same manner and from the same population as the experimental groups?
		5. Were the various treatments (including control) assigned at random to the groups?
		6. Did the experiment include a replication?
		7. Was the level of significance necessary for rejection of the null hypotheses specified before the data was collected or analyzed?

C. The Procedure
   1. Were the treatments and methods of collecting data described so that an independent investigator could replicate the experiment?
   2. Were the size and characteristics of the sample adequately described?
   3. Were the treatments administered so that extraneous sources of error were either held constant for all treatment and control groups or randomized among subjects within all groups?

D. The Analysis
   1. Was the criterion measure appropriate?
   2. Was any evidence of the reliability of the criterion measure given for the experimental sample?
   3. Were the statistical assumptions which are necessary for a valid test of the null hypotheses satisfied?

E. The Interpretation
   1. Were the conclusions consistent with the results obtained?
   2. Were generalizations confined to the population from which the sample was drawn?

The check list will be used in the present review to evaluate only those studies that are specifically related to the present study.

II. GENERAL BACKGROUND STUDIES OF INSTRUCTIONAL TELEVISION

Instructional television research has been sponsored by several agencies and groups, foremost of which are colleges and universities, secondary schools, and the military services. Some of the research has been supported by the agency itself and some has received support from foundations. Most of this research has been concerned with comparisons of instruction by live television, kinescope projections, and conventional teaching methods in the classroom. The majority of these studies have pertained to achievement and retention of course content

and to opinions and evaluation of television instruction. For purposes of understanding the background for the present study, a summary of some of the more pertinent studies will follow.

Studies of Achievement and Retention of Course Content

A study to determine the comparative effectiveness of television, kinescope recordings, and classroom instruction reported by Rock, Duva, and Murray (48) was conducted at the Special Devices Center of the United States Navy beginning in 1949. Nine naval air stations with approximately 40 Naval Air Reservists at each station were used in the experiment. Three stations received instruction by means of live television, three stations were presented kinescopes of the same lesson, and three stations were instructed by local instructors so that for each treatment there were approximately 100 to 120 men. The course content consisted of two series of eight lessons each; one series for officers and one series for enlisted men. In the live television groups, sixteen-inch receivers were used with ten men to each receiver, and a feedback microphone was provided so that the men could ask questions during the lecture. In the groups viewing recordings, the films were projected through a motion-picture projector on a screen with from 20 to 40 men assigned to a room. In the groups taught by local instruction, the instructor gave his lesson prepared from a lesson-plan rather than a verbatim script of the lecture. Pre-test and post-tests were given to all subjects. The tests were of the multiple-choice type and consisted of 30 items on the pre-test and post-test plus an additional 30 items on the post-test. Results of the study are given in

terms of percentages. In comparing the live television and local instruction groups, the television groups were superior for 50% of the officers and 53% of enlisted men, and equal to local instruction for 38% of the officers and 20% of the enlisted men. Television was inferior for 13% of the officers and 27% of the enlisted group. The authors conclude that 80% of the comparisons show television as good as or better than local instructors. In comparing the television recordings with local instruction, the authors conclude that the recordings were superior or as good as local instruction in 75% of the comparisons. In comparisons of live television and recorded television it is reported that live television is superior to recordings in a great majority of the cases. Several criticisms can be made of this study. Comparing live television received on a sixteen-inch receiver with recorded television projected onto a movie screen introduced uncontrolled variables. Comparing the same instructor in the live television and recorded television groups with three different local instructors also introduced a variable that was not controlled. There is no evidence in the report that the men in the three groups were randomly assigned to the experimental groups, nor is there evidence that indicates the men in the nine naval air stations were from the same population. The additional 30 items in the post-test could not be used in a comparison of pre-tests and post-test scores. The statistical model used was inappropriate. The design of the experiment was such that a treatment by levels design analyzed by covariance technique would have been more appropriate. No conclusive results can

be drawn from the percentage figures reported by the authors.

In another experiment reported by Rock, Duva, and Murray (49) an attempt was made to determine the effect of television on achievement and retention of instruction. A series of eight one-hour telecasts were broadcast to 160 groups of approximately 3,000 U. S. Army reservists ranging from private to colonel. The reservists viewed the programs in 10 different cities in groups of varying sizes. Receivers were of various types ranging from 7" screens to theater-type projection screens. Pre-tests and post-tests were given and a delayed recall test was administered three to six weeks after the original lesson. Each of the eight programs was produced with professional actors and was concerned with the phases of an Army's division operations. The authors report, "All grades of officers and enlisted men made higher scores on test questions after the telecasts than they did before the telecasts." Figures for the gains are not shown. On retention tests it was reported that officers retained 85% of newly learned material and enlisted men retained 65% of newly learned material for a period of six weeks. Much the same criticisms made of the previous study by these same authors can be made of this experiment. No attempt was made to equate or randomize the subjects. The effect of using different sized groups and different receiving methods was uncontrolled. The fact that all grades made higher scores after receiving the telecasts would be expected, but whether or not these gains were significantly different for different ranks, under different modes of reception, and in different sized groups was not answered. The investigation is more nearly a

survey than it i

A study

jects by televi

was given to fo

tion, and to si

The course was

true-false test

The same instru

tion. Students

possible 32, a

26.5. The aut

In comparing t

that the class

indicates that

scores. It is

attributed to

the subjects w

by academic st

In a cl

at Annapolis

shipmen each

The battalions

vision instruc

In the second

tion. All pa

survey than it is a scientific study.

A study of the feasibility of teaching Quartermaster Corps subjects by television was conducted by Allen (1) in 1954. Instruction was given to forty-seven ROTC students by means of television instruction, and to sixty ROTC students by regular classroom instruction. The course was four hours in length. A 32 item multiple-choice and true-false test was given four days after the conclusion of the course. The same instructor was used for both groups in any one day's instruction. Students in the regular classroom group scored 27.7 out of a possible 32, and the students receiving television instruction scored 26.5. The author gives no data on the significance of the difference. In comparing the academic standing of the two groups the author found that the classroom group was higher than the television group, and indicates that this could account for the differences in final test scores. It is possible, however, that such small differences can be attributed to chance. The findings in this study are limited because the subjects were not randomly assigned to the two groups nor equated by academic standing prior to the experiment.

In a closed-circuit television experiment at the Naval Academy at Annapolis reported by Boone (8), six battalions of about 140 midshipmen each were given training in two courses of 25 minutes length. The battalions were split into two groups: one group receiving television instruction and one group receiving normal lecture instruction. In the second course the groups were reversed for the mode of instruction. All participants were administered a 10 minute prognostic test

to determine t
courses all we
lecturer for t
television gro
analyze the re
significantly
second course
the television
nesses. An ur
instructors in
methods on the
adequate test
Finally the fa
television ins
indicates eith
tion effect th
television.

Dowell
Force to evalu
instructional
last three day
The trainees w
struction by t
classroom inst
two instructor

to determine the initial learning level and upon completion of the courses all were given post-tests. Instruction was by the regular lecturer for the lecture groups but by a different instructor for the television groups. An analysis of covariance technique was used to analyze the results. In the first course the television group was significantly superior at beyond the 1% level of significance. In the second course the non-television group was significantly higher than the television group at the 5% level. This study has several weaknesses. An uncontrolled variable was introduced by using different instructors in the different groups. Comparing the results of the two methods on the basis of a 25 minute class session does not comprise an adequate test of either television or regular classroom procedures. Finally the fact that one group did better first under conditions of television instruction and then under conditions of lecture instruction indicates either an initial difference in the two groups or a motivation effect that was carried-over from the novelty of exposure to television.

Dowell (15) reports the results of an experiment by the Air Force to evaluate the effectiveness of closed-circuit television as an instructional medium. A group of 266 Air Force trainees were given the last three days of an eight week course in Electronics Fundamentals. The trainees were divided into two groups: one group receiving the instruction by television and the second group receiving conventional classroom instruction. Each of the groups were further split between two instructors so that each instructor taught the trainees by both the

television and the conventional methods. An experimental group was taught by one of the instructors and the experimental population was approximately doubled two weeks later by repeating this procedure with the second instructor. A control group, composed of students equally matched with those in the experimental group, was selected from subsequent classes in the course. The students were matched between groups by using grades obtained in phase I and II of their earlier training. The matching variable was obtained by computing simple and multiple coefficients of correlation between tentative matching variables--grades, aptitude scores, and combinations of these--and the criterion variable for the control group. The criterion variable was a 50-item multiple choice achievement test. The correlation coefficient between the matching and criterion variables was given as .66.

The report of findings in this study are poorly organized and confusing, and will be discussed in some detail. The author first reports on a methods experiment using two methods and two instructors with 90 students in each of the TV groups and Control groups and 90 students for each of the two instructors; a total of 180 students. The author then refers to the precision with which students were matched between instructors by reference to the following table:

Table IV

Mean Scores on Matching Variable

Instructor	TV Gp	Control Gp
A	46.73	50.62
B	46.76	50.71

Reference is next made to an analysis of variance table as follows:

> An F ratio test was used in analyzing the variance of criterion scores. In this analysis, which follows, the apparent variance of the effects of methods is misleading. It must be remembered that matching was between instructors and not between "methods groups." As can be seen from the mean scores, the control group was the superior group.

Table V

Source	d.f.	Variance	Obtained F	Required 1% level
Methods	1	466.52	8.14	6.84
Instructors	1	6.0	.10	6.84
Methods x Instructors	1	58.48	1.02	6.84
Within Classes	176	57.3		
Total	179			

It is not clear whether this table presents an analysis of variance of criterion scores--as the author indicates--or an analysis of scores on the matching variable as given in the preceeding table. If it refers to criterion scores, a highly significant difference existed between the TV group and the Control group. If it refers to the scores on the matching variable, a highly significant difference existed between the two groups on Phase I and II grades and indicates that the groups were not equally matched initially. Since no mean scores are presented for the criterion variable, the author's reference to mean scores is apparently to those on the matching variable presented in Table IV.

Dowell further reports on a methods experiment involving two methods and three levels. Again the original study is quoted:

> The number of students utilized for analyzing the variance of criterion scores on a methods experiment involving 2 methods and 3 groups of students classified according to initial ability levels on the matching variable is shown in Table VI. The subjects taught by both instructors were pooled since the variance between instructors and the interaction between instructors and

methods was found to be insignificant.

Table VI

Number of Students

Group	No. of Subjects	Mean Score on Matching Variable
TV	91	49.2
Control	91	49.3
Total	182	

The number of students in this table is 182 as compared to 180 in the previous reference, and the mean scores on the matching variable are not in agreement with those given in Table IV. Apparently this experiment involves different students than those used in the previous methods experiment. If this is true then the reference to pooling subjects taught by both instructors is not clear. Is the author referring to Table V when he says the variance between instructors and the interaction between instructors and methods was found to be insignificant, or is he referring to a different analysis for which no tables are given? If the former, he is using data collected in one experiment to perform operations in a second experiment, and if the latter, the tables should have been printed to make clear his methods.

Dowell presents the analysis of criterion scores for a second methods experiment involving two methods and three levels. It is reproduced below.

Table VII

Source	d.f.	Variance	Obtained F	Required 1% Level
Methods	1	47.57	1.31	6.78
Levels	2	1606.15	44.3	4.73
Methods x Levels	2	42.57	1.17	4.73
Within Classes	176	36.23		
Total	181			

As the above table indicates a highly significant difference was found to exist between students grouped by initial ability level. Since no mean scores are given for the groups by ability levels the table has little meaning. Moreover, inasmuch as this experiment apparently comprised a different group of students than those in the previous experiment it reveals nothing about the differences that exist between television and conventional instruction and therefore adds nothing to the original purpose of the experiment.

Referring again to the original study:

> The criterion measure was administered as a pre-test to a third group of students immediately before they entered the experimental block of instruction. The pre-test scores for this group were compared with post-test criterion scores achieved by matched experimental students and with matched control students.
>
> The analysis of variance of criterion scores between matched pre-test groups and post-test control groups is shown below:

Table VIII

Value of t for criterion scores between matched pre-test and post-test control groups

	Pre-test Group		Post-test Control Group
No. of students in each group	98		98
Mean Score on matching variables	50.1		50.3
S.D. of matching variable	6.3		6.2
Mean Score on criterion variable	15.36		28.39
S.D. of criterion variable	3.97		7.0
Difference in means criterion variable		13.03	
Standard error of means criterion variable	.349		.530
Coefficient of correlation matching variable & criterion variable	.49		.66
Standard error of difference in means of criterion variable		.635	
t:		20.05	
p:		.001%	

The analysis of variance of criterion scores between matched

pre-test group and post-test experimental group is shown below:

Table IX

Value of t for criterion scores between matched pre-test and experimental post-test groups

	Pre-test Group		Post-test Experimental Group
No. of students in each group	89		89
Mean Score on matching variable	49.3		49.2
S.D. matching variable	6.2		6.2
Mean Score on criterion variable	15.26		26.45
S.D. on criterion variable	3.58		7.4
Difference on mean criterion scores		11.19	
Standard error of means criterion variable	.332		.592
Coefficient of correlation matching variable & criterion variable	.49		.66
Standard error of difference in means of criterion variable		.679	
t:		16.48	
p:		.001%	

The study is even more confused at this point. The author refers to an analysis of variance in reference to the above two tables, but presents figures based on a "t" computation. Were the analysis of variance tables omitted or did the author make a mistake in terminology? Here again the numbers of students does not agree with the number as reported in the previous tables. Evidently these figures are based on studies with different students than the previous experiment. Why were comparisons made between the pre-test and post-test experimental group and the pre-test and post-test control group? It would have been much more meaningful to have compared the post-test experimental and post-test control groups. The present writer computed a "t" value for the difference of the mean scores on the criterion variable for the control and experimental groups from the data given in the above tables, and

found that there was a significant difference beyond the 5% level of confidence in favor of the control group. Considering the ambiguousness of this study, however, no conclusions can be drawn from the above computations. The conclusions for the above study follow in the words of the author:

> The use of live television proved to be equally effective as conventional methods of instruction. Use of the television medium of communication was limited to transmission of lecture-demonstration for viewing by students in remote classrooms. This evidence tends to further confirm the premise that unlimited numbers of students can learn satisfactorily by viewing televised programs of good instructors performing as they normally would in the classroom. This proved to be true for all levels of student ability. The amount of learning that took place in the experimental block of instruction was significant for both the control and experimental students.

In view of the many questionable features and practices in the above experiment, such conclusions are totally unwarranted. The students were not randomly assigned to the experimental and control groups, but "...a control group composed of students equally matched with those in the experimental group was selected from subsequent classes." This practice and the fact that the control group was shown to be superior to the experimental group (Table VI) on the matching variable leads one to suspect that the students were from different populations. The method of matching is not entirely clear in itself. A coefficient of correlation of .66 was obtained between the matching variable and the criterion variable, and this matching variable was then used to match-out subjects to use in the control group. Such methodology seems certain to bias experimental results. Although the author states, "In this experiment, equal results were obtained by matched groups of stu-

dents who received training by the same instructor but by different methods," there is no evidence presented to warrant such a statement. Finally, there is no evidence presented to show that the control groups and experimental groups were actually compared, but instead a third pre-test group is compared with the post-test experimental group and the post-test control group. These differences are found to be highly significant, as one would expect them to be, but this in no way proves that the experimental group performed less well, equally, or better than the control group, and in fact the differences were shown by the present writer to be significantly in favor of the control group on the basis of the data given. The experiment was poorly designed, the procedures were dubious, and the analyses and conclusions were ambiguous and unwarranted.

An experiment in army training is reported by Kanner, Runyon and Desiderato (27). This study was designed to investigate differences between television and regular instruction, differences between kinescope and regular instruction, differences in retention between television and regular instruction, differences between high and low aptitude trainees on learning and retention through television and regular instruction, and the effects of kinescope review on retention compared with effects of no further training for high and low aptitude trainees. Fourteen hours of training were selected from the first eight weeks of basic training. Two experienced instructors were chosen for each of the fourteen hours of training. For 11 of the 14 hours one instructor taught by television while the second instructor taught in the conven-

tional manner.

instructors reve

instructor taught

ously. Basic tra

into two groups,

Test, and were us

Similarly other c

instruction and r

were used in this

structed for the

choice, fill-in,

tests only were g

tion in a large le

into small groups

rooms.

In comparis

groups it was foun

scores on 12 of th

ference was found

were found between

aptitude group. F

were found in 10 o

When retention tes

significant differ

the television gro

tional manner. When a second group was taught the same subjects, the instructors reversed their roles. For the remaining three hours, one instructor taught both the television and the control group simultaneously. Basic training companies used for the experiment were split into two groups, matched for scores on the Army General Classification Test, and were used to compare television and conventional instruction. Similarly other companies were matched for comparisons of kinescope instruction and regular instruction, however only 7 of the 14 hours were used in this comparison. Seventeen criterion tests were constructed for the fourteen hours of instruction, including multiple-choice, fill-in, picture identification, and performance tests. Post-tests only were given. The regular instruction group received instruction in a large lecture hall, and the television group was sub-divided into small groups of 12 to 16 men and received instruction in proctored rooms.

In comparisons of the live television and regular instruction groups it was found that there were no significant differences on mean scores on 12 of the 17 tests. In five of the tests a significant difference was found in favor of the television groups. No differences were found between the television and regular instruction for the high aptitude group. For the low aptitude group significant differences were found in 10 of the 17 tests favoring the television instruction. When retention tests were administered one month after the instruction significant differences were found in four of the 14 tests in favor of the television group. Retention tests were not made for high aptitude

men, but for low aptitude men on retention tests, significant differences were found for 4 of the 14 tests favoring television instruction.

In comparing kinescope instruction and conventional instruction significant differences were found in only one test, and this favored the kinescope group. Using kinescope review prior to taking retention tests resulted in the low aptitude men scoring significantly higher than on their immediate post-tests. The authors conclude that televised instruction was at least as effective as regular instruction, that televised instruction was more effective for low aptitude groups, and that televised instruction was remembered at least as well as regular instruction. The authors also conclude that television effectiveness may be related to subject-matter content. A criticism of this study is the manner in which comparisons were made between companies. Although the groups were matched within companies the fact that the live television, kinescope television, and conventional instruction groups were not selected from the same companies introduced uncontrolled variables such as differences in morale between companies. The fact that the regular group received instruction in a large lecture hall whereas the televised group was divided into small groups of 12 to 16 trainees for one received in a proctored room would in itself account for some of the differences found.

A study conducted by the Educational Testing Service for the American Red Cross is reported by Shimberg (52). The study was designed to measure the differences in learning and opinion of the course under three conditions: (1) instruction by television, (2) instruction

by television plus a weekly practice session, and (3) instruction in a regular classroom group without television. The two television groups received 13 half-hour programs over a period of seven weeks with the practice group receiving in addition a weekly practice session of one-hour. The regular classroom group attended two one-hour sessions weekly for a period of seven weeks. Subjects were members of organized Red Cross groups. The television groups viewed the programs at home. In terms of performance scores the television instruction was found to be as effective as classroom instruction and no differences were found in the two television groups. On the written test the television group did slightly less well than the classroom group and again there were no differences between the two television groups. No data on tests of significance are reported. Inasmuch as the populations in this study were volunteers and no attempt was made to equate or randomize the subjects between groups, no conclusions can be made. Another weakness in this study is the fact that the classroom group received almost twice as much classroom time as the television groups.

Seipman (51) reports an experiment in teaching elementary school music in the public schools. Three-hundred sixth grade pupils in 7 different schools participated in the experiment with 177 pupils in the experimental group and 123 students in the control group. Pre-tests and post-tests were given to all groups. The results of this investigation are reported in terms of the percent of gain between the pre-test and post-test examinations. The average percent of gain for the experimental group was 20.5 and for the control group was 21.8. It was

concluded that the control group was numerically superior to the experimental group. Although the original publication of the study could not be obtained, there are several criticisms that seem warranted. Apparently no randomization nor equating was made of the subjects within the two groups. No attempt was made to control the variable of different schools. The method of analyzing the data was inadequate. No statistical analysis was made to determine if the differences obtained were significant, and although the study lent itself to a covariance design of testing methods by schools by classes the results are merely reported in terms of percentage gained. For the purpose of generalization, the conclusions drawn are not warranted and the facts are inconclusive.

Another study conducted in the public schools was reported by Anderson and Vander Meer (2). The purpose of the study was to compare television and regular instruction on the use of the slide rule. Matched groups from five classes of high school sophomores were formed on the basis of scores on the California Test of Mental Maturity and the Stanford Achievement Test. Forty-one students comprised the television group and another forty-one students were in the conventional group. Six one-half hour programs were taught over a period of 6 weeks. The same instructor was used for both groups. A five-item test was given at the end of each class session and a final examination consisting of the items on the daily tests was given at the end of the series. No significant differences were found between the two groups on the basis of final examination scores, nor were there significant

differences between final scores when analysis was made to take into account differences in sex and intelligence.

A study conducted in the Chicago Public Schools on the teaching of algebra and physics was reported by Willis (60). Nineteen different schools were used in the evaluation of algebra and 24 schools were used in the teaching of physics. Ten lessons were given in all followed by an examination over the subject-matter covered. Instruction was by one instructor for the television sections but by different instructors for the classroom groups. There were no significant differences between the television and classroom groups in either algebra or physics. When the scores were adjusted for the ability levels of the students it was found that the television students in both courses did slightly better, but the difference was not significant. The report states that instruction by television is slightly more effective than instruction in the regular classroom. This conclusion is not justified on the basis of the reported data. The study may be criticized in that comparisons were made without attempting to control the variable of different instructors and different schools. The results are inconclusive.

Comparisons between classes under four different conditions of instruction is reported by Husband (25). The subject was a course in the psychology of adjustment at Iowa State College and the four conditions were: (1) television instruction received at home, (2) television in the classroom received by a monitor, (3) kinescope instruction at a later time, and (4) normal classroom instruction. The instructor was the same for all groups. The television sessions were 30 minutes in length while the classroom sessions were 50 minutes in

length. All groups received the same number of sessions. The group receiving television at home were not regularly enrolled college students, their median age was 37, and the length of time since graduation from high school was a median 20 years prior to taking the course. The other students were regularly enrolled college students. Grade point averages for the course were used for comparisons. The kinescope group did the best, the group receiving television at home next best, and the other groups were next in achievement. No tests of significance are reported. Inasmuch as the comparisons were made between groups receiving differential amounts of instruction, and between groups that were obviously not from the same population no conclusions can be made from this study.

An interesting experiment by Williams (59) was designed to compare instruction by television, lecture, radio, and by reading. One hundred and eight undergraduate students were divided into four groups so that each group contained an equal number of high, average and low ability students. Each of the four groups was assigned to one of four classes: (1) lecture, (2) television, (3) radio, and (4) reading mimeographed copies of the lecture. The subject "Thinking Through Language," was not familiar to the students. The same lecturer presented his material simultaneously to the lecture, television and radio groups, while the reading group read the lecture at the same time. The lecture group was in the studio where the lecture originated. Examination consisted of 19 multiple-choice questions and an essay-type question to be answered in 200-300 words. The results were tested by analysis of variance and "t" tests on the objective part of the examina-

tion only. It was found that television instruction was superior to radio instruction beyond the 1% level of significance, that radio was significantly better than reading at the 5% level, and that there was no significant difference between the reading and lecture groups. Testing the data on the basis of academic ability, it was found that the same order obtained for the television, radio and reading groups. However, the lecture group was last in the amount learned by the high and low ability groups, but as high as television for the average groups. This study is well designed and its principal weakness, as the author points out, is the fact that the lecture group was not a "true" lecture group, but rather a studio group and was undoubtedly influenced by the distractions of lights, cameras and equipment.

Paul and Oglivie (45) conducted a follow-up of the Williams study reported above, eight months after the original study. Using the same multiple-choice test as was used in the first study, they administered the test to 74 of the 108 original students. An analysis of variance of the results showed that the television group was still highest, that the studio group (previously last) was second on the retention test, followed by the radio and the reading group. Again the authors in this study state that no conclusions can be made about the studio group, because of the conditions obtaining in the studio.

Another study conducted on the university level was reported by Evans, Roney and McAdams (19). Comparisons were made of two classes. In an elementary psychology class three groups of students were compared: (1) ninety-six students in an on-campus lecture section, (2) a television lecture session plus correspondence work consisting of 17

subjects, and (3) thirty students enrolled in a television lecture plus discussion section. In the biology course there were 78 subjects matched for college class, grades in college and sex. The criterion in the psychology course consisted of scores obtained on an 150 item multiple-choice type examination administered at the end of the course. Mid-semester grades on a 70 item test were used as the criteria for the biology course. One instructor taught each course for all groups. No significant differences were found between groups in either course. Inasmuch as no attempt was made to randomize or equate students within groups in the psychology course, it is impossible to draw any conclusions from the experiment in that course. In the biology course the practice of using mid-term grades rather than final examination grades is questionable. It is conceivable that different results could have been obtained had final examination scores been used.

Pasewark (44) conducted a study to determine the effectiveness of television as a medium in teaching typewriting. Forty-four college students were split into two matched groups on the basis of scores received on the American Council for Education Psychological Examination. One group received instruction by television and the other group received instruction in the conventional classroom. The same instructor taught both groups. Instruction was given one-half hour per day, five days a week for a total of 48 days. Achievement was measured by a timed typing test given at the end of the course. Students in the television section typed significantly faster than did students in the regular class section. There was no significant difference in the number of mean errors between the two groups. The only criticism to be

made of this study is that in the television section there were no proctors or instructor, whereas the instructor was present in the conventional classroom. The absence of an instructor or proctor may have provided less disturbance and thereby enabled the television group to perform better in this type of course.

A study reported by Tannenbaum (55) was conducted at the University of Illinois Medical School. Students in a basic physiology course were split into two groups equated on the basis of mid-term grades. Three 50-minute lectures were given to the two groups of students. One group received conventional lecture in the presence of television cameras and the second group received instruction by means of monitors. One week after the lectures an examination was given consisting of 19 multiple-choice items. Analysis of variance was used to analyze the results. The television group was found to be superior to the lecture group at the 7% level of confidence. Although this is not the usual level of acceptance, the author suggests that the novelty of television learning may have accounted for the difference. The author suggests also that "nearness" to the instructor was a factor in learning reasoning that the television students were "nearer" to the instructor than those students sitting in the rear of the lecture room, however such a hypothesis was not supported.

A series of studies conducted at Purdue University (13) in closed-circuit television instruction was carried on in classes in political science, analytic geometry, general chemistry, and general bacteriology. In only one of these courses, that in political science, were the students randomized between the television and conventional

classroom groups. In the other three courses comparisons were made with other sections enrolled in the same course. In none of the courses were significant differences obtained between students in the two groups. These were considered preliminary studies and the results are not conclusive.

Studies of Student Opinion of Television Instruction .

In the previously reported study by Rock, Duva, and Murray (48) an attempt was made to evaluate the television instruction by asking the participants to make comments about the course. It is reported that comments by the trainees were favorable to the course. The staff felt that more visual emphasis should be made and the talkback microphone was felt to be inadequate because many trivial questions were asked. No objective data was obtained for the basis of making a statistical study of the evaluation.

In the second study by the same authors (49) a series of questions were asked the reservists about the course. Four-fifths of the group reported the series as interesting or very interesting, and a majority of the group said the series was good or excellent. About three-fourths of the group felt that they would rather be taught by television than by the regular method. Again no objective data or analysis is made.

In the study by Allen (1) a questionnaire was used to evaluate the course. A majority of the ROTC students reported that the television instruction was as interesting or more interesting than other types of instruction. Thirty-one out of fifty-three said the material

was

lear

visi

anal

who

this

in t

mate

key

post

with

tele

enro

cours

would

gener

over

tech

diffe

was about as easy to learn as regular instruction, 11 thought television learning was easy or very easy, and 11 stated that learning by television was more difficult than conventional instruction. No statistical analysis was made of the results of the questionnaire.

Shimberg (52) in the American Red Cross study states that those who viewed the television instruction were overwhelmingly in favor of this type of instruction.

In the study reported by Willis (60) on teaching by television in the Chicago Public Schools, the students reported that too much material was presented by television, but they favored the summary of key points at the end of the lesson. They felt that television was most effective when followed by discussion afterward in the classroom with the regular class teacher.

Evans attempted to evaluate the attitude of students towards television by a questionnaire (19). He found that 70% of the students enrolled in an elementary psychology course would enroll in another course taught by television, 16% were undecided, and 13% said they would not enroll again in such a course.

In the Purdue study previously cited (13), the students were generally favorable to the demonstrations in chemistry and bacteriology over television, but in the courses in political science and analytic geometry the students were not so favorable. No statistical tests of differences were made.

III. SPECIFICALLY RELATED STUDIES OF INSTRUCTIONAL TELEVISION

Only a few studies have been designed to measure the effective-

ness of television instruction in relation to classroom discussion. All of the studies in which this has been one of the purposes have been sponsored by educational institutions, and in the majority of instances these have been universities and colleges. These studies have been mainly concerned with achievement and retention of course content, and with student opinion of the course. A review of several studies closely related to the present study will be examined in the following section.

Studies of Achievement and Retention of Course Content

Kumata (31) reports a study by Parsons designed to compare outcomes in a psychology of child development course under three different learning conditions. Forty university students were used as subjects and were randomly assigned to all experimental groups. Twenty students were assigned to a control group with no instructor. The remaining twenty students were assigned to three groups. One group was designated as a kinescope-correspondence study group without an instructor; a second group was composed of students studying in a conventional classroom with an instructor; and the third group was an independent correspondence study group. All subjects used the same textbook, a workbook, and manuals. The kinescope group met weekly to view the kinescope lectures and the classroom group met twice a week with the instructor at which time class discussions were held in addition to the lecture. Standard examinations were given to all groups. Pre-tests, post-tests, and a retention test was given approximately four months after the close of the course. There were no significant differences among the three groups on pre-tests and post-tests and all three groups were signifi-

cantly higher than the control group on the post-test. On the retention test the correspondence group was significantly higher than the classroom group, and the kinescope group scored lower than the correspondence group and higher than the classroom group, but was not significantly different from either. An evaluation of all written work for the course revealed no differences in any of the groups.

A measure of group cohesiveness administered to the kinescope and classroom groups only--since the others had little contact--revealed no significant differences. A sociometric test designed to measure the group structure showed that the total number of choices increased significantly from pre-test to post-test administrations for the classroom group, and decreased slightly but not significantly for the kinescope group.

The original copy of this study could not be obtained so it is impossible to make a critical evaluation of the study. The study appears to be well designed and carefully controlled except for one feature. The classroom group met twice a week with the instructor while the kinescope group met only once a week without the instructor. This feature of the study introduced a variable that could not be controlled. The author offers no explanation concerning the superior retention of the correspondence group. It may be that this group being an independent group became more highly motivated during the course of study, and reinforced their learning during the period intervening between the post-test and the retention test. This finding might also be explained by hypothesizing that overlearning occurred for this group because of

the manner in which they were required to study the course, therefore their rate of forgetting was less.

During the year 1955-56, experimentation in the use of closed-circuit television was conducted at New York University (46). The purpose of the experimentation was to learn how to make the best use of television in college instruction, to compare the quality of instruction in televised and non-televised courses, and to investigate costs of instruction using television. Two courses were selected for experimentation, a freshman College Composition course and Literature of England--a sophomore course. Students were those who regularly enrolled in the course. No attempt was made to randomize students within the two groups, and comparisons were made between several sections in each course, some of which received television instruction and some of which received normal instruction. The composition course met three times per week, receiving a televised lecture for two of the meetings while the third hour was a "tutorial" hour during which an instructor worked with the students on preparation of papers and to lead class discussions. The normal classroom groups met three times per week with their regular instructor. The literature course was handled in the same manner. The television lecturers were members of the staff. No attempt was made to have one lecturer for all programs, but instead lecturers were alternated so that as many as eight different people participated in the lectures for one term.

Achievement was evaluated by comparing final grades at the end of the instruction with final grades of students in non-televised sec-

tions during the same day and term. It was concluded that students performed about equally well in the television and non-television sections. It is suggested tentatively that the average and poor students in the composition course under television instruction did not perform as well as the average and poorer students in the conventional sections. In the literature course, however, the average and poorer students did better than their counterparts in the non-television sections. As the report states, however, these conclusions are merely tentative and suggestive, and more experimentation is needed with a more carefully controlled evaluation of results.

It is difficult to make a critical review of the above study inasmuch as it was not considered to be a final conclusive experimental study, but more nearly a "pilot" study to precede a more carefully controlled study. It was not a carefully controlled study: subjects were not randomized nor equated, the instructor variable was uncontrolled, the criterion measure was not adequate, and no statistical measures were used to analyze the results. Even as a preliminary study it could have been considerably improved.

A carefully planned and conducted study was made at Pennsylvania State University in 1954-55 (12). This experiment had several purposes but for the present study the discussion will be concerned with comparisons of the relative effectiveness of conventional instruction with the same instruction presented over closed-circuit television. Three courses were used in the study; General Psychology, General Chemistry and Psychology of Marriage. Four sections of General Chemistry were

used in the experiment. Students in these sections were matched on the basis of interests (major curricula) and grades in the preceeding chemistry course (Chemistry 1). Two sections of 120 students each were designated as television sections, and two sections of 100 students each received the lecture in the regular manner. Three lecturers were used in such a manner that the students in all sections were exposed to all of the lecturers an equal number of times and for the same subject-matter area of the course. One of the experimental groups was subdivided into four viewing classrooms, and one met in the lecture-originating room with the television equipment present.

The students in General Psychology were randomly assigned to experimental or control groups. Two instructors were involved. One instructor lectured to a group of 40 students in the television originating room, and to two classes of 40 students in the television receiving rooms, while a second instructor lectured to a control group of 40 students in another classroom. Later in the day the roles were reversed for the same procedure with a different group of students.

The Psychology of Marriage course was not originally included in the experiment, but enrollment was so large that it was decided to divide the class into four equal groups of about 30 each with one group in the originating room and the others in the receiving rooms. There was no control group in this course. One lecturer was used for the entire course.

Proctors were provided for all television viewing sections whose duties were to take roll, adjust monitors, and perform other procedural

duties, but who were instructed to do no teaching nor to answer questions. Questions were referred to the regular lecturer.

The criterion measures in the courses consisted of examinations given throughout the course and the final examinations, all of the objective type. In addition a measure of student attitudes towards the courses were given, and these will be discussed in the next section of this review. (See page 58) Analysis of results was computed by analysis of covariance and analysis of variance techniques. Reliabilities of tests were computed by Kuder-Richardson or by analysis of variance of item scores. In addition to the above measures, an Opinion or F Scale was given to the students in General Psychology; The Marriage Happiness Prediction Inventory was given to the students in the Psychology of Marriage class; and a Student Reaction Schedule was administered to all students.

Except for one test administered during the term there were no significant differences found in any of the courses on the basis of achievement test scores. The study suggests that differences may have existed but that the measured samples of total learning may have lacked the appropriateness and sensitivity to detect the full differences. It is also suggested that sources of information such as library books, informal student discussions, and other resources may have reduced the probabilities of getting measurable results. It was also proposed that the television classes may have assumed more responsibility for their own learning and hence compensated for any differences in the two modes of instruction. The general conclusion made was, "It is logical to con-

clude that there is no basis in the evidence found on information learning for rejecting the use of instructional television for teaching courses and students like those used in this experiment." It is also concluded that the evidence available indicates that decisions to use or not to use instructional television in such courses as were studied and for defined student populations must be made on the basis of administrative policies, acceptability of televised instruction to students and faculty, and other practical considerations.

The above study has much to commend it. It has the following strong features:

1. The problem was clearly defined and was framed in the form of objectives which were experimentally verifiable.

2. The statistical design was appropriate to the conditions.

3. The population and the sample was clearly specified and the manner of drawing the sample was clearly indicated.

4. The control groups were chosen from the same population and in the same manner as the experimental groups.

5. The level of significance for test was set at the beginning of the experiment.

6. The treatments and methods of collecting the data were adequately described.

7. Extraneous sources of error were controlled.

8. The criterion measure reliability was given.

9. The conclusions were consistent with the results obtained.

10. Generalizations were confined to the populations and condi-

tions of the experiment.

The study could have been improved in that it displayed the following weaknesses:

1. The experiment did not include a replication.

2. The characteristics of the sample were not fully described.

3. The assumptions underlying the statistical tests used were not tested.

In spite of these limitations, the Pennsylvania State University study is one the most complete and carefully executed studies that has been attempted.

Another thorough and well-designed study is reported by Miami University of Ohio (39). The purpose of the study was primarily for studying the effectiveness of certain types of large group instruction and to demonstrate the feasibility of these procedures at the college level. Courses included in the study were courses taught by television, large courses other than television courses, courses taught by graduate students, and conventional (control) courses consisting of 25-35 students. The first three types were considered as experimental courses. The control sections were taught by the same instructor who taught the corresponding television or large class section. The control sections in the graduate student phase of the study were taught by full-time faculty members, but were of comparable size to the graduate student assistant sections. The variables employed for equating students included the scores on the Cooperative Test of English Achievement, Cooperative Mathematics Placement Test, and the American Council for

Education Examination for College Freshman. The equating was completed after registration and the experimental and control groups represented only samples of the total enrollment. Neither the instructors nor the students were aware of the students who were not included in the sample. It is important to note also, that decisions about collection of data and sampling technique were made prior to the collection of data. Most of the courses were full year courses, and the same procedures were employed during the second semester as were used in the first semester except that students to be included in the sample the second semester must have been used during the first semester. Evaluation was made over four areas: achievement, student reaction to the course content and the instructor, student attitudes about instruction, and instructor attitudes about teaching television and large course sections. As with the previous study only the achievement area will be considered leaving student attitudes for the next section. Criterion tests for achievement were measured by subject-matter knowledge tests; tests on synthesis, problem-solving, and critical thinking; course related tests, such as "Stereotypes in Social Studies" and "Misconceptions in Psychology." Some of these measures were objective and others were essay tests. Essay tests were graded by multiple readers after precautions were taken to guard against "halo effects". Graders were not aware whether the essays were for students in the experimental or control sections. Reliabilities were computed for all tests. Analysis of covariance and "t" ratios were used in analyzing the data.

The findings relative to the achievement area are as follows:

In the acquisition of subject-matter it was found that in general, students in television sections perform about as well as students in control sections. Exceptions to this were observed in the second semester of full year courses which was due possibly to the motivational decline as the novelty of television instruction was dissipated. Achievement in large course and control sections was about the same. In the comparisons of achievement in critical thinking, problem solving, and synthesis it was found that television instruction was significantly inferior to conventional instruction in one course (Economics) but that no differences existed in the other courses. In the large course sections no differences were found as compared to control sections, except again in one economics course. The achievement of course related attitudes was not investigated in the television sections. In the large course sections differences favoring the conventional type of instructions were found in only one course (Economics), and in all the others there were no differences.

In the comparison of large course sections and conventional sized classes, it was found that the large sections did not do as well in the Introductory Psychology course. This finding was also true for the high ability students in Business and Government, but not for the low ability students for they perform as well in large classes as in conventional size classes. In the Social Studies course there were no differences in achievement. In analysis of achievement results in which the interaction of academic ability and type of instruction was considered there were no differences reported except in one section. In general it was

concluded that no interaction existed.

This study has many good characteristics.

1. The problem was clearly defined.

2. The problem was stated in the form of questions which were experimentally verifiable.

3. The statistical design was appropriate.

4. The method of drawing the sample was clearly specified.

5. The control groups were chosen in the same manner as the experimental groups.

6. The treatments were randomly assigned.

7. The level of significance was specified prior to experimentation.

8. The methods of collecting data were adequately described.

9. Extraneous sources of error were held constant or randomized.

10. The criterion measures were appropriate.

11. The reliability of criterion measures was given.

12. The conclusions were consistent with the results.

13. Generalizations were confined to the setting of the experiment.

The study has the following weaknesses:

1. The population was not clearly defined.

2. The characteristics of the sample were not given.

3. The experiment did not include a replication.

4. The assumptions underlying the statistical tests were apparently not made or tested.

In spite of these weaknesses the Miami University study is one of the most thorough and comprehensive to date because there was an attempt to find the relationship between ability and achievement in large and small classes.

The final study to be considered in this section is one conducted at the State University of Iowa during the academic year 1956-57 (5). American Government, a three hour course meeting for three fifty-minute sessions per week, was selected for the experiment. The course is required of majors in the department of political science as well as students in other areas particularly those majoring in teacher training. It is open to all classes, but the largest enrollees are Freshmen and Sophomores. The study was designed to answer several questions:

1. Can knowledge of and certain attitudes toward American Government be taught better by the discussion method or by the lecture method?

2. Does the method of instruction differentially affect the acquisition of knowledge and attitudes of students at different levels of academic ability?

3. What methods of instruction are preferred by instructors of the American Government course?

4. What methods of instruction are preferred by students in the American Government course?

5. Can a class with an enrollment of sixty to eighty be taught by the discussion method and achieve the same results in course examinations as a class of twenty to twenty-five taught by this method?

6. Can an instructor, using closed-circuit television, promote more discussion and secure better results in course examinations with the same number of students as in a conventional classroom situation?

7. Will the performance of students who can see and hear a television discussion, but who cannot participate orally in it, be differently affected as compared to those of students who can participate?

The treatment groups consisted of five types of sections:

1. Television discussion sections composed of a comparatively large number of students--about sixty to ninety--and divided into three groups. Group One consisted of fourteen students in the television originating room whose voices and images were transmitted to the viewing rooms. Groups Two and Three were in two of the viewing rooms and could communicate with the instructor and students in the originating room vocally but not visually. When they wished to enter the discussion a signal was sent to the instructor via remote control and he could then call on them. Students were rotated among the three groups approximately every two weeks, and all students spent approximately an equal amount of time in each group.

2. Television observation sections composed of twenty to thirty students who could see the instructor and the students in Group One in the originating room, and could hear the discussion between Groups One, Two, and Three. Students in this group, however, could not participate in the discussions.

3. Small group discussion sections varying in size from eighteen

to twenty-nine. These sections were considered the control.

4. Large group discussion section composed of seventy-five students.

5. Lecture sections composed of one hundred and twelve students one semester, and one hundred and thirty-two students the second semester. No systematic discussion of topics was allowed by the size of this section although the instructors allowed questions and interruptions during the class hour.

Random assignments of students to the various sections could be followed only in the case of the television observation and television discussion sections. Assignments of the other students depended on the hour at which the course was chosen. The assumption was made and tested that students who registered for the course at different times of the day did not differ significantly on any of the variables that would affect the criterion measures.

Four teachers were involved in the experiment. One instructor taught all of the small group discussion sections and one of the large lecture sections during the first semester. A second instructor taught a lecture section the second semester. The third instructor taught a lecture session each of the two semesters. The large group discussion sections and the television sections were taught by the fourth instructor. Although this arrangement was an obvious weakness in the experiment, scheduling difficulties necessitated such an organization.

In order to promote discussion and to provide a common core of topics covered in the course, a syllabus including assignments, objec-

tives, and questions which served as the basis for discussion on any given day in the discussion sections was prepared by the members of the department.

Evaluation was based on four criteria:

1. Achievement as measured by two mid-term and one final examination. Each mid-term examination consisted of forty multiple-choice items plus an essay question. The final examination for the first semester was composed of eighty multiple-choice questions and one essay question. The second semester final examination was made-up of seventy multiple-choice items plus one essay question.

2. Attitudes toward concepts of "liberal democracy," as measured by an instrument under development by one of the faculty members of the political science department. The measure contains forty-eight Likert-type questions.

3. Opinions concerning the various methods of instruction and the perceived effects of each method. The opinions were appraised on a questionnaire.

4. Attitudes of the discussion instructors toward these methods of instruction. These were measured on the basis of subjective impressions submitted by the instructors.

Replication of the study was obtained by comparing the results of the study for two identical semesters.

In order to test the assumption that the various groups did not differ significantly on any of the variables assumed to affect the criterion measures, the homogeneity of the groups on the basis of the En-

trance Composite Percentile Rank, which each student had received when he entered the university, was tested by analysis of variance. In none of the groups was a significant difference obtained and the assumption of homogeneity was accepted. The study does not explain what tests are used for computing the Entrance Composite Percentile Rank.

An analysis of variance of mean scores on the three achievement examinations and the total of the three examinations revealed no significant interaction between the measures and the sections and no differences among sections on achievement tests. The only significant difference was that between measures--tests--and this was expected inasmuch as a much higher score was possible on the third achievement examination.

In order to examine the relationship of ability level with the method of instruction, a "treatments by levels" analysis of variance was made. Two separate analyses were made using final examination scores as the criterion. In one analysis, the Entrance Composite Percentile Rank was used as the control variable and in the other the Pretest examination score was used. Three levels were used for each of the control variables, the top quartile, the two middle quartiles, and the bottom quartile. The analyses revealed no differences and the hypothesis of no differences between the effectiveness of the various methods of instruction was accepted. A significant difference was found between the various ability levels as was expected.

It was concluded that achievement in the American Government course, as measured by the mid-term and final examinations for the

course was improved by all methods of instruction, and there were no differences due to the method of instruction.

Analysis of the results on the attitude criterion showed no significant differences between groups during the first semester. The attitude data for the second semester was analyzed by levels using the pre-attitude score as the control variable and the post-attitude score as the criterion measure. A significant difference between levels was found, as was expected. The differences between groups were also found to be significant. One of the lecture groups obtained the highest mean and the lowest was obtained by the television observation group. None of the other groups differed significantly. Inasmuch as only one of the two lecture groups was significantly different, and a comparison of these results with first semester results showed no similar trends, it was suggested that this difference was probably a chance effect. The authors conclude that there were no significant differences between methods of instruction on the learning of attitudes towards the concepts of "liberal democracy" as measured by the criterion instruments used in this study.

The above study has much to commend it.

1. The problem was clearly defined, and was framed in the form of questions which were easily transferrable to null hypotheses.

2. The statistical design was appropriate to the methods and hypotheses under test.

3. The population and the method of drawing the sample was clearly specified.

4. The experiment was replicated.

5. The level of significance for rejection of the null hypotheses was specified before the data was collected.

6. The treatments and the method of collecting data was adequately described.

7. The criterion measures employed were appropriate to the purposes of the study.

8. The statistical assumptions underlying the tests of null hypotheses were satisfied.

9. The conclusions were consistent with the results, and generalizations were confined to the population from which the sample was drawn.

The following weaknesses in the study should be indicated:

1. The control group was not chosen in the same manner as the experimental groups.

2. The characteristics of the sample were not adequately described in that the various curricula, classes, age and sex of the sample was not given.

3. The fact that different instructors were used in different groups introduced an extraneous source of error that was not controlled.

4. There was no evidence that the reliability of the criterion measures was computed.

Although the subjects in this study were not randomly assigned to all groups, the fact that all subjects were tested for homogeneity of variance on the Entrance Composite Percentile Rank and found to be homogeneous indicates partial control of this variable. The fact that

one instructor taught the television sections and the large group discussion sections lessens the seriousness of the criticism of the uncontrolled instructor variable, for the comparisons between the above groups.

Studies of Student Opinion of Television Instruction

In the Parsons study (43) previously cited, student evaluation of the course was rated on a three point scale. A score of 1.0 on the scale was defined as "more valuable", 2.0 was the midpoint, and 3.0 was defined as "less valuable." Mean ratings for the classroom group were 1.0, the correspondence group scored a 1.4 and the kinescope group averaged 1.7. No tests of significance were reported for these figures.

A student questionnaire was used to assess student opinion about the course in the New York University study (46). Although no objective statistical summary was made of the results, the following general summary is quoted from the study:

> 1. There was no change during the year in the students' attitude toward the usefullness of television as an educational device except for an increase in the favorable attitude of the Literature of England group. Both groups were favorable toward the use of television as an educational device.
>
> 2. The College Composition group shifted during the year toward a preference for one instructor. There was no change in the Literature of England course. Both groups were preponderantly favorable to the idea of several instructors.
>
> 3. Both groups indicated dissatisfaction with the course. During the year the College Composition group shifted to an even more negative attitude.
>
> 4. There was a tendency for students generally negative to commercial television to be less satisfied with the course and more negative toward television as an educational medium. In general, they were less favorable toward all aspects of the program.

5. No difference occurred between the Composition and Literature groups with respect to their satisfaction with the course, and in general, there were no significant differences between these two groups with respect to the whole program.

6. No differences occurred between any of the groups with respect to the felt 'closeness' of the subject matter. The majority did feel that television prevented close contact with subject matter.

Several measures were used to evaluate the students' attitudes toward television instruction in the Pennsylvania State University study (12). One of these, a Student Reaction Schedule, was administered to the students in the television receiving room only. The students were asked to rate their estimated amount of learning and interest in the course compared to what it probably would have been in a conventional face-to-face classroom situation. They were also asked to write a paragraph explaining the reason for their ratings, and the advantages and disadvantages of the course. Although a majority of the students estimated they were learning "about the same" or "a little less" by television no clear trends were evident. This scale was used in all three courses (General Psychology, Psychology of Marriage, and General Chemistry) and the results were similar from course to course.

In the General Psychology course the students were asked to complete a questionnaire purportedly distributed by a "Curriculum Survey Committee" in which they were asked to rank all of the courses they were taking during the semester of the experiment in the order in which (1) the course was contributing to their academic objectives and (2) the courses were liked by them. The results of this rating showed that the students in the conventional classroom ranked the course significantly

higher than did students in the television receiving room both in the contribution of the course and in their liking for the course. No differences were found between groups in the originating classes and those in the standard classes.

A Personal Relevance scale was administered to the students in the General Psychology course. This scale was a Guttman-type scale and was designed to measure the perceived immediacy or remoteness of the use of the information by the students. Analysis of variance of the results revealed no significant differences between the three methods. There was a significant interaction between the method of instruction and the instructor. The authors suggest that more investigation needs to be done before definite conclusions can be made about this interaction.

An F Scale designed to measure "authoritarianism" was also given to the students in the General Psychology course. No significant difference was found among the experimental and control groups on this measure.

The strengths and weaknesses of the Pennsylvania State University study that were listed in the section pertaining to the achievement of course content can also be made of this part of the study. In general it is an excellent study.

In the study at Miami University (39) which was previously cited, several evaluative criteria were used to measure the student's attitude about the course and the instructor, and about television and large class instruction.

1. The C-Scale, designed to measure the students' rating of the course, was a Thurstone-type attitude scale with a nine-point continuum. Five was the neutral point on the scale; scores below 5.0 were favorable and scores above 5.0 were unfavorable.

2. The I-Scale was used to rate the instructors. It consisted of twenty-four items of teaching effectiveness about which the students rated the instructor. Scores below 5.0 were favorable and scores above 5.0 were unfavorable.

3. A TV-Scale required the students to evaluate the effectiveness of television instruction in comparison to conventional small class instruction.

4. An LC-Scale was used to compare the effectiveness of large class instruction with conventional instruction.

5. A TV-LC-Scale required the students to compare television and large class instruction.

The TV and LC Scales were administered in all courses at the end of the first semester and readministered in certain courses at the end of the second semester. The results lead to the conclusion that as a group students assigned to television and large course sections prefer assignment to a conventional section. In two of the courses where the scales were administered at the end of both semesters it was found that the attitude toward television at the end of the second semester was less favorable than at the end of the first semester. In one course the students were exposed to television instruction one semester and large course the next semester with the same instructor. In comparisons

of these modes of instruction and of small class instruction, it was found that the students favored small classes over large classes and favored large classes over television classes. The study reports the following conclusions regarding student attitudes about television and large class instruction.

1. Students assigned to TV or LC (large class) sections generally do not like them as well as conventional (small) classes.

2. It is possible to teach a TV course in such a way that students actually prefer it to conventional instruction. This, however, requires a unique combination of instructor, course material, and the full use of the potentialities of television as an instructional medium.

3. Students in at least one course (wherein the problem was investigated) preferred large class instruction to TV instruction, although they tended to prefer small class instruction to either large or TV classes.

4. Students in TV courses tend to become disenchanted with television as a means of instruction during the course of the year. Most students reported that they neither learned as much nor were as attentive during the course as they had originally anticipated. Students in LC courses, however, are better able to anticipate their end-of-the-year reactions than those in TV courses.

5. Attitudes about the means of instruction are much more variable between TV courses than between LC courses. The prevailing attitude toward LC instruction as compared to control instruction is mildly unfavorable. The prevailing attitude toward TV instruction as compared to control instruction ranges from strong enthusiasm to extreme displeasure.

6. The instructor is a major determinant of how students will react to TV and LC instruction. There is a pronounced tendency for students who dislike their instructor to dislike TV (or LC) classes and vice versa.

7. The majority of students would enroll in a TV or LC section (even though they prefer small classes) if it meant that they would be assured of being taught by an excellent instructor.

8. In general, attitudes about TV and LC instruction are independent of level of academic ability. Two exceptions to this generalization were apparent during the spring 1956 semester wherein an

inverse relationship between academic ability and attitude about TV instruction was obtained.

The analysis of the results obtained after administration of the C-Scale and I-Scale are summarized in the study as follows:

1. Blanket generalizations about student attitudes regarding the worth of a course as a function of class size are not justified. Other factors, including course content and the ability of the instructor to handle larger groups of students interact with class size to affect these attitudes.

2. Student motivation and interest in the subject-matter is not significantly diminished when the course is presented on television or in large classes.

3. There is a pronounced tendency for instructors to be rated as more effective when they teach conventional or small sections than when they teach TV or large classes. Again, however, this generalization does not hold for all instructors. Some teachers are able to teach large groups as effectively as smaller ones.

4. Some of the specific "intangible" benefits often associated with a low student-instructor ratio need not be sacrificed as a result of large group instruction. These intangibles are, however, achieved somewhat more satisfactorily by faculty members than by graduate assistants.

The Miami University study was critically evaluated as to the achievement portion of the experiment in the preceeding section of this chapter. The strength and weaknesses indicated in that evaluation apply as well to the above portion of the study dealing with student attitudes. The only criticism that can be made in addition to the above is that the data and the design were well adapted to an analysis of variance test, instead of the "t" tests that were used.

The final study to be reviewed in this section is the State University of Iowa experiment (5). Student Opinions concerning the various methods and the effects of each method were gathered during the first semester only. In analyzing the opinions of each group toward the alter-

native methods of teaching the course, two results were shown. One was the importance of some form of discussion for the students and the other was the inclination of those students who had some familiarity with television instruction to prefer discussion by this method. A comparison of the groups on the estimated relative preparation required for the course taught by each method showed that the students in the discussion sections felt that the course required more preparation than did those in the lecture section, the television section or the television observation section. The differences were significant between all groups.

In reply to a question about which group had learned the most, it was found that the students in the small discussion section felt that they had learned the most and those in the television discussion were only slightly lower than the small discussion group. Students in the television observation group felt that they had learned the least. These differences were all significant at the 5% level of confidence. There were also significant differences between the groups on the extent to which they felt motivated to greater thought in comparison with comparable lecture and discussion courses.

The major results as summarized in the study are as follows:

1. Students in American Government preferred the course taught by the discussion method or a combination of lecture and discussion rather than lecture alone.

2. Students who had experienced a course taught by the discussion method on television were more favorable toward the use of television for instructional purposes than were students who had not had

this experience.

3. The students' ratings or self-evaluations of the amount they learned as well as the degree to which they were stimulated by the courses were positively related, in the television sections, to the students' opportunities to participate.

One of the weaknesses of this study, as indicated earlier, was that of not controlling the instructor variable. In making comparison of discussion sections with other types of instruction, this is an important criticism. The ability to use the discussion method is more closely related to personal characteristics of the instructor rather than the ability to enumerate factual data characteristic of the lecture method. It would seem, therefore, that the results on the opinion criteria can be only accepted with reservations until further verification is made.

IV. SUMMARY

The literature about instructional television reviewed in this chapter was concerned with general background studies in which the efficacy of television as a medium of instruction was explored, and with specifically related studies in which television instruction was used in discussion methods of instruction. The results of the review are summarized in Table I, page 65.

Inspection of Table I reveals that of the general background studies in which the achievement of course content was used as a criterion only four of the sixteen studies were satisfactory in terms of

TABLE I

NUMBER OF STUDIES REVIEWED, CLASSIFIED BY CONCLUSIONS, ADEQUACEY OF METHODOLOGY AND PURPOSE.

Purpose	Findings Favorable to Television Instruction		Findings not Conclusive		Findings Not Favorable to Television Instruction	
	Critical Evaluation		Critical Evaluation		Critical Evaluation	
	Satis-factory	Unsatis-factory	Satis-factory	Unsatis-factory	Satis-factory	Unsatis-factory
General Background Studies						
Achievement of course content as criterion	2	5	2	5	0	2
Opinion of course as criterion*	0	5	0	2	0	0
Specifically Related Studies						
Achievement of course content as criterion	0	0	3	2	0	0
Opinion of course as criterion*	0	0	2	1	1	1

*The studies reported in this row are parts of the studies reported in the row above using the achievement criterion. Not all of the studies used both an opinion and achievement criterion hence the difference in numbers.

experimental procedures, methods, techniques and analyses. Of these four, two were found favorable to television instruction and in two the findings were inconclusive. In the seven studies in which the opinion of the course was used as a criterion, all were found unsatisfactory in terms of experimental procedures.

In the review of studies specifically related to the present study in which the achievement of course content was investigated, three of the five were found satisfactory and the findings were inconclusive. In the same five studies in which the opinion of course was used as a criterion, two were found to be inconclusive and in one the findings were unfavorable toward television instruction. The other two studies were unsatisfactory in terms of experimental procedures.

Because of the many differences in designs of the studies reviewed, the diversified backgrounds of the subjects, the variety of subject-matter courses used in the studies, and the varying conditions prevailing from one study to another, it is difficult to make specific generalizations about television instruction. Probably the most that can be said at the present time is that there is no conclusive evidence to show that television instruction is either more or less satisfactory than ordinary classroom instruction.

In spite of such findings many educators and writers have overlooked the weaknesses of the many studies of instructional television and have assumed that if television is not a better medium of instruction than the more conventional methods, it is at least just as good and can be used to reach a greater number of people. Such conclusions

have resulted in the installation of television equipment in many schools and colleges as a solution to the problems of teacher shortages. Such practices may be justified where television is to be used for the transmission of factual information, visual demonstrations, movies and the like, but there is no clear evidence to indicate that television can substitute for conventional classroom methods where discussion methods of instruction are the principle ways of transmitting information and teaching critical thinking, problem-solving, and the like.

It is clear that more experimentation is needed along these lines in the use of television instruction. Inasmuch as discussion techniques are an important part of teaching, the present study is an attempt to evaluate the use of instructional television where discussion is used in the teaching methods.

The design of the present experiment will be described in Chapter III.

CHAPTER III

THE DESIGN OF THE EXPERIMENT

There are a number of statistical principles and various designs available for the modern experimenter in education. The present experiment employed a four-way factorial design. The use of such a design made it possible not only to determine the effects of the two discussion methods, but also to explore the effects of certain associated variables and their interaction with the discussion methods and with each other. The control of extraneous variables by randomization procedures were also provided for in the design. In addition, by the use of analysis of covariance it was possible to statistically control for initial differences of the subjects on the criterion variables.

I. THE DESIGN

The students were randomly assigned to two groups. One group was designated the television (TV) group and received discussion by means of closed-circuit television. The second group was designated the non-television (NTV) group and discussion was in the face-to-face classroom situation. The randomization and discussion methods will be described in detail in Chapter Four.

The Independent Variables

In addition to the discussion variables three other independent variables with two categories in each variable were chosen for study.

The independent variables were: (1) the discussion techniques, (2) the ability level of the students, (3) the preference for instruction of the students, and (4) the sex of the students.

The discussion techniques were, of course, the primary variables under study. There were an almost unlimited number of factors that could be considered for other variables. Among these were age, interests, socio-economic status, personality classification, college grade-point average, class in college, and many others. The choice of the number and the particular factors for study in any experiment is dependent upon the purposes and design of the experiment, upon established and conventional practices, upon the interests and preferences of the experimenter, and upon the practical limitations of partitioning the variables. In the present study it was decided that more than four variables would become unwieldy in terms of statistical procedures and would be difficult to interpret if higher-order interacting effects were obtained. In the choice of the particular factors to use an attempt was made to choose variables that would be of practical use and significance for the present study and for future investigators of instructional television. Thus in addition to the discussion variables, it was decided that differences in ability level of students, the differences in the preference for instruction that students exhibit, and the differences in males and females would be the most practical variables to study in terms of educational outcomes under the conditions of television and non-television discussion.

The decision to use two categories or levels for each of the

variables was made in order to determine any interacting effects that might exist among the variables. The choice was, of course, limited to two levels in the case of the discussion and sex variables. In the case of the preference for instruction variable, and the ability level variable the number of levels was arbitrarily set at two in order to minimize the expense and labor involved in the statistical analysis.

The Dependent Variables

In the choice of criterion variables the experimenter must make decisions about what effects he wishes to measure and then select or devise instruments to measure these effects. In many cases the choice of dependent variables will be influenced or limited by the types of measuring instruments available, the conditions of the experiment, and the existing facilities. In the present study the major concern was in measuring the effects of the discussion conditions on educational outcomes. Four criterion variables were chosen. These were: (1) student achievement of course content, (2) student attitude toward teaching, (3) student misconceptions about education, and (4) student opinion about the course.

The use of four criteria was considered adequate to give a variety of measures for determining the effectiveness of the two discussion methods and their relationship to the other control variables. The criteria selected were realistic in terms of the purposes and conditions of the experiment. The experiment was conducted in a course in beginning educational psychology; it was required of all students majoring in education; and the majority of the students were sophomores.

In addition to the usual criteria of course content it was thought important to learn about the effect of the discussion techniques on the attitudes toward teaching of beginning students in educational psychology. Likewise, it was considered of interest and importance to note any relationships that might exist between the discussion methods and the ability level, preference for instruction, and sex of the students on his attitude toward teaching. The misconception criteria was chosen to include an area not measured by the usual achievement and attitude test. Misconceptions so conceived are in the nature of irrational beliefs or attitudes not based on facts and not attitudes in the usual sense. The opinion criteria was selected to determine the student's opinions about the course under the conditions of the two discussion techniques and their relationship to the other control variables. In addition the opinion variable would provide a measure of the morale of the students.

The instruments for obtaining measures of the independent and dependent variables will be fully discussed in Chapter Four.

Diagrammatic Plan of the Design

The combination of four independent variables with two levels of classification for each variable as employed in the present study is known as a four-way factorial design. It is sometimes described in quantitative terms as a 2 x 2 x 2 x 2 (read as two by two by two by two) or simply a 2^4 design. Such a four-dimensional design as contrasted to a two or three-dimensional design, is difficult if not impossible to represent or visualize pictorially. It is possible, how-

ever, to represent the relationship and inter-relationship of the four variables in a diagrammatic figure, and these relationships in the present study are shown in Figure 1 on page 73. By splitting each of the two discussion groups (TV and NTV) into two groups on the basis of scores on the ability variable; by splitting them into two groups on the basis of scores on the preference for instruction variable; and by further dividing them on the basis of sex it can be been that sixteen subgroups can be formed for purposes of analysis. Thus by using variance or covariance analysis techniques, it is possible to examine the effects of various combinations of variables at different levels of each variable and to test the significance of each. It was thus possible to simultaneously measure the main and interaction effects of the four control variables for each of the four criteria measures without resorting to the tedious procedure of several independent experiments of the single variable type.

The Control of Extraneous Variables

The design of the present experiment provided a means of control of many extraneous variables that would be difficult to control in a simpler type design. Individual differences of subjects introduces many variables into the experimental situation. In many experiments an attempt is made to control these variables by equating subjects on those factors that it is thought will bias the results such as intelligence, age, academic rank, interests, and many others. It is probably impossible to select and equate subjects on all of the many variables that might affect the experimental results. It is possible, however,

The Independent Variables															
High Ability Level								Low Ability Level							
High Preference for Instruction				Low Preference for Instruction				High Preference for Instruction				Low Preference for Instruction			
Male		Female		Male		Female		Male		Female		Male		Female	
TV	NTV	TV	NTV	TV	NTV	TV	NTV	TV	NTV	TV	NTV	TV	NTV	TV	NTV
1 *	2 *	3 *	4 *	5 *	6 *	7 *	8 *	9 *	10 *	11 *	12 *	13 *	14 *	15 *	16 *

* Scores on each of the four criterion measures are entered in the sixteen cells of this diagram for analysis.

FIGURE 1. Relationship of the four independent variables with two levels for each variable and the sixteen subgroups formed by the various combinations.

to minimize many sources of error or bias by randomization procedures, thus randomizing the uncontrolled error variables among the treatment groups. In the present study the subjects were randomly placed in either the TV or NTV groups at the outset of the experiment. In this manner any variations and sources of error due to individual differences were randomized among the two methods. The randomization procedures will be discussed in Chapter Four.

Another source of error in studies of the single variable type in which a series of experiments are performed, is that introduced by the conditions of the experimental situation changing from experiment to experiment and from the risk of not obtaining samples from the same population for subsequent experiments. Or, if the same sample is used from one experiment to another there is no way of controlling variations due to intervening activity from one experiment to another. In the present investigation such sources of variation were controlled by the design of the experiment itself, because it was possible to obtain simultaneous measurements of the effects of the desired variables.

Differences in instructors is also a possible source of error in experiments of teaching methodology. In the present experiment the instructor variable was controlled by alternating the instructors between the TV and NTV groups from one class period to another, thus any error introduced by this variable was randomized among all subjects and treatment groups.

No control section in the sense of traditional experimentation

was provided for in the present investigation. The experimental design was planned to measure the relative effectiveness of the two discussion methods, each group in effect was both a control and an experimental group in relation to the other.

Inasmuch as students bring a wide variety of abilities, interests, attitudes, aptitudes and experiences to the experimental situation it is unwise for the experimenter to assume, even for the most naive subjects, that all subjects are at the same level initially. In order to control this source of error, beyond the control afforded by randomization procedures, pre-test and post-test measures were obtained for all of the criterion variables except the opinion variable. By the use of analysis of covariance it was possible to statistically adjust for differences in initial standing on the criterion variables. In the case of the opinion variable it was thought that the administration of this measure at the beginning of the experiment would influence the post-test measures. Analysis of variance, chi-square, and the "t" test were used to test the results on this measure. A discussion of the statistical treatment will be found in section three of the present chapter.

Replication was not provided for in the present experiment. Replication could have been provided by duplicating the experiment with a second section of the same course, by repeating the experiment in a subsequest term, or by altering the present design and using the present subjects and course facilities. Because of administrative decisions in the matter of scheduling courses, instructor assignments, and classroom facilities the first two possibilities were not feasible. Altering the

design of the present experiment would have resulted in a reduction of the number of subjects for each replication or a reduction in the number of variables for study. Inasmuch as this was an initial experiment of its kind, there was no way of knowing at the outset what problems and conditions would be met during the course of the experiment. A replication would have enhanced any problems encountered and would have increased the cost of the experiment in time and money. The decision was made, therefore, to maintain the design in its present form and accept the limitations of no replication. In the future more refined investigations of a similar nature should provide for replication in the experimental design.

II. THE NULL HYPOTHESES

Modern experimental procedure usually concerns itself with the purpose of testing a "null" hypothesis or the hypothesis that there is no true differences among the experimental treatments as far as the criterion is concerned. In the present design with two levels for each of four independent variables or "treatments", four main effects and eleven interacting effects may be derived. Thus a total of fifteen different null hypotheses may be formulated for test on each of the criterion measures. These effects and interactions and their corresponding null hypotheses are:

Main Effects

1. Ability

   There are no differences between mean scores of the criterion instruments attributable to the ability level of the students.

2. Preference for instruction

There are no differences between mean scores of the criterion instruments attributable to the preference for instruction of the students.

3. Sex

There are no differences between mean scores of the criterion instruments attributable to the sex of the students.

4. Discussion techniques

There are no differences between mean scores of the criterion instruments attributable to the two discussion techniques.

First-order Interaction Effects

5. Ability and preference for instruction

There are no differences between mean scores of the criterion instruments attributable to the interaction effect of the ability level and preference for instruction.

6. Ability and sex

There are no differences between mean score of the criterion instruments attributable to the interaction effect of the ability level and sex.

7. Ability and discussion techniques

There are no differences between mean scores of the criterion instruments attributable to the interaction effect of the ability level and the two discussion techniques.

8. Preference for instruction and sex

There are no differences between mean scores of the criterion instruments attributable to the interaction effect of the preference for instruction and sex.

9. Preference for instruction and discussion techniques

There are no differences between mean scores of the criterion instruments attributable to the interaction effect of the preference for instruction and the two discussion techniques.

10. Sex and discussion techniques

There are no differences between mean scores of the criterion instruments attributable to the interaction effect of sex and the two discussion techniques.

Second-order Interaction Effects

11. Ability and preference for instruction and sex

There are no differences between mean scores of the criterion instruments attributable to the interaction effect of ability, preference for instruction and sex.

12. Ability and preference for instruction and discussion techniques

There are no differences between mean scores of the criterion instruments attributable to the interaction effect of ability, preference for instruction and the two discussion techniques.

13. Ability and sex and discussion techniques

There are no differences between mean scores of the criterion instruments attributable to the interaction effect of ability, sex, and the two discussion techniques.

14. Preference for instruction and sex and discussion techniques

There are no differences between mean scores of the criterion instruments attributable to the interaction effect of the preference for instruction, sex, and the two discussion techniques.

Third-order Interaction Effect

15. Ability and preference for instruction and sex and discussion techniques

There are no differences between mean scores of the criterion instruments attributable to the interaction effect of ability, preference for instruction, sex, and the two discussion techniques.

In the present experiment the primary concern was with those hypotheses in which the discussion technique was either a main or interacting factor although all hypotheses were tested. The statistical tools used to test the hypotheses and the assumptions underlying each are discussed in the next section.

III. THE STATISTICAL TREATMENT

The analysis of variance, the analysis of covariance, the chi-square, and the "t" test were the principal statistical tools used in the present study. By the use of these tools the experimenter is able to analyze the total variance into independent components for each

experimental variable and test its significance statistically. It is important, however, that the assumptions involved in the use of these techniques be satisfied or the experimenter cannot be sure that the obtained differences are "true" differences, but rather result from a failure to satisfy the assumptions underlying the use of the statistical tests. The assumptions underlying the analysis of variance are:

1. The groups are selected at random from the same population. This assumption was accepted in the present study on the basis of the randomization methods used and described in Chapter Four.

2. The distribution of criterion scores within groups or cells is normally distributed. This assumption of normality of distribution is seldom tested and it was not tested in the present study. It has been shown by the Norton study (38, pp. 78-90) that the F-distribution is relatively insensitive to the form of the distribution, and Walker and Lev (56, p. 230) report that "...empirical studies of samples from non-normal populations indicate that a considerable departure from normality does not invalidate the methods described" in a discussion of the use of statistical methods including the analysis of variance.

3. The groups are homogeneous in variability. This assumption was tested in the present investigation by the Bartlett test as described in Lindquist (38, pp. 87-88).

The assumptions underlying the analysis of covariance are:

1. The groups are selected at random from the same population. This assumption was accepted on the basis of the randomization methods used.

2. The distribution of criterion scores within groups or cells is normally distributed. The assumption was not tested and was discussed in connection with the analysis of variance above.

3. The groups are homogeneous in variability. This assumption was tested by the Bartlett test previously cited.

4. The regression of post-test scores on pre-test scores is the same in all groups or cells. The assumption of homogeneity of regression was tested by the method described in Walker and Lev (56, pp. 390-393).

5. The regression of post-test scores on pre-test scores is linear for all groups or cells. The assumption of linearity of regression was tested by the method described in Walker and Lev (56, pp. 396-397).

One other important condition in designs of this general type is that the number of cases from cell to cell in the same row or column be proportional. Although it is not imperative that this condition be observed, the statistical computations become very involved and burdensome if it is not observed. In the present study proportionality within cells was obtained by randomly selecting a portion of the sample from each of the television and non-television groups. In analysis involving the Chi-square and "t" tests the total sample was used.

The level of significance for accepting or rejecting all null hypotheses was set at the 5% level of confidence prior to the beginning of the experiment.

For the tests of assumptions underlying the analysis of variance

and covariance techniques, the level of confidence was set at the 1% level prior to the beginning of the experiment.

IV. SUMMARY

A four-way factorial design with four independent variables and two levels of classification for each variable was used in the present investigation. The primary variable in the study was the discussion variable with the two levels represented by a television (TV) group and a non-television (NTV) group. The other variables were the ability level, the preference for instruction, and the sex of the students. Four criterion variables were selected to measure the effects and interactions of the control variables. These were achievement of course content, attitude toward teaching, misconceptions about education, and the opinion of the course. Null hypotheses were derived for the corresponding effects and interactions of the independent variables. The basic statistical tools were the analysis of variance, the analysis of covariance, the Chi-square test, and the "t" test. The assumptions underlying the use of the statistical tools and the methods or tests for satisfying these assumptions were discussed. The level of significance for testing the null hypotheses and the assumptions were set at 5% and 1% levels respectively.

CHAPTER IV

EXPERIMENTAL METHODS AND PROCEDURES

The methods and procedures used in conducting an experiment are determined in part by the experimental design and in part by the physical setting and facilities. In the present chapter the course and the facilities, the population and the sample, the randomization procedures, the discussion techniques, and the instructors will be fully described.

I. THE COURSE AND THE FACILITIES

The investigation was conducted in section 4 of Foundations of Education 200, "The School and the Individual," a six-credit course in beginning educational psychology offered in the College of Education at Michigan State University. The section met three times weekly for two consecutive periods of fifty minutes each. The catalog description of the course follows (41);

> Mental, emotional, physical, and personality development of the human organism related to problems of formal education. Concepts of learning, motivation, memory, habits, individual similarities and differences as they condition the learning situation.

The section was chosen for the investigation because it provided a large enrollment, varying from approximately 150 to 200 students depending on the term, and because it had been previously taught by closed-circuit television and many of the administrative problems involved were known. In addition it is a required course for all students majoring in education and probably plays an important part in the for-

mation of student attitudes toward education and teaching. It was thus an ideal course for studying some of the effects under consideration in the present investigation.

The originating room for the television presentation was a large ampitheater-type lecture room with a seating capacity of 331 students. It was provided with two television cameras with three rotating lenses, necessary lighting facilities, microphones and loud-speakers, a blackboard, an easel, a speaker's lectern, and a table and chairs.

The control room in the basement was equipped with a 16mm movie projector, a 2x2 slide projector, and monitoring and switching apparatus for the program director and engineers.

Two viewing classrooms each with a capacity for 96 students were provided for the experimental subjects. Each viewing room contained two 24" television monitors, two loud-speakers, a microphone, and the usual classroom equipment. The microphone and loud-speakers provided two-way audio communication with the originating room.

II. THE POPULATION AND THE SAMPLE

The population

All of the subjects in the present investigation were students enrolled in Michigan State University during Fall Term 1957, and registered in section 4 of Foundations of Education 200. The students were believed to be representative of all students enrolling in Foundations of Education 200. Typically the course is composed of male and female students from the Basic College and the College of Agriculture, Business

and Public Service, Communication Arts, Engineering, Home Economics, Science and Arts, and Education with the College of Education having the largest representation.

The sample

Two-hundred and eighteen students originally registered for the course. As there was seating capacity for only ninety-six students in each of the viewing rooms, twenty-six students were assigned to the originating room. These twenty-six students were not included in the experiment. Of the remaining 192 students, three dropped the course during the second week of the term. Twenty-five additional students were excluded from the sample because they either failed to report during the first day of class when pre-tests were administered or because their college aptitude test scores were not available at the University Evaluation Services. The remaining 167 students were included in the sample for the present study, and the distribution of these students by college class, sex, and major curricula is shown in Table II on page 85 In the table it is shown that the largest number of students were sophomores and were majoring in education and that females predominantly outnumbered males.

III. THE RANDOMIZATION PROCEDURES

On the first day of class the students completed a Personal Data Card (Appendix A, p. 147) containing their name, age, class, major and other pertinent data. These cards were collected and before the next class meeting they were thoroughly shuffled and dealt into two piles

TABLE II

NUMBER OF STUDENTS IN THE SAMPLE CLASSIFIED BY CLASS, SEX AND MAJOR CURRICULA

Major Curricula	Fresh-man		Sopho-more		Junion		Senior		Special		Total		
	M	F	M	F	M	F	M	F	M	F	M	F	
No Preference	1	2	3	9	1	0	0	0	1	0	6	11	17
Agriculture	0	0	4	1	0	0	0	0	0	0	4	1	5
Business	0	0	0	0	0	0	1	0	0	0	1	0	1
Business Education	0	0	0	2	1	9	0	1	0	0	1	12	13
Political Science	0	0	0	0	1	0	0	0	0	1	1	1	2
Social Work	0	0	0	1	0	0	0	0	0	0	0	1	1
Engineering	0	0	0	0	0	0	1	0	0	0	1	0	1
Home Economics	0	0	0	7	0	4	0	0	0	0	0	11	11
Fine Arts	0	0	2	1	1	1	0	2	0	0	3	4	7
Language & Literature	0	1	0	1	0	4	0	0	0	0	0	6	6
Biological Science	0	0	1	0	1	0	2	0	0	0	4	0	4
Physical Science	0	0	1	0	1	2	0	0	0	0	2	2	4
Social Science	0	4	4	7	6	2	1	1	0	0	11	14	25
Pre-Medical	0	0	0	0	1	0	0	0	0	0	1	0	1
Speech	0	0	0	1	1	0	0	0	0	0	1	1	2
Education	1	3	1	42	2	9	0	1	0	0	4	55	59
Physical Education	0	1	6	1	0	0	0	0	0	0	6	2	8
TOTAL	2	11	22	73	16	31	5	5	1	1	46	121	167
	13		95		47		10		2		167		

corresponding to the television and non-television groups. No bias was introduced by this method as each card had an equal chance of being placed in one of the two groups. The television and non-television groups were designated by the toss of a coin. In order not to exceed the seating capacity of each viewing room the last twenty-six students on the class list were assigned to the originating room and were included in the sample. Although this procedure would seemingly introduce error into the randomness of the sample, inasmuch as these twenty-six people were from the latter part of the alphabet, an unpublished study by Krumboltz (29) based on a sample of 12,000 Air Force recruits indicates that no bias is introduced by alphabetical sampling.

At the end of the second class meeting, during which pre-tests and an introductory lecture were given, the students were assigned to their respective groups and rooms and for the remainder of the class meetings reported to their assigned rooms.

III. THE DISCUSSION TECHNIQUES

The discussion method as used in the present study may be defined as student participation by the interchange of opinion, information, illustration, question and answer, and pro and con arguments for the purpose of understanding and clarification, with the instructor playing the role of group leader and moderator. The instructor usually defined the area for discussion, posed questions for consideration, supplied information and illustration where necessary or when asked to do so, clarified and related meanings where necessary, and helped maintain an

orderly sequence of discussion aimed around a central theme or problem.

The two types of discussion used were six-by-six and panel discussions. In the six-by-six discussion the class was organized into groups of not more than six students per group. The groups discussed the topic, problem, or question under consideration for about six to ten minutes after which the recorder from each group reported the views of the group on the topic, and these were further discussed by the class with the instructor acting as the moderator and group leader. During the six-by-six discussion phase the instructor moved from group to group to help in clarifying problems and to act as a resource person.

In the panel discussions the students were grouped into panels of not more than six persons. The panels were assigned topics for presentation to the class, and gave a fifteen to twenty minute presentation. After the panel presentation the class further discussed the topic or problem posed by the panel, or engaged in question and answer and discussion with the panel. The instructor acted as a resource person for the panel and helped the panel to organize their topics. During the post-panel discussion the instructor allowed the panel members to conduct the discussion except to help clarify and to maintain the discussion on the topic.

In the use of each of these types of discussion the last few minutes of the class period were utilized by the instructor for a summary of the day's discussion. The summary consisted of a brief statement of the topic, an enumeration of the main points discussed, a final clarification of difficult and misunderstood points, and an emphasizing

of important points, concepts or principles. Related problems were also mentioned. The use of these discussion methods will be discussed under the description of the television and non-television techniques.

The class meeting was divided into two fifty-minute periods. The first fifty-minutes of the class consisted of a lecture over the closed-circuit television system to both the television and non-television groups. The lectures were supplemented by demonstrations, kinescopes, movies, and guest speakers. A ten minute intermission followed the first period. During the second fifty-minute period the class participated in discussion.

Television Discussion

During the discussion period the instructor remained in the originating room while an assistant was present in the viewing room. The discussion was conducted between the classroom and the originating room with feedback provided by the two-way audio and one-way visual communication system. In the six-by-six discussion at the completion of the six to ten minute group discussions the group recorders reported to the originating room where they presented reports to the class over television. The instructor remained in the originating room and acted as a moderator for the ensuing discussion, while the assistant in the classroom proctored the class.

In the panel discussion the panels presented their discussion to the class over the television system after which the discussion was conducted between the class and the panel members over the television with the instructor again acting as moderator.

Non-television Discussion

All discussions in the non-television group were conducted in the conventional face-to-face manner in the classroom with the instructor in the classroom. The topics for six-by-six and panel discussion were coordinated with the television section so that for any day the two groups were discussing the same topics. The primary difference between these two groups was in the feedback arrangement. In the non-television group feedback was obtained in the face-to-face situation while in the television group feedback was obtained by the television and loud-speaker arrangement. The first and second period arrangements are shown in the schematic diagram on page 90.

IV. THE INSTRUCTORS

The instructors in the course were two full-time Assistant Professors in the College of Education. The lecture presentations were shared equally by the two instructors. Two assistants aided in the administrative duties such as roll-taking, adjustment of television monitors, and proctoring in the two viewing classrooms. One was a part-time Instructor and one was a Graduate Assistant in the College of Education. The Instructors and Assistants were alternated between the two classrooms to minimize any differential bias due to instructor differences.

V. SUMMARY

The investigation was conducted in Foundations of Education 200,

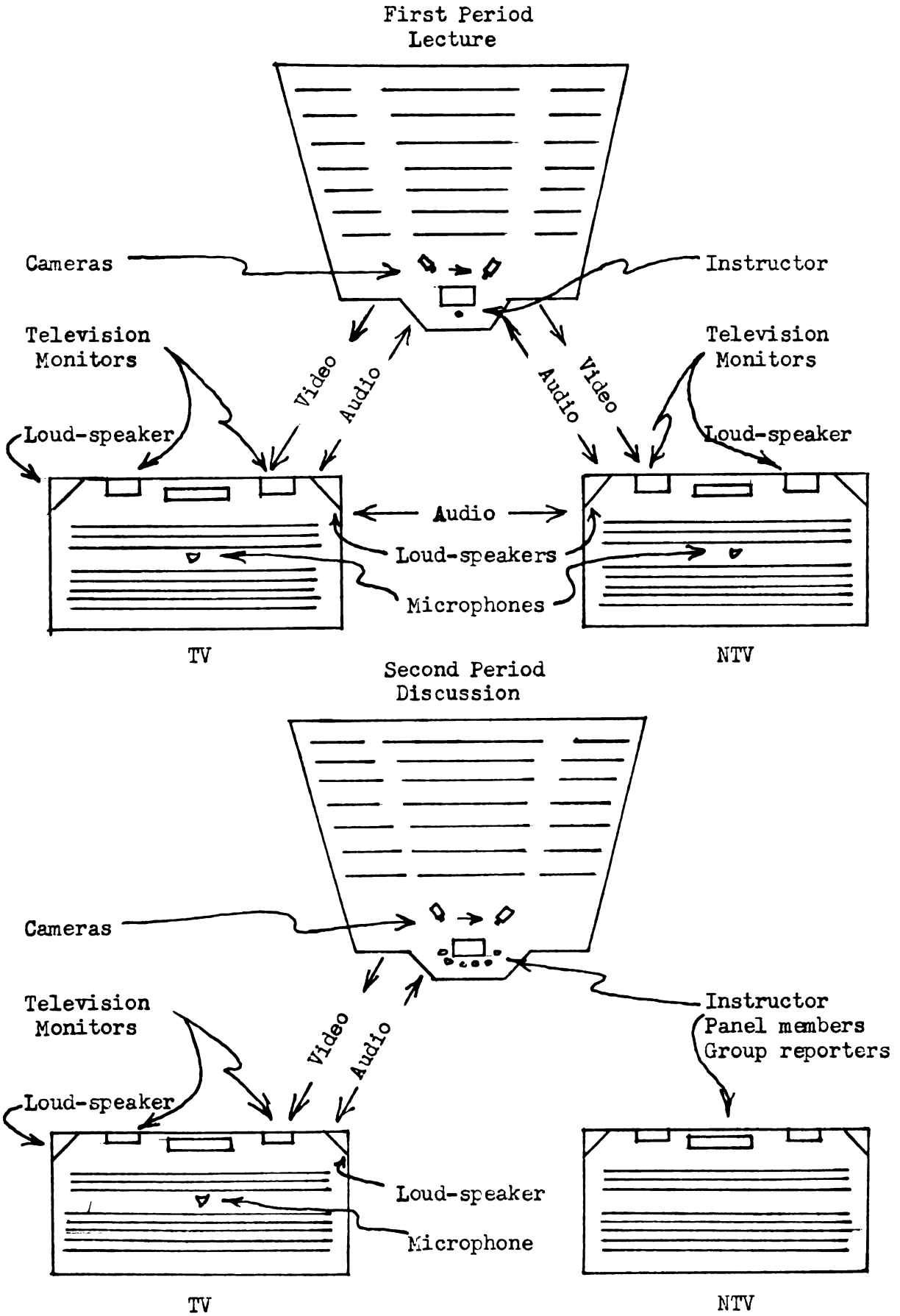

Figure 2. Schematic diagram of the first and second period arrangements of the TV and NTV groups.

Section 4, a course in beginning educational psychology, required of all students majoring in education. The physical facilities included a television originating room, a control room and two viewing classrooms equipped with the necessary cameras, monitors, lighting and technical equipment necessary for television production and viewing.

The sample consisted of 167 students and was believed to be representative of the students enrolling in Foundations of Education 200. The students were randomly placed in the television (TV) or non-television (NTV) groups.

The first fifty minutes of the class period consisted of a television lecture to both groups. The second fifty-minute period of each class consisted of six-by-six and panel discussions followed by general class discussion with the instructor acting as a moderator. In the television group the discussions were conducted over the television system between the class and the panel members or group reporters. In the non-television group all discussions were held in the classroom in a face-to-face situation.

The instructors for the course were two full-time assistant professors aided by two assistants. All instructors were alternated between the two groups to minimize differences.

CHAPTER V

EVALUATION INSTRUMENTS

The experimenter can frequently use existing instruments for the purposes of the investigation, but often it is necessary to construct or adapt available instruments appropriate to the objectives being sought in the study. In the present investigation some of the measuring instruments were in existence and some of the instruments were constructed by the experimenter. The present chapter will contain a description of the evaluation instruments used in the study, a discussion of the administration of the instruments and collection of the data, and an explanation of the procedures used in recording the data.

I. MEASURES OF THE INDEPENDENT VARIABLES

It will be recalled that the independent variables were (1) the discussion techniques, (2) the ability level of the students, (3) the students' preference for instruction, and (4) the students' sex. Since the discussion techniques were fully described in Chapter Four and the students' sex was obtained from the Personal Data Card only the measures of ability and preference for instruction will be discussed here. The American Council on Education Psychological Examination for College Freshmen was used as the measure of the students' ability and the Preferred Instructor Characteristic Scale was used to measure the students' preference for instruction. A description of the above instruments follows.

American Council on Education Psychological Examination (ACE)

The American Council on Education Psychological Examination for College Freshmen is designed for and standardized on entering college freshmen and is used by about 300 colleges and universities. As this test is so well known and is so widely used no description of its contents will be given. The test yields three scores: a quantitative (Q) score, a linguistic (L) score and a total (T) score. Only the T-score was used in the present study. A study of the reliability of the test by the authors, L. L. and T. G. Thurstone, as reported in Super (54), indicated an odd-even reliability of .95 for the total score. The ACE is administered to all entering students at Michigan State University by the University Office of Evaluation Services, and the scores for the students in the present study were obtained from the files of that office.

Preferred Instructor Characteristic Scale (PICS)

The Preferred Instructor Characteristic Scale is a measure designed by Farquhar (20) and Krumboltz (30). The authors postulated a "cognitive-affective" continuum of instructor characteristics. In the words of Krumboltz (30) "a cognitive instructor was defined as one concerned with the intellectual, abstract, subject-matter goals of teaching; the affective instructor was defined as being more concerned with emotional adjustments and student interaction in the classroom." A number of statements believed characteristic of the two types of instructors were proposed. After validation and refinement the final form of the scale consisted of six cognitive and six affective statements each

paired with the other to make a total of 36 items in the scale. The subject is required to choose one of each pair of statements in responding to the test items. The PICS is scored so that a high score indicates the student prefers a cognitive type of instruction and a low score indicates a preference for an affective type of instruction. The authors report a test-retest reliability coefficient of .88 and a reliability computed by Hoyt's Analysis of Variance technique of .90.

The PICS was administered to the students in the present study on the first class meeting of the term. The scale and the scoring key appears in Appendix A.

II. MEASURES OF THE DEPENDENT VARIABLES

The dependent variables in the present investigation were (1) student achievement of course content, (2) student attitude toward teaching, (3) student misconceptions about education, and (4) student opinion about the course. In the same order as listed above these variables were measured by a Pre and Post Achievement Test (PAT), the Minnesota Teacher Attitude Inventory (MTAI), the Misconceptions about Education Scale (MES), and the Course Evaluation Scale (CES). The description of each of these instruments follows.

Pre and Post Achievement Test (PAT)

The Pre and Post Achievement Test was constructed from test items in the Instructor's Handbook accompanying the textbook for the course, "Educational Psychology" by Cronbach (14). Fifty multiple-choice items were selected as representative of the content of the course. Cronbach

has ranked the items according to difficulty into three classifications and the items were selected so that approximately one-half were in the easiest and most difficult range and about one-half were of medium difficulty.

The corrected split-half reliability coefficient based on the odd-even items was found to be .53. A second reliability measure computed by a Kuder-Richardson formula using the difficulty level for each item and the standard deviation of the test was .52, agreeing almost perfectly with the first value. Such a reliability is considerably lower than is considered useful for individual testing purposes and would reduce the standard error of measurement by only approximately 70% of the value if the reliability were zero. However, for purposes of group measurement as were employed in the present experiment a reliability of about .50 is considered adequate. Kelley states, "If group measurement is all that is undertaken, scores which are individually reliable are not demanded, for the reliability of an average score is such greater than that of the single score....a test so unreliable that it will not be serviceable in making individual diagnosis may be very serviceable for group diagnosis." (28). Assuming that the errors in measurement are distributed in both directions from a true value, in a large number of cases they would tend to cancel each other, the group mean would not be adversely affected by a low reliability, and the test would measure adequately enough for the purposes of the experiment.

The Pre-Achievement Test was administered to the students during the first class-meeting of the term, and the Post-test during the final

examination period. The test and the scoring key will be found in Appendix A.

Minnesota Teacher Attitude Inventory (MTAI)

The Minnesota Teacher Attitude Inventory is published by the Psychological Corporation and was developed by Cook, Leeds, and Callis. The Inventory is based on the assumption that the attitudes of the teacher are the key to the type of classroom he will be able to maintain and that these attitudes are the result of the interaction of a multitude of factors including social intelligence, personality traits, values, academic intelligence, general knowledge, and teaching techniques. The Inventory is designed to measure those attitudes of a teacher which predict how well he will get along with pupils in interpersonal relationships and indirectly how well satisfied he will be with teaching as a vocation. The Inventory consists of 150 attitude statements about which the subject is asked to respond on a five-choice scale of "strongly agree", "agree", "undecided or uncertain", "disagree", and "strongly disagree". Scoring is on the basis of "right" minus "wrong" answers. The instrument was validated on teachers in the field and norms appear in the test booklet. The split-half reliability of the inventory in three separate studies was found to be .93, .88, and .93. The MTAI was administered to the students as a pre-test during the second class meeting and as a post-test during the final examination period of the term.

Misconceptions about Education Scale (MES)

The Misconceptions about Education Scale was designed by the

writer and a colleague[1] in an attempt to measure an area that is not ordinarily measured by the usual attitude or achievement test. Misconceptions about Education refer to irrational beliefs or attitudes not based on factual knowledge that are believed to be held by the layman, beginning students, and neophyte teachers. Such beliefs or attitudes are akin to superstitious beliefs or "old wives' tales" but not superstitious in the usual meaning of the word. A number of statements were constructed by examining psychology and educational psychology textbooks for discussions of such fallacies. The Wickman study (58) was also a source for the construction of many of the statements. The statements were submitted to seven full-time faculty members of the College of Education and all items about which there was disagreement were discarded. Ambiguous statements were reworded. The final scale consisted of 61 statements such as "A high forehead usually indicates intellectual superiority." The subjects were required to answer each statement as a true-false item. A corrected split-half reliability was computed and found to be .63. The scale was administered as a pre-test to the students during the second class-meeting and as a post-test during the final examination period. A copy of the scale and the scoring key will be found in Appendix A.

---

[1]Mr. Clarence Williams, formerly Part-time Instructor in the Department of Foundations of Education, Michigan State University.

Course Evaluation Scale (CES)

The Course Evaluation Scale was an adaptation of a similar instrument used in instructional television research at Pennsylvania State University. The scale consisted of 22 multiple-choice items about the course, the instruction, the evaluation procedures, about the ways in which television was used, and about the mechanics of the course. The multiple-choice items were designed after the Likert technique (36). Each response was given a weight on a five-point scale so that a five indicated the most favorable response, a three was a neutral response, and a one indicated a least favorable response. In addition to the 22 items, space was provided for comments, criticisms, and suggestions. The scale was so scored that a high score indicated a relatively favorable opinion of the course and a low score indicated a relatively unfavorable opinion of the course. No reliability was computed for this instrument. The scale and the scoring key appear in Appendix A. The scale was administered to the students on the last class-meeting of the course.

III. THE COLLECTION AND RECORDING OF THE DATA

The pre-tests were administered during the first two class periods with all subjects present in one room. No time limit was set for any of the instruments and all students completed the tests before the end of the two-hour period. Instructions were given by one of the instructors and the assistants aided in distributing and collecting

materials and in proctoring the tests.

The post-tests were administered in the respective rooms of the TV and NTV groups during the two final examination periods of the course. The procedures for administering the tests were established beforehand and both groups began the tests at the same time. Again no time limit was set for completion of the tests and all subjects completed the tests before the end of the two-hour period.

All answer sheets were machine-scored by the University Office of Evaluation Services. The scores were recorded on the Personal Data Card (see Appendix A) and were later key-punched into IBM cards for machine computation.

IV. INTERCORRELATION OF INSTRUMENTS

The intercorrelations of the measuring instruments for the TV and NTV groups are presented in Table III on page 100. The correlations are based on the scores of 86 students in the TV group and 81 students in the NTV group, the 167 students for whom a complete set of scores were obtainable.

In comparing the correlation between pre and post scores on the MTAI, the MES and the PAT it will be noted that the MTAI is the highest, the MES is next and the PAT has the lowest correlation of the three. The reliability coefficients for these three instruments (reported in the previous section) ranked in the same order indicating a consistency in this respect. The significant correlations on the post-test measures of these three instruments seems to indicate that a common factor is

TABLE III

INTERCORRELATIONS OF THE EVALUATION INSTRUMENTS

	ACE		PICS		PRE PAT		POST PAT		PRE MTAI		POST MTAI		PRE MES		POST MES		CES	
	TV	NTV	TV	NTV	TV	NTV	TV	NTV	TV	NTV	TV	NTV	TV	NTV	TV	NTV	TV	NTV
ACE			.18*	.01	.17	.37**	.13	.15	.08	.10	.13	.13	.23*	.19*	.22*	.10	-.11	-.04
PICS	.18*	.01			-.12	-.36**	.13	.02	.06	.01	-.03	.22*	.11	.01	.07	.39**	.07	.23*
PRE PAT	.17	.37**	-.12	-.36**			.32**	.28**	.12	.01	.22*	.15	.25*	.32**	.32**	.03	-.12	.01
POST PAT	.13	.15	.13	.02	.32**	.28**			.01	.31**	.15	.35**	.27**	.19*	.38**	.38**	.16	-.03
PRE MTAI	.08	.10	-.06	.01	.12	.01	.01	.31**			.64**	.69**	.36**	.43**	.39**	.44**	.05	.03
POST MTAI	.13	.13	-.03	.22*	.22*	.15	.15	.35**	.64**	.69**			.21*	.41**	.37**	.46**	.16	.06
PRE MES	.23*	.19*	.11	.01	.25**	.32**	.27**	.19*	.36**	.43**	.21*	.41**			.59**	.50**	.03	-.10
POST MES	.22*	.10	.07	.39**	.32**	.03	.38**	.38**	.39**	.44**	.37**	.46**	.59**	.50**			.18*	.03
CES	-.11	-.04	.07	.23*	-.12	.01	.16	-.03	.05	.03	.16	.06	.03	-.10	.18*	.03		

* Significantly different from zero at the 5% level. TV - N = 86 5% = .177 1% = .248
** Significantly different from zero at the 1% level. NTV - N = 82 5% = .181 1% = .253

ACE - American Council on Education Psychological Examination for College Freshmen
PICS - Preferred Instructor Characteristics Scale
PRE PAT - Pre-Achievement Test
POST PAT - Post Achievement Test
PRE MTAI - Pre Minnesota Teacher Attitude Inventory
POST MTAI - Post Minnesota Teacher Attitude Inventory
PRE MES - Pre Misconceptions about Education Scale
POST MES - Post Misconceptions about Education Scale
CES - Course Evaluation Scale

being measured by these instruments.

It is interesting to note that the correlation between the PICS and the CES was significant in the NTV group but not in the TV group, indicating that the students with a preference for a cognitive type of instruction held a more favorable opinion toward the course in the NTV group than in the TV group.

The ACE, the PICS and the CES, except in a few instances, were not significantly correlated with the other measures nor with each other, and thus were measuring independently of each other.

V. SUMMARY

The T-score on the ACE was used as a measure of the ability of the students in the present investigation. It is a widely used test and has a reliability that is very high.

The PICS was used as a measure of the students' preference for a cognitive or an affective type of instruction. The test-retest reliability of this instrument is high.

The PAT was designed as a measure of the achievement of the course content. It is a 50 item multiple-choice type examination selected from the Instructor's Handbook accompanying the course textbook. The split-half reliability of this instrument was low, but was considered adequate for the measurement of group means used in the investigation.

The Minnesota Teacher Attitude Inventory, a commercially published instrument, was employed to measure the students' attitudes toward teaching. The inventory contains 150 attitude statements about

which the subject responds on a five-point scale of agreement or disagreement. The reliability of this instrument has been found to be consistently high.

The MES was designed to measure the students' misconceptions about education. Misconceptions in the context used here relates to irrational beliefs or attitudes believed to be held by the laymen, beginning students and the neophyte teacher. The scale consisted of 61 true-false items and it was found to be of medium reliability.

The CES was designed to measure the students' opinions about the course. It consisted of 22 Likert-type items covering the instruction, evaluation, use of television, and the mechanics of the course. No reliability was computed for this instrument.

The pre-tests were administered during the first two class meetings of the course, and the post-tests were administered during the final examination periods. ACE scores were obtained from the University Office of Evaluation Services. The scores for all measures were machine-scored and entered on the Personal Data Card.

The intercorrelations of the measuring instruments were computed and the results presented in tabular form. The correlations between the post-test scores for the MTAI, the MES and the PAT were found to be consistent with the reliabilities reported for these instruments. The ACE, the PICS and the CES were found to be measuring independently of the other instruments, while the MTAI, the MES and the PAT were found to have low but significant relationships.

CHAPTER VI

ANALYSIS OF THE EXPERIMENTAL RESULTS

A description of the computational procedures, the results of the tests of assumptions underlying the statistical techniques, and the results of the statistical tests for each of the dependent variables will be discussed in the present Chapter.

I. THE COMPUTATIONAL PROCEDURES

The basic data for each individual including his name, student number, sex, curriculum, class, the experimental group (TV or NTV), and the scores on all measuring instruments was key-punched into an IBM card. The cards for each experimental group were then sorted on ACE scores and the cutting-score for high and low groups determined. The cards were next combined and the cutting-scores for the high and low PICS were obtained in the same manner. The cards for each experimental group were then sorted into the sixteen subgroups corresponding to the sixteen cells in the experimental design (See Chapter III, Figure 1). In order to maintain proportionality between subgroups for the variance and covariance analysis, using a table of random numbers, cards were randomly selected from the subgroups so that each male subgroup contained 3 subjects and each female subgroup contained 11 subjects for a total of 56 subjects in each of the TV and NTV groups. This reduced the TV group by 30 students and the NTV group by 25 stu-

dents. The subgroup numbers were then punched in the cards for identification. The basic data with the exception of the name and student number is tabulated in Appendix B, "Original Data."

The calculation and summation of scores, squares, and cross-products was accomplished by passing the cards through a 604 IBM Electronic Calculating Punch. As a check on the programming of the machine the scores for one subgroup were calculated by hand on a mechanical computer, and as a further check all cards were fed through the machine for two passes. The data from the above calculations were automatically punched on a summary card for each subgroup.

Appendix C, "Explanation of Computational Procedures," contains a series of tables showing the manner in which the data for the sixteen subgroups were combined to test each of the null hypotheses. A separate table is shown for each of the main and interaction effects. Each table shows the identification number of the subgroup and the number of subjects in each subgroup and the total number of subjects. By referring to these tables in Appendix C and the original data in Appendix B, the reader can compute the sums of squares for the pre and post measures, the cross-products, the adjusted variances, and the F ratios. For the test of the null hypotheses the sums of scores, sums of squares, and sums of cross-products for each subgroup were combined and entered in its proper place in the table corresponding to the null hypothesis under test. The adjusted scores, variances and F ratios were then computed and entered in the analysis of covariance tables for each of the criterion tests. All of the above calculations were

accomplished on a mechanical computer.

The above procedures apply to the covariance analysis of the measures for which pre and post tests were obtained. Essentially the same procedures were applied to the variance analysis of the CES measure except, of course, the computation of cross-products and adjusted scores was eliminated as this was a post-test measure only.

For Bartlett's Test of Homogeneity of Variance it was necessary to compute an adjusted post-test score for each of the sixteen subgroups for the three criterion measures involving this assumption. The adjusted scores were computed on a hand calculator and are tabulated in Appendix D, "Tests of Assumptions of Variance and Covariance." The data for all other tests of assumptions also will be found in Appendix D. The assumptions underlying the analysis of variance and covariance were discussed in Chapter III, Section 3.

II. ANALYSIS OF THE ACHIEVEMENT TEST (PAT)

The results of the tests of assumption of covariance for the Achievement Test (PAT) are summarized below in Table IV.

TABLE IV

RESULTS OF TESTS OF ASSUMPTIONS OF COVARIANCE FOR THE ACHIEVEMENT TEST (PAT)

Assumption	Obtained Value	Value at 1% level	Null Hypothesis
Homogeneity of Regression	0.56	2.28	Accept
Linearity of Regression	1.03	2.32	Accept
Homogeneity of Variance	35.66	30.58	Reject

The assumption that the variance from subgroup to subgroup is the same, except for chance, must be rejected. Lindquist (38, pp. 96-7) has shown that heterogeneity of variance will result in a considerably larger portion of high F value for tests of significance than the proportions given in the table for F. In effect this would result in the lowering of the level of confidence. Quoting from Lindquist,

> In general, when the heterogeneity in form or variance is 'marked' but not 'extreme', allowance may be made for this fact by setting a higher 'apparent' level of significance for the tests of treatment effects than would otherwise be employed. In cases of very marked heterogeneity, for example, if one wishes the risk of a Type I error not to exceed 5%, he might require the effect to be 'significant' at the 2.5% level, or if he wants the risk of a Type I error not to exceed 1%, he might set the 'apparent' level of significance of the test at 0.1%. (38, p. 86)

Unfortunately it is impossible to state what the limits are for a "marked" or "very marked" heterogeneity, or to what level one must change the level of significance to compensate for the heterogeneity. In the analysis of the Achievement Test that follows it is impossible to ascertain the size of the F ratio required for rejection of the null hypotheses at the 5% level, and therefore none of the null hypotheses can be rejected. However, if the covariance analysis reveals no significant F ratios, the null hypotheses of no differences can be accepted.

The analysis of covariance of the PAT will be found in Table V on page 107. Inspection of Table V reveals that there were no significant differences on the Achievement Test due to the ability level of the students, the students' preference for instruction, the students' sex, or the two discussion techniques, nor were there significant differences due to the interaction effects of these variables.

TABLE V

ANALYSIS OF COVARIANCE
PAT POST-TEST SCORES ADJUSTED FOR PRE-TEST SCORES

Nature of Variation	df	Sum of Squares	Variance	F*	Null Hypothesis
ACE (A) Ability	1	3.59	3.59	0.19	Accept
PICS (P) Preference for Instruction	1	61.32	61.32	3.29	Accept
SEX (S)	1	66.83	66.83	3.58	Accept
TECH (T) Discussion Techniques	1	0.83	0.83	0.04	Accept
A x P	1	9.91	9.91	0.53	Accept
A x S	1	0.00	0.00	0.00	Accept
A x T	1	11.59	11.59	0.62	Accept
P x S	1	10.08	10.08	0.54	Accept
P x T	1	11.25	11.25	0.60	Accept
S x T	1	20.73	20.73	1.11	Accept
A x P x S	1	44.44	44.44	2.38	Accept
A x P x T	1	18.57	18.57	1.00	Accept
A x S x T	1	4.02	4.02	0.22	Accept
P x S x T	1	18.39	18.39	0.99	Accept
A x P x S x T	1	14.21	14.21	0.76	Accept
Between Cells	(15)	(306.51)	(20.43)		
Within Cells	95	1772.46	18.66		
Total	110	2078.97	18.90		

*$F_{5\%}$ (1,95) 3.94

III. ANALYSIS OF THE MINNESOTA TEACHER ATTITUDE INVENTORY (MTAI)

The results of the tests of assumptions of covariance for the MTAI are tabulated in Appendix D and are summarized in Table VI below.

TABLE VI

RESULTS OF TESTS OF ASSUMPTIONS OF COVARIANCE FOR THE MINNESOTA TEACHER ATTITUDE INVENTORY (MTAI)

Assumption	Obtained Value	Value at 1% level	Null Hypothesis
Homogeneity of Regression	0.81	2.28	Accept
Linearity of Regression	1.57	2.32	Accept
Homogeneity of Variance	10.57	30.58	Accept

As indicated in Table VI the assumptions underlying the use of the covariance analysis were met for the data obtained on the MTAI, and the data may be validly tested by the covariance technique.

The results of the analysis of covariance of the MTAI data are presented in Table VII, page 109. Inspection of the table shows that a significant interaction effect exists between the Preference for Instruction (PICS) and the Discussion Techniques (TECH) variables. For all other effects the null hypotheses of no differences were accepted.

The pre-test mean scores, the post-test means scores and the adjusted post-test mean scores for the PICS x TECH interaction are presented in Table VIII, page 110. Examination of the adjusted post-test scores shows that the "cognitive"-NTV group obtained the highest scores,

TABLE VII

ANALYSIS OF COVARIANCE
MTAI POST-TEST SCORES ADJUSTED FOR PRE-TEST SCORES

Nature of Variation	df	Sum of Squares	Variance	F*	Null Hypothesis
ACE (A) Ability	1	648.53	648.53	1.72	Accept
PICS (P) Preference for Instruction	1	1298.38	1298.38	3.44	Accept
SEX (S)	1	1301.04	1301.04	3.44	Accept
TECH (T) Discussion Techniques	1	118.62	118.62	0.31	Accept
A x P	1	0.84	0.84	0.00	Accept
A x S	1	23.35	23.35	0.06	Accept
A x T	1	15.73	15.73	0.04	Accept
P x S	1	264.84	264.84	0.70	Accept
P x T	1	2281.04	2281.04	6.04	Reject
S x T	1	30.81	30.81	0.08	Accept
A x P x S	1	345.19	345.19	0.91	Accept
A x P x T	1	7.25	7.25	0.02	Accept
A x S x T	1	91.94	91.94	0.24	Accept
P x S x T	1	767.12	767.12	2.03	Accept
A x P x S x T	1	1020.84	1020.84	2.70	Accept
Between Cells	(15)	(8347.16	(552.57)		
Within Cells	95	35898.05	377.87		
Total	110	44245.21			

*$F_{5\%}$ (1,95) 3.94

the "affective"-TV group scored next highest, and the "cognitive"-TV and "affective"-NTV groups scored the lowest. Inspection of the differences in adjusted mean scores reveals that the "cognitive" students obtained much higher scores in the NTV group than the "cognitive" students in the TV group, whereas the differences for the "affective" students in the TV and NTV groups is much smaller. It will also be noted that the differences between the "affective" and "cognitive" students in the NTV group is large, while the differences between the "affective" and "cognitive" students in the TV groups is considerably smaller.

TABLE VIII

MEAN SCORES AND ADJUSTED MEAN SCORES ON THE MTAI FOR THE PICS x TECH INTERACTION

	High PICS "Cognitive"			Low PICS "Affective"		
	Pre	Post	Adjusted Post	Pre	Post	Adjusted Post
Technique TV	27.1	50.3	45.8	11.4	43.8	48.3
Technique NTV	22.0	58.7	58.0	24.5	44.1	43.4

Apparently under the conditions of the present experiment the combination of "cognitive" type students and face-to-face discussion promoted the best scores in attitudes toward teaching as measured by the MTAI. The television discussion technique was not nearly as effective for either the "cognitive" or the "affective" students although it favored the "affective" students slightly.

The interpretation of these findings is somewhat hazardous. The

rejection of a null hypothesis of no differences indicates only that it is unlikely the observed differences are due to chance effects and does not indicate the cause of the differences. The differences may be due to uncontrolled variables in the design of the experiment itself. In interpreting the above findings it is not possible to say with certainty that the differences were due to the discussion techniques, because the instructor variable could have biased the results. Although the instructors were rotated between sections to equalize any differences due to instructors, it is possible that all instructors favored the "cognitive" type of instruction rather than the "affective" type, and if this were the case, it would favor the students who designated a preference for a "cognitive" type of instruction. Even this bias would not be serious if it were constant for both groups, as both "affective" and "cognitive" students would be affected equally. It is possible, however, that the instructors in the present investigation displayed different degrees of "cognitiveness" or "affectiveness" in the face-to-face and television situations which would introduce a third uncontrolled interacting factor.

Therefore, in interpreting the above findings, it can only be said that under the conditions of the present experiment, the evidence indicates that the students' attitudes toward teaching as measured by the Minnesota Teacher Attitude Inventory are most affected by the combination of a preference for a "cognitive" type of instruction and the face-to-face discussion technique.

IV. ANALYSIS OF THE MISCONCEPTIONS ABOUT EDUCATION SCALE (MES)

The data for the MES was tested for the assumptions and the results are shown in Appendix D and are summarized in Table IX below.

As indicated in Table IX all tests of assumptions were satisfied for this data, and the analysis of covariance is a valid test for the data.

TABLE IX

RESULTS OF TESTS OF ASSUMPTIONS OF COVARIANCE FOR THE MISCONCEPTIONS ABOUT EDUCATION SCALE (MES)

Assumption	Obtained Value	Value at 1% level	Null Hypothesis
Homogeneity of Regression	0.85	2.28	Accept
Linearity of Regression	0.80	2.32	Accept
Homogeneity of Variance	17.41	30.58	Accept

The analysis of covariance for the MES will be found in Table X on page 113. The analysis shows that all null hypotheses can be accepted and that there are no differences, other than chance differences, on the Misconceptions about Education Scale due to the main or interacting effects of the independent variables.

V. ANALYSIS OF THE COURSE EVALUATION SCALE (CES)

Inasmuch as the CES was employed as a post-test only, the analysis of variance was used to test the data and the only assumption tested

TABLE X

ANALYSIS OF COVARIANCE
MES POST-TEST SCORES ADJUSTED FOR PRE-TEST SCORES

Nature of Variation	df	Sum of Squares	Variance	F*	Null Hypothesis
ACE (A) Ability	1	2.77	2.77	0.18	Accept
PICS (P) Preference for Instruction	1	15.54	15.54	0.98	Accept
SEX (S)	1	0.35	0.35	0.02	Accept
TECH (T) Discussion Techniques	1	1.76	1.76	0.11	Accept
A x P	1	4.76	4.76	0.30	Accept
A x S	1	15.39	15.39	0.97	Accept
A x T	1	3.50	3.50	0.22	Accept
P x S	1	39.03	39.03	2.47	Accept
P x T	1	27.43	27.43	1.73	Accept
S x T	1	0.58	0.58	0.04	Accept
A x P x S	1	12.51	12.51	0.79	Accept
A x P x T	1	37.36	37.36	2.36	Accept
A x S x T	1	26.69	26.69	1.69	Accept
P x S x T	1	0.80	0.80	0.05	Accept
A x P x S x T	1	1.64	1.64	0.10	Accept
Between Cells	(15)	(187.52)	(12.50)		
Within Cells	95	1501.79	15.81		
Total	110	1689.31			

*$F_{5\%}$ (1,95) 3.94

was that of homogeneity of variance. The results are tabulated in Appendix D and summarized in Table XI below.

The assumption of homogeneity of variance was met and the analysis of variance is a valid test for this data.

TABLE XI

RESULT OF THE TEST OF HOMOGENEITY OF VARIANCE FOR THE COURSE EVALUATION SCALE (CES)

Assumption	Obtained Value	Value at 1% level	Null Hypothesis
Homogeneity of Variance	20.07	30.58	Accept

The results of the analysis of variance for the CES are presented in Table XII, page 115. For this analysis the results on items 7, 10, 11, 12, 15, and 16 were omitted in computing the total score as these items refer to particular course procedures and cannot be evaluated in terms of the opinion for cross-comparison purposes. It will be noted that the F value for the effect of the discussion techniques was sufficently large for rejection of the null hypothesis at the 5% level of confidence, and that there is a significant difference between the TV and NTV groups on CES scores. Since the effect involves only the difference between two means, it was tested by the "t" test for verification and the results are shown in Table XIII which reveals that the difference between the obtained means for the two discussion groups is significant at the 5% level.

TABLE XII

ANALYSIS OF VARIANCE
CES SCORES

Nature of Variation	df	Sum of Squares	Variance	F*	Null Hypothesis
ACE (A) Ability	1	46.29	46.29	0.24	Accept
PICS (P) Preference for Instruction	1	6.04	6.04	0.03	Accept
SEX (S)	1	37.85	37.85	0.19	Accept
TECH (T) Discussion Techniques	1	880.32	880.32	4.50	Reject
A x P	1	66.04	66.04	0.34	Accept
A x S	1	14.31	14.31	0.07	Accept
A x T	1	38.39	38.39	0.20	Accept
P x S	1	0.02	0.02	0.00	Accept
P x T	1	5.14	5.14	0.03	Accept
S x T	1	5.00	5.00	0.03	Accept
A x P x S	1	6.91	6.91	0.04	Accept
A x P x T	1	137.28	137.28	0.70	Accept
A x S x T	1	151.44	151.44	0.26	Accept
P x S x T	1	132.83	132.83	0.68	Accept
A x P x S x T	1	18.71	18.71	0.10	Accept
Between Cells	(15)	(1447.07	(96.44)		
Within Cells	95	18788.36	195.71		
Total	110	20235.43			

*$F_{5\%}$ (1,95) 3.94

TABLE XIII

RESULTS OF THE "t" TEST OF THE DIFFERENCE BETWEEN MEAN SCORES FOR THE TV AND NTV GROUPS ON THE CES

Group	N	Mean	SD	SE_M	Diff.	t	$t_{5\%}$
Television	56	54.27	9.71	1.03	5.71	2.30	2.00
Non-Television	56	59.98	15.43	2.06			

The non-television (face-to-face) discussion students had a more favorable opinion of the course than the students in the television group. This being the case it would be of further interest to learn on what particular items of the Course Evaluation Scale these differences were significant, and therefore each of the items on the CES was tested for differences between the TV and NTV groups by the "t" test.

For the computations in these tests the total sample of 167 students was used. It will be recalled that in order to maintain proportionality between cells in the analysis of variance and covariance tests the TV group was reduced by 30 students and the NTV group by 25 students. Since the "t" test places no restrictions on the proportionality of subjects these 55 students were included. The two groups were first tested for the difference between total means scores and the result is shown in Table XIV on page 117. The result of this test agrees with the findings in Table XII, and further substantiate the findings that the NTV group displayed a more favorable opinion toward the course than the TV group.

The "t" test of the individual items on the CES will be found in

TABLE XIV

RESULT OF THE "t" TEST OF THE DIFFERENCE BETWEEN MEAN SCORES FOR THE TV AND NTV GROUPS ON THE CES FOR THE TOTAL SAMPLE

Group	N	Mean	SD	SE_M	Diff	t	$t_{5\%}$
Television	86	53.34	8.71	0.94	6.83	3.34	1.96
Non-Television	81	60.17	15.95	1.77			

Table XV on page 118. The items are so scored that a mean value of 3.00 would indicate a completely neutral response, a score of above 3.00 a favorable response, and a score below 3.00 an unfavorable response. The Course Evaluation Scale will be found in Appendix A.

For ten of the twenty-two items on the CES a significant difference was found in favor of the NTV group, and no significant differences were obtained favoring the TV group. In examining the results for each item it will be noted that for Item 1 the NTV group considered the information in the course more useful than did the TV group and they also felt it to be more interesting, as revealed by the responses to Item 3. Inspection of Item 4 indicates that the NTV group were happier about having taken the course than the TV group, and according to Item 5 were more certain about taking an advanced course in Educational Psychology. The TV group felt that they had learned somewhat less than they would have learned in a similar course taught without television, while the NTV group felt they had learned slightly more as indicated by responses to Item 8. Analysis of Item 17 reveals that the NTV group felt that television instruction was somewhat less disturbing than ordinary in-

TABLE XV

COMPARISON OF THE TV AND NTV GROUPS ON THE INDIVIDUAL ITEMS OF THE CES

Test Item	TV Mean	NTV Mean	Diff.	SE_D	t	$t_{5\%}$	Null Hypothesis
1. As far as the information in this course is concerned, a. I think it will be very useful in my future teaching. b. I think it will be somewhat useful in my future teaching. c. I am not sure whether or not it will be useful in my future teaching. d. I think it will be of little use in my future teaching. e. I think it will be of no use in my future teaching.	4.24	4.85	0.61	0.11	2.33	1.96	Reject
2. In the future whenever there is an opportunity to find out more about the subject-matter in this course, a. I will positively do so. b. I will probably do so. c. I don't know what I will do. d. I will probably not do so. e. I will positively not do so.	3.73	3.90	0.17	0.11	1.55	1.96	Accept
3. The information in this course is, a. very interesting to me. b. somewhat interesting to me. c. neither interesting or uninteresting to me. d. somewhat uninteresting to me. e. very uninteresting to me.	4.03	4.38	0.35	0.11	3.18	1.96	Reject

TABLE XV (cont'd.)

Test Item	TV Mean	NTV Mean	Diff.	SE_D	t	$t_{5\%}$	Null Hypothesis
4. Now that I have had some experience in this course, a. I am very glad I took the course. b. I am glad I took the course. c. I am neither glad nor sorry I took the course. d. I am sorry that I took the course. e. I am very sorry I took the course.	3.48	3.92	0.44	0.14	3.14	1.96	Reject
5. If I have the opportunity to take an advanced course similar to this one, a. I will certainly do so. b. I will probably do so. c. I don't know what I will do. d. I will probably not do so. e. I will certainly not do so.	3.22	3.60	0.38	0.15	2.53	1.96	Reject
6. The subject-matter in this course is, a. very practical. b. somewhat practical. c. neither practical nor impractical. d. somewhat impractical. e. very impractical.	4.18	4.28	0.10	0.11	0.91	1.96	Accept
7. In comparing the lecture and discussion portions of this course, I feel that the lecture, a. is much more informative than the discussion. b. is somewhat more informative than the discussion. c. and discussion are equally informative. d. is somewhat less informative than the discussion. e. is much less informative than the discussion.	4.40	4.47	0.07	0.13	0.54	1.96	Accept

TABLE XV (cont'd.)

	Test Item	TV Mean	NTV Mean	Diff.	SE_D	t	$t_{5\%}$	Null Hypothesis
8.	Do you think that you have learned as much over television as you would have learned in a similar course, but taught without the use of television? a. Much more. b. A little more. c. About the same. d. A little less. e. Much less.	2.71	3.04	0.33	0.16	2.06	1.96	Reject
9.	Do you think this course is more or less valuable than it would have been if television had not been used? a. Much more valuable. b. Somewhat more valuable. c. About the same. d. Somewhat less valuable. e. Much less valuable.	3.00	3.29	0.29	0.16	1.81	1.96	Accept
10.	As far as class participation is concerned I would have liked to participate, a. much more. b. a little more. c. no more than at present. d. a little less. e. much less.	4.17	4.18	0.01	0.11	0.09	1.96	Accept

TABLE XV (cont'd.)

	Test Item	TV Mean	NTV Mean	Diff.	SE_D	t	$t_{5\%}$	Null Hypothesis
11.	In comparing the panel discussion groups and 6 x 6 discussion groups, I feel that the 6 x 6 discussion groups a. were much more helpful than the panel discussion groups. b. were somewhat more helpful than the panel discussion groups. c. were neither more nor less helpful than the panel discussion groups. d. were somewhat less helpful than the panel discussion groups. e. were much less helpful than the panel discussion groups.	4.03	4.11	0.08	0.14	0.07	1.96	Accept
12.	As far as the time devoted to discussion in the course, I would have liked to have spent a. much more time on discussion. b. a little more time in discussion. c. no more nor no less time in discussion. d. a little less time in discussion. e. much less time in discussion.	3.47	3.39	0.08	0.16	0.50	1.96	Accept
13.	Aside from the technical difficulties involved in presenting movies over television, I feel that the movies a. were very useful. b. were somewhat useful. c. were neither useful nor useless. d. were somewhat useless. e. were very useless.	4.43	4.21	0.22	0.12	1.83	1.96	Accept
14.	As far as the reading and reporting of journal articles once weekly is concerned, a. I think it is very helpful. b. I think it is somewhat helpful. c. I am not sure if it was helpful. d. I think it was of little help. e. I think it was of no help.	3.28	3.21	0.07	0.18	0.39	1.96	Accept

TABLE XV (cont'd.)

	Test Item	TV Mean	NTV Mean	Diff.	SE_D	t	$t_{5\%}$	Null Hypothesis
15.	As far as the participation in panels was concerned, I would have liked a. to participate much more in panel discussions. b. to participate somewhat more in panel discussions. c. to participate neither more nor less in panel discussions. d. to participate somewhat less in panel discussion. e. to participate much less in panel discussion.	3.60	3.82	0.22	0.14	1.57	1.96	Accept
16.	As far as the subject-matter of the lecture in comparison to textbook material was concerned, a. they were too much alike. b. they were somewhat alike. c. they were about right. d. they were somewhat unlike. e. they were too much unlike.	3.16	3.29	0.13	0.12	1.08	1.96	Accept
17.	As far as the television instruction is concerned, I think it is a. much less disturbing than ordinary courses. b. somewhat less disturbing than ordinary courses. c. no more disturbing than ordinary courses. d. somewhat more disturbing than ordinary classes. e. much more disturbing than ordinary courses.	2.70	3.24	0.54	0.18	3.00	1.96	Reject
18.	As far as television instruction is concerned, I think that I can concentrate a. much better than in ordinary classes. b. somewhat better than in ordinary classes. c. no more nor less than in ordinary classes. d. somewhat poorer than in ordinary classes. e. much poorer than in ordinary classes.	2.96	3.40	0.44	0.19	2.32	1.96	Reject

TABLE XV (cont'd)

	Test Item	TV Mean	NTV Mean	Diff.	SE_D	t	$t_{5\%}$	Null Hypothesis
19.	As far as the instruction is concerned, I feel that it is a. well above average. b. above average. c. about average. d. below average. e. well below average.	3.46	3.91	0.45	0.12	3.75	1.96	Reject
20.	If I had it to do again, knowing what I now know, I would a. positively choose a large television section again. b. probably choose a large television section again. c. have no preference. d. probably would not choose a large television section again. e. positively would not choose a large television section again.	2.36	3.00	0.64	0.20	3.20	1.96	Reject
21.	I feel that the objectives of the course have been a. very well met. b. well met. c. satisfactorily met. d. poorly met. e. very poorly met.	3.53	3.71	0.18	0.13	1.38	1.96	Accept
22.	As far as the instructors in this course are concerned, I think they are a. among the best. b. better than most. c. about average. d. poorer than most. e. among the poorest.	3.58	3.99	0.41	0.12	3.42	1.96	Reject

struction, whereas the TV group considered it somewhat more disturbing. Much the same feelings are expressed in Item 18 in which the TV group felt that television instruction interfered with their ability to concentrate, but the NTV group felt that they could concentrate somewhat better than in classes without television. Although both groups indicated that the instruction was somewhat above average, Item 19, a significant difference was found again in favor of the non-television group. Item 20 shows that the non-television students had no preference about choosing or not choosing a large television section again, but the television students tended to feel they would not choose a large television section again, the differences being significant. In comparisons of the instructors in the course with other instructors both groups indicated that they were above average, but again the NTV group was significantly higher than the TV group.

Although both sections were taught by the same instructors, the students in the television section were not as favorable in their opinion of the instructors as were the students in the non-television section. Neither did the TV students feel that they had learned as much as the NTV students although the results on the Achievement Test showed no differences in this respect. It is apparent that the students in the television group were not as happy with the course as were the students in the face-to-face situation.

The free comments on the Course Evaluation Scale were tabulated into three groups as follows: (1) Complimentary Comments such as "I think the instructors in the course were among the best", (2) Helpful

Comments such as "Discuss more practical problems", and (3) Complaining Comments such as "I found television boring". The tabulation of these results is shown in Table XVI below. Some of the students made comments that were classifiable into more than one category and these were so classified. A Chi-square analysis of these results shows no differences between the TV and NTV groups.

TABLE XVI

CHI-SQUARE ANALYSIS OF THE FREE COMMENTS ON THE COURSE EVALUATION SCALE FOR THE TV AND NTV GROUPS

	N	Complimentary Comments	Helpful Comments	Complaining Comments	Total	X^2	$X^2_{5\%}$
TV group	86	9	26	25	60	0.33	6.00
NTV group	81	9	19	19	47		
Total	167	18	45	44	107		

Comparing the number of students in each section who responded by making a comment with the total number of students in each section who could have responded, revealed a Chi-square value significant at the 1% level of confidence, as shown in Table XVII on page 126. Evidently the students in the television section felt more of a need to make some kind of a comment about the course than did the students in the non-television section, indicating a dissatisfaction with the course.

From the results of the findings on the Course Evaluation Scale it is clearly indicated that the students undergoing discussion by means of the television technique held a significantly less favorable

opinion about the course than did the students in the face-to-face situation. It may be inferred that the inter-personal relationships

TABLE XVII

CHI-SQUARE ANALYSIS OF CES FOR TV AND NTV GROUPS ON THE BASIS OF COMMENTS AND NO COMMENTS

	TV	NTV	Total	X^2	$X^2_{1\%}$
Comment	48	37	85	7.88	6.64
No Comment	38	44	82		
Total	86	81	167		

between the students and the instructors in the television group were not as good as those displayed in the non-television group, and that the students in the former group displayed more hostility as evidenced by the significantly greater number of comments made. Such findings bear out the argument presented in Chapter One that the absence of maximum feedback would result in the lowering of class morale and a corresponding increase in the amount of hostility and insecurity displayed by these students.

VI. SUMMARY

The basic data for each individual was key-punched on IBM cards and the basic calculations were accomplished by the use of electronic calculating machine. All other computations were performed on a hand mechanical calculator. The original data for each student and summarization of subgroup data is shown in the Appendixes.

Analysis of the data obtained on the Pre-test and Post-test measures of the Achievement Test indicates that there were no significant differences due to the ability level of the students, the preference for instruction of the students, the sex of the students, or the two discussion techniques. Nor were there significant differences due to the interaction effects of these variables.

Analysis of the Pre-test and Post-test data obtained on the Misconceptions about Education Scale indicates no significant differences due to any of the main or interaction effects, and all null hypotheses were accepted.

Analysis of the Pre-test and Post-test data obtained on the Minnesota Teacher Attitude Inventory showed a significant interaction effect between the preference for instruction and the discussion techniques variables. Interpretation of this finding seems to indicate that under the conditions of the present investigation the combination of face-to-face discussion techniques and a preference for a cognitive type of instruction produces the most favorable attitudes toward teaching as measured by the Minnesota Teacher Attitude Inventory. All other null hypotheses for this criterion were accepted.

Analysis of the Course Evaluation Scale indicates a significant difference between the students due to the discussion techniques. A significant difference between the students in the television and non-television groups were found for the total scale and for ten of the twenty-two items on the scale all of which favored the non-television students. Further analysis of the free comments by the students indi-

cated that the students in the television section made a significantly greater number of comments than the students in the face-to-face section. Interpretation of these findings seems to indicate that the students in the face-to-face situation held a more favorable opinion of the course than the students in the television discussion situation. The morale of these students was higher and they displayed less insecurity and hostility than the television students as inferred from the fewer number of comments. All other null hypotheses concerning the Course Evaluation Scale were accepted.

CHAPTER VII

SUMMARY AND CONCLUSION

I. SUMMARY

The Problem

It was the primary purpose of the present study to determine the relative effectiveness of face-to-face discussion and discussion by means of television on educational outcomes in a large class. The use of instructional television has increased rapidly in the past few years, but little is known about its efficacy in the discussion situation. Research evidence indicates that the discussion method is a superior instructional technique for teaching problem-solving techniques and where the objective is the learning of abstract materials. Learning theory and the evidence from at least one study in communication theory suggests that feedback between the instructor and the students is an important factor in the learning of abstract material. It was theorized that under conditions of television discussion the absence of maximum feedback would prove a handicap to both the instructor and the student, and would result in less satisfactory educational outcomes as well as a lowering of student morale and an increase in insecurity and hostility on the part of the student.

A review of the instructional television literature revealed that the type and quality of studies conducted so far has contributed little to an understanding of the problem and that few studies have been

concerned with the discussion techniques.

The Design

The study employed a four-way factorial design with two levels in each of the classification variables. The four independent variables were: (1) the discussion techniques, (2) the ability level of the students, (3) the preference for instruction of the students, and (4) the sex of the students. The combination of two levels for each of the independent variables allowed the formulation of fifteen testable null hypotheses of which four were single or main effects and eleven were interacting effects. Educational outcomes were measured on the basis of four dependent variables: (1) achievement of course content, (2) attitudes toward teaching, (3) misconceptions about education, and (4) opinion of the course. The design made it possible to test each of the dependent variables for the fifteen null hypotheses thus allowing for a total of sixty testable null hypotheses. The control of variable errors was afforded by randomization of subjects, rotation of instructors between the discussion groups, and the use of covariance statistical analysis.

Experimental Procedures

The experiment was conducted in an Educational Psychology course. Facilities included an originating room, two viewing classrooms, and the necessary equipment for television production. The population consisted of all Michigan State University students enrolling in Section 4, of Foundations of Education 200 in the Fall Term of 1957. The final sample

consisted of one-hundred and sixty-seven students. The students were randomly assigned to a television and a non-television group. In the television group discussion was conducted over the closed-circuit television system and in the non-television group discussion was conducted in the classroom in a face-to-face situation.

Evaluation Instruments

The ability level of the students was measured by the American Council on Education Psychological Examination. The preference for instruction was measured by the Preferred Instructor Characteristics Scale, an instrument designed to measure the students' preference on an affective-cognitive continuum wherein "affective" denotes a preference for the inter-personal, student-centered type of instruction and "cognitive" denotes a preference for a subject-centered, intellectual type of instruction. The PICS has a reliability of about .90.

Achievement was measured by a fifty-item multiple-choice achievement test over the course content. Reliability of this instrument was found to be about .52. Attitudes toward teaching was measured by the Minnesota Teacher Attitude Inventory, an instrument designed to measure the students' score on attitude statements about teaching in comparison with successful teachers. The MTAI has a reliability of about .90.

Misconceptions about education was measured by the Misconceptions about Education Scale, designed by the writer and a colleague to measure the students conceptions about irrational beliefs in education and psychology. The MES has a reliability of about .63.

Student opinion about the course was measured by the Course Evalu-

ation Scale, adapted from a similar instrument used in research at Pennsylvania State University. It consisted of twenty-two items about the instruction and the course designed after the Likert technique. Additional space was provided for free comments.

The basic data was summarized by use of an electronic computer and completed by use of a mechanical calculator. The data was tested by the analysis of variance and covariance, the "t" test, and the Chi-square test. The level of confidence for all test of assumptions was set at the 1% level of confidence, and the level of confidence for the criterion tests was set at the 5% level.

Results

1. Achievement criterion - There are no significant differences in achievement due to the ability level of the students, the preference for instruction of the students, the sex of the students, or the discussion techniques, nor are there any significant interaction effects due to the above variables and all null hypotheses are accepted.

These findings agree substantially with other research in the use of instructional television. Many experimenters have interpreted such findings to mean that television instruction is equally effective as conventional instruction. It may be, however, that the measuring instruments are not sensitive enough to detect differences, and this was possibly the case in the present investigation since the Achievement Test had a low reliability.

2. Misconceptions about Education criterion - There are no significant differences due to the ability level, preference for instruc-

tion, or sex of the students, or the discussion techniques, nor are there any significant interaction effects due to the above variables and all null hypotheses are accepted.

In only one other study has a similar instrument been used. A Misconceptions about Psychology test was used in the Miami University study (39), and in comparisons of large class and control sections of small classes a significant difference was found favoring the control classes.

In the present study the lack of significant differences may have been due in part to the fact that the instrument was weak from the standpoint of reliability.

3. Attitude toward Teaching criterion - There is a significant interaction effect due to the preference for instruction and the discussion techniques variables, and the null hypothesis concerning the PICS x TECH interaction is rejected. All other null hypotheses were accepted. Interpretation of this interaction indicates that the combination of students with a preference for a "cognitive" type of instruction and the face-to-face discussion technique is most conducive to the promotion of attitudes toward teaching as measured by the Minnesota Teacher Attitude Inventory under the conditions of the present experiment.

No other studies to the writer's knowledge have used these instruments in instructional television research, and therefore no comparisons can be made. The effect obtained in the present study could have been obtained on a chance basis alone, however, the phenomenon

does merit more study.

4. Opinion about the course criterion - There is a significant effect due to the discussion techniques variable and the null hypothesis concerning the discussion variable is rejected. All other null hypotheses are accepted. Interpretation of this finding indicates that the students under the conditions of face-to-face discussion held a more favorable opinion of the course than the students under the conditions of television discussion. Comparisons of the two groups on individual items of the scale verified these findings. Comparisons of free comments made by the students indicated a significantly greater number of criticisms were made by the students under the conditions of television discussion. It is suggested that this is an indication of the lower morale of the students in the television discussion group and evidence of greater hostility and insecurity.

Limitations of the Study

The present study was not without its limitations. Replication was not provided for in the present study. Verification of results and less hazardous generalizations are afforded by replication. It was previously discussed that replication in the present investigation was impractical, nevertheless the study was limited in this respect and its generalizability less certain.

A second limitation of the present study was the weaknesses of two of the criterion instruments, the PAT and the MES, as indicated by their low reliability. It is possible that more significant findings would have been obtained had these instruments been more carefully

planned and tested at the outset of experimentation. The findings of any study even negative findings, are more satisfying to the experimenter and more important to his profession, when he is certain that the instruments he is using are yielding maximum results. The most obvious method for rectifying this limitation would be to lengthen both tests.

No matter how carefully the experimenter attempts to determine and control sources of error prior to experimentation, frequently during the course of the experiment or in the analysis and interpretation process, additional sources of bias are discovered. As was suggested in the discussion relative to the findings of a significant interaction between the PICS x TECHNIQUES variables, it was possible for instructor bias to have contaminated these results. The instructors may have displayed different degrees of "cognitiveness" and "affectiveness" under the conditions of face-to-face and television discussion. However, even if this bias did exist it bears out the original argument that under conditions of partial feedback both the instructor and the student are handicapped and, in effect, merely intensified the conditions of partial feedback already existing.

The effect of the lecture period in the present study cannot be separated from the other effects although it was equalized for both groups of students. A decisive study of instructional television discussion must limit the experimental conditions to discussion only. The presence of the lecture may have been the factor that results in no differences being found on the achievement and misconceptions criterion. Inasmuch as the experiment was an attempt to study the problem in a real-

istic and practical situation, this is not considered a serious limitation.

II. CONCLUSIONS

Conclusions based on the findings of the present study cannot logically be carried beyond the present experimental situation without further replication and research. With these limitations in mind the following conclusions seem warranted:

1. The results of the achievement criterion and the misconceptions about education criterion are inconclusive. There is no evidence to indicate the superiority of either the face-to-face discussion or the television discussion methods for producing educational outcomes on these criteria. Neither is it possible to generalize from this evidence that the two discussion methods are equally effective in producing educational outcomes on these criteria.

2. The combination of face-to-face discussion method and "cognitive" students, as measured by the Preferred Instructor Characteristic Scale, is much more effective in increasing student attitudes toward education, as measured by the Minnesota Teacher Attitude Inventory, than is the combination of face-to-face discussion and "affective" students.

3. Television discussion is about equally effective for "cognitive" and "affective" students in increasing attitudes toward teaching, but is much less so than face-to-face discussion with "cognitive" students.

4. Face-to-face discussion produces more favorable student

opinions about the course than does television discussion.

5. Face-to-face discussion results in fewer student criticisms of the course than does television discussion.

III. SUGGESTIONS FOR FUTURE RESEARCH

The majority of instructional television experimentation to date, in comparing television and conventional classroom instruction, has found no differences to exist in the achievement of course content. This finding has been interpreted as indicating that the two methods are equally effective. Where achievement is defined as the comprehension of factual and informational data about the course, and it is so defined in the vast majority of studies, such an interpretation is probably relevant. In another context, however, achievement may be defined as the ability to apply principles and concepts, to draw inferences, to synthesize, to perceive relationships, and to evaluate. Is television instruction as effective as conventional classroom methods where the above objectives are to be met? No definitive studies have as yet been made to answer this question, mainly because the criterion instruments have not been constructed to measure such objectives. Future research should be concerned with developing more sensitive instruments to measure the above objectives and in applying these instruments to the problem of the effectiveness of the two types of instruction.

The closely related problem of attitude changes would seem a needed area for future research. Evidence from the present study indicates that face-to-face instruction is more effective than television

instruction for certain kinds of students in producing attitude changes. This interrelationship needs to be investigated much more thoroughly. What subject-matter attitudes are most affected by different types of instruction? What types of students are most affected in their attitude changes? What is the interrelationship between attitudes, student preferences, and types of instruction? These are but a few of the questions that need to be answered.

What is the relationship between student morale and learning? Does the fact that a student has an unfavorable opinion of a course mean that he will not learn as well as students whose opinion is more favorable? If such a relationship exists, is it as true for the learning of factual information as for the learning of more emotionally laden attitudes? Is good classroom morale more important for some students than for others? How important are teacher-student relationships? Do poor relationships actually interfere with learning for all types of objectives? How do they affect teachers? Are some teachers affected differently than others? The answers to these questions would seem to be important for television instruction.

The problem of feedback in itself is an intriguing one. Does the instructor acquire cues that alter his instruction in a face-to-face situation that he would not acquire in a less personal situation? What cues are important? Are different cues important for the learning of different kinds of materials? Are some students better able to communicate cues to the instructor than other students? Are some instructors more capable of recognizing important cues than other instructors?

Again the answers to these questions are directly related to the problems of television instruction and discussion methods.

The problems suggested by the Preference for Instruction Characteristic Scale are likewise intriguing. This is as yet a new and untried instrument that seems to have promise for research. Is there a relation between the students' emotional needs and his preference for certain kinds of instruction? In what particular kinds of instruction will students with different preferences profit most? What is the relationship between intellectual ability and preference for instruction? Are some instructors better able to meet the needs of certain students than other instructors? Do different instructors prefer different kinds of students? Is there a relationship between the students' preferences on this scale and his vocational choice?

There are many more problems that need to be answered in the use of television as an instrument for instruction than have been answered thus far by research. The use of television in the schools will no doubt continue to grow, and it is vitally important that its strengths and weaknesses be much more thoroughly understood than they are understood at the present time.

BIBLIOGRAPHY

1. Allen, M. R. Quartermaster Training Command Educational-Television Study. Fort Lee, Va.: Quartermaster School, Quartermaster Training Command, June 4, 1954.

2. Anderson, G. R. and Vander Meer, A. W. "A Comparative Study of the Effectiveness of Lessons on the Slide Rule Presented via Television and in Person," The Mathematics Teacher. 1954, 47: 323-27.

3. Asch, Morton Jay. "Nondirective Teaching in Psychology: An Experimental Study," Psychological Monographs, 1951, Vol. 65, No. 321.

4. Bane, Charles L. The Lecture in College Teaching. Boston: Richard G. Badger, 1931.

5. Becker, S. L., Murray, James N. Jr., and Bechtoldt, H. P. Teaching by the Discussion Method. A Study supported in part by a grant from the Fund for the Advancement of Education. Iowa City, Iowa: State University of Iowa, 1958.

6. Bloom, B. S. "Thought-Processes in Lecture and Discussion," Journal of General Education, 1953, 7: 160-69.

7. Bond, Betty Wells. "The Group Discussion—Decision Approach—An Appraisal of its Use in Health Education," Dissertation Abstracts, 1956, 16: 903-4.

8. Boone, W. F. Evaluation of the U. S. Naval Academy Educational Television as a Teaching Aid. Annapolis: United States Naval Academy, 7010 7-26-54, October 29, 1954.

9. Bovard, Everett W., Sr. "The Psychology of Classroom Interaction," Journal of Educational Research, 1951, 45: 215-24.

10. Brinkley, S. G. "Mental Activity in College Classes: Student Estimate of Relative Value of Ten Learning Situations," Journal of Experimental Education, 1953, 20: 373-78.

11. Carlson, Carl Raymond. "A Study of the Relative Effectiveness of Lecture and Directed Discussion Methods of Teaching Tests and Measurements to Prospective Air Force Instructors," Dissertation Abstracts, 1953, 13: 1112-13.

12. Carpenter, C. R. and Greenhill, L. P. An Investigation of Closed-Circuit Television for Teaching University Courses. Instructional Television Research Project Number One. Project sponsored by the Fund for the Advancement of Education. University Park, Pennsylvania: The Instructional Film Research Program, Pennsylvania State University, July 1955.

13. Closed-Circuit Television Instruction. Four duplicated preliminary reports. Lafayette: Purdue University, 1956.

14. Cronbach, Lee J. Educational Psychology. New York: Harcourt, Brace and Company, 1954.

15. Dowell, E. C. An Experiment in Training by Television. Keesler Air Force Base, Mississippi: 3380th Technical Training Group, Project 53-32, February 1954.

16. Dunham, Franklin and Lowdermilk, Ronald R. Television in our Schools. U. S. Department of Health, Education and Welfare, Bulletin 1952, No. 16. Washington: U. S. Government Printing Office, 1956.

17. Edmitson, R. W. and Braddock, R. W. "A Study of the Effect of Various Teaching Procedures upon Observed Group Attention in Secondary Schools," Journal of Educational Psychology, 1941, 32: 665-72.

18. Eglash, Albert. "A Group Discussion Method of Teaching Psychology," Journal of Educational Psychology, 1954, 45: 257-67.

19. Evans, R. E., Roney, H. B. and McAdams, W. J. "An Evaluation of the Effectiveness of Instruction and Audience Reaction to Programming on an Educational Television Station," Journal of Applied Psychology, 1955, 39: 277-79.

20. Farquhar, William W. An Investigation of the Relationship of Three Teaching Methods to Student Behavior in a How to Study Course. Unpublished Ph.D. Dissertation, University of Minnesota, 1955.

21. Farquhar, William W. and Krumboltz, John D. "A Check List for Evaluating Experimental Research in Psychology and Education," Journal of Educational Research, in press.

22. Faw, Volney. "A Psychotherapeutic Method of Teaching Psychology," American Psychologist, 1949, 4: 104-9.

23. Gerberich, J. R. and Warner, K. O. "Relative Instructional Efficiencies of the Lecture and Discussion Methods in a University Course in American National Government," Journal of Educational Research, 1936, 29: 574-79.

24. Hudelson, Earl. "Class Size at the University Level," Problems of College Teaching. Minneapolis: University of Minnesota Press, 1928, pp. 403-20.

25. Husband, R. W. "Television versus Classroom for Learning General Psychology," American Psychologist, 1954, 9: 181-83.

26. Johnson, Donald M. and Smith, Henry C. "Democratic Leadership in the College Classroom," Psychological Monographs, 1953, 68: No. 361.

27. Kanner, J. H., Runyon, R. P., and Desiderato, O. Television in Army Training: Evaluation of Television in Army Basic Training. Washington: Human Resources Research Office, The George Washington University, Technical Report No. 14, November 1954.

28. Kelley, Truman Lee. Interpretation of Educational Measurements. New York: World Book Co., 1927.

29. Krumboltz, John D. Alphabetical Bias in Certain Aptitudes by Geographical Areas of the United States. Unpublished study. 1957.

30. Krumboltz, John D. An Investigation of the Effect of Three Teaching Methods on Motivational Outcomes in a How to Study Course. Unpublished Ph.D. Dissertation, University of Minnesota, 1955.

31. Kumata, Hideya. An Inventory of Instructional Television Research. Ann Arbor, Mich.: Educational Television and Radio Center. December 1, 1956.

32. Leavitt, Harold J. and Mueller, Ronald A. H. "Some Effects of Feedback on Communication," Small Groups: Studies in Social Interaction, ed. by A. Paul Hare, Robert Bales, and Edgar F. Borgotta. New York: Alfred A. Knopf, 1955.

33. Levine, Jacob and Butler, John. "Lecture vs Group Decision in Changing Behavior," Journal of Applied Psychology, 1952, 36: 29-33.

34. Lewin, Kurt. "Group Decision and Social Change," Readings in Social Psychology, ed. by Theodore M. Newcomb and Eugene Hartley. New York: Henry Holt and Company, 1947, pp. 330-44.

35. Lewin, Kurt, Lippitt, Ronald and White, Ralph K. "Patterns of Expressive Behavior in Experimentally Created 'Social Climates'," Journal of Social Psychology, 1939, 10: 271-99.

36. Likert, R. "A Technique for the Measurement of Attitudes," Archives of Psychology, No. 40, 1932.

37. Lindquist, E. F. Statistical Analysis in Educational Research. New York: Houghton Mifflin Co., 1940.

38. Lindquist, E. F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin Co., 1953.

39. Macomber, F. G., and others. Experimental Study in Instructional Procedures. Project under grant from the Fund for the Advancement of Education. Oxford, Ohio: Miami University, October 1, 1957.

40. Maier, Norman R. F. "An Experimental Test of the Effect of Training on Discussion Leadership," Human Relations, 1953, 6: 161-73.

41. Michigan State University Catalog. East Lansing, Michigan: Michigan State University Publication, April 1958, Vol. 52, No. 10.

42. Packman, Martin. "Educational Television," Editorial Research Reports, May 18, 1954, Vol. 1, No. 19, p. 376.

43. Parsons, T. S. A Comparative Analysis of Some Outcomes of Instruction by Kinescope, Correspondence Study and Classroom Procedures. Duplicated report of a paper read at the APA convention, San Francisco, Sept. 1955. Ann Arbor, Michigan: School of Education, University of Michigan, 1955.

44. Pasewark, W. R. The Effectiveness of Television as a Medium of Learning Typewriting. Unpublished Ph.D. dissertation, New York University, 1956.

45. Paul, J. and Ogilvie, J. C. "Mass Media and Retention," Explorations, 1955, 4: 120-23.

46. Pollock, T. C. Closed-Circuit Television as a Medium of Instruction, 1955-56: A Report on New York University's First Year of Experimentation with Television in College Classrooms. Project sponsored by the Fund for the Advancement of Education. New York: New York University, October 1956.

47. Rickard, Paul B. "An Experimental Study of the Effectiveness of Group Discussion in the Teaching of Factual Content," Summaries of Doctoral Dissertations, Northwestern University. 1946, 14: 72-7.

48. Rock, R. T. Jr., Duva, J. S. and Murray, J. E. Training by Television: The Comparative Effectiveness of Instruction by Television Recordings and Conventional Classroom Procedures. Port Washington, L. I., N. Y.: Special Devices Center, SDC Report 476-02-2 (NAVEXOS P-850-2) No date.

49. Rock, R. T. Jr., Duva, J. S. and Murray, J. E. Training by Television: A Study in Learning and Retention. Port Washington, L. I., N. Y.: Special Devices Center, SDC Report 476-02-3 (NAVEXOS P-850-3) No date.

50. Ruja, Harry. "Outcomes of Lecture and Discussion Procedures in Three College Courses," Journal of Experimental Education, 1954, 22: 368-94.

51. Seipman, Charles A. Television and Education in the United States. Paris: United Nations Educational, Scientific and Cultural Organization. 1952.

52. Shimberg, B. Effectiveness of Television in Home Nursing. Princeton, N. J.: Educational Testing Service, Research Bulletin (RB-54-19), August 1954.

53. Spence, Ralph B. "Lecture and Class Discussion in Teaching Educational Psychology," Journal of Educational Psychology, 1928, 19: 454-62.

54. Super, Donald M. Appraising Vocational Fitness by Means of Psychological Tests. New York: Harper and Brothers, 1949.

55. Tannenbaum, P. H. Instruction Through Television: A Comparative Study. Urbana: Institute of Communication Research, University of Illinois, June 1956.

56. Walker, Helen and Lev, Joseph. Statistical Inference. New York: Henry Holt and Co. 1953.

57. Ward, John N. "Group-Study versus Class Discussion in Teaching Educational Psychology," Journal of Educational Psychology, 1956, 14: 197-210.

58. Wickman, E. K. Children's Behavior and Teachers' Attitudes. New York: Commonwealth Fund, 1928.

59. Williams, D. C. "Mass Media and Learning--An Experiment," Explorations, 1954, 3: 75-82.

60. Willis, B. C. Evaluation Report of the Two Week Experiment of Direct Teaching on Television. Chicago: Chicago Public Schools, 1956.

SUPPLEMENTARY REFERENCES

Barr, A. S., Davis, R. A. and Johnson, P. O. Educational Research and Appraisal. New York: J. B. Lippincott Co., 1953.

A Manual of Style. 11th ed. Chicago: The University of Chicago Press. 1949.

Campbell, William G. Form and Style in Thesis Writing. Boston: Houghton Mifflin Co., 1954.

Dixon, W. J. and Massey, F. J. Jr. Introduction to Statistical Analysis. New York: McGraw-Hill Book Co. 1951.

Garrett, Henry E. Statistics in Psychology and Education. New York: Longmans, Green and Co. 1947.

Good, Carter V. and Scates, Douglas E. Methods of Research. New York: Appleton, Century-Crofts, Inc. 1954.

Hillway, Tyrus. Introduction to Research. Boston: Houghton Mifflin Co. 1956.

Johnson, P. O. Statistical Methods in Research. New York: Prentice-Hall, Inc. 1949.

Perrin, P. G. Writer's Guide and Index to English. New York: Scott, Foresman and Co. 1950.

APPENDIX A

EVALUATION INSTRUMENTS

F S J S SP M F TV NTV

Name: _____ Student No. _____ Grade Pt. Ave. _____
Last First

Birth Date: _____ 19____ Seat No. _____ Room No. _____

College Address: _____ Telephone No. _____

Permanent Address: _____ City: _____ State: _____

	Pre	Post	Pre	Post
Ach.	___	___	___	___
ACE	___	___	___	___
MTAI	___	___	___	___
MES	___	___	___	___
PICS	___	___	___	___

Major: _____

Educ. Area: El. Ed. _____

Sec. _____

Higher _____

Spec. _____

Other (indicate)
_____

PERSONAL DATA CARD

CODING

F - Freshman
S - Sophomore
J - Junior

S - Senior
SP - Special
M - Male

F - Female
TV - Television
NTV - Non-Television

Preferred Instructor Characteristics Scale

Directions:

What kind of an instructor do you prefer? In the following items you will find two instructor characteristics paired. From each pair choose the one characteristic you most prefer. Then mark your choice in the proper column on the special answer sheet. Do not omit any items. This is to find out your preferences. There are no right or wrong answers.

I prefer an instructor who:

1. a. is an expert.
   b. treats us as mature people.
2. a. makes the classroom pleasant.
   b. thinks logically.
3. a. understands our point of view.
   b. is well known in his field.
4. a. is dedicated to his students.
   b. is dedicated to his subject.
5. a. thinks logically.
   b. is friendly.
6. a. is well known in his field.
   b. makes the classroom pleasant.
7. a. is interested in us.
   b. covers all the material.
8. a. is dedicated to his students.
   b. knows the theoretical background of his subject.
9. a. thinks logically.
   b. treats us as mature people.
10. a. is friendly.
    b. is well known in his field.
11. a. covers all the material.
    b. understands our point of view.
12. a. is interested in us.
    b. is dedicated to his subject.
13. a. is an expert.
    b. is dedicated to his students.
14. a. is well known in his field.
    b. treats us as mature people.
15. a. covers all the material.
    b. makes the classroom pleasant.
16. a. understands our point of view.
    b. is dedicated to his subject.
17. a. is interested in us.
    b. knows the theoretical background of his subject.
18. a. is friendly.
    b. covers all the material.

I prefer an instructor who:

19. a. makes the classroom pleasant.
    b. is dedicated to his subject.
20. a. knows the theoretical background of his subject.
    b. understands our point of view.
21. a. is interested in us.
    b. is an expert.
22. a. is dedicated to his students.
    b. thinks logically.
23. a. treats us as mature people.
    b. covers all the material.
24. a. is dedicated to his subject.
    b. is friendly.
25. a. makes the classroom pleasant.
    b. knows the theoretical background of his subject.
26. a. is an expert.
    b. understands our point of view.
27. a. is dedicated to his students.
    b. is well known in his field.
28. a. is dedicated to his subject.
    b. treats us as mature people.
29. a. is friendly.
    b. knows the theoretical background of his subject.
30. a. is an expert.
    b. makes the classroom pleasant.
31. a. thinks logically.
    b. is interested in us.
32. a. treats us as mature people.
    b. knows the theoretical background of his subject.
33. a. is an expert.
    b. is friendly.
34. a. thinks logically.
    b. understands our point of view.
35. a. is interested in us.
    b. is well known in his field.
36. a. is dedicated to his students.
    b. covers all the material.

Check to see if you left any blanks.

Scoring Key

Preferred Instructor Characteristics Scale (PICS)

Item	Cognitive Response	Item	Cognitive Response
1	a	19	b
2	b	20	a
3	b	21	b
4	b	22	b
5	a	23	b
6	a	24	a
7	b	25	b
8	b	26	a
9	a	27	b
10	b	28	a
11	a	29	b
12	b	30	a
13	a	31	a
14	a	32	b
15	a	33	a
16	b	34	a
17	b	35	b
18	b	36	b

P A T - F.E. 200

1. People are especially likely to remember
   (a) startling facts.
   (b) facts which they expected.
   (c) facts about which they had no previous background or expectation.

2. Interference between two learning situations can best be avoided
   (a) by making sure that a correct response is made in the first situation.
   (b) by warning the learner against transferring responses.
   (c) by teaching why a certain response is successful in the first situation.
   (d) by presenting the situations only after readiness has developed.

3. A psychologist trying to understand the significance of a delinquent's acts would be especially concerned about
   (a) his religious knowledge and beliefs.
   (b) his history of school achievement.
   (c) his attitudes toward authority.
   (d) his understanding of principles of good conduct.

4. Tension during an activity is eliminated if the person
   (a) has readiness for the activity.
   (b) has emotional support from classmates and teachers.
   (c) is certain he will reach his goals.
   (d) is genuinely interested in reaching the goals set.

5. When group decision is used to set goals,
   (a) the group members act in accord with the decision.
   (b) goals are set to fit the slower members of the group.
   (c) performance is close to capacity.

6. The chief fault of ability grouping is that
   (a) it makes no provision for social-emotional differences.
   (b) it works an emotional hardship on the less able group.
   (c) it cannot markedly reduce the range of individual differences.

7. Compared to older drill methods modern activity methods
   (a) place less emphasis on learning through active practice.
   (b) place more emphasis on learning through active practice.
   (c) make more provision for evaluation and reinforcement of responses.
   (d) require more active responses to meaningful complex situations.

8. An action does not show good adjustment if
   (a) it is unlikely to attain its purpose.
   (b) the person acts without considering the facts.
   (c) the person chooses a course of action which increases his tension.
   (d) the person leaves some of his needs unsatisfied.

9. Overlearning is primarily the result of
   (a) extra practice.
   (b) strong motivation.
   (c) pressure to pass examinations.
   (d) fear of failure.

10. The accepting teacher
    (a) does not judge whether the pupil's word is good.
    (b) considers the pupil's feelings and judgments.
    (c) tries to provide the pupil with continual success.

11. Constant criticism of the pupil's original ideas is quite likely to teach him
    (a) to insist on high standards.
    (b) to rebel against adults.
    (c) to follow others' suggestions.
    (d) to take responsibility for himself.

12. The basic vocabulary of fourteen-year-olds is in the neighborhood of _____ words.
    (a) 4,000
    (b) 10,000
    (c) 30,000
    (d) 40,000

13. Good adjustment is a matter of
    (a) contentment.
    (b) confidence.
    (c) co-opertiveness.
    (d) conscientiousness.

14. A traditional testing procedure may have the following bad effect:
    (a) it provides little incentive to study.
    (b) it provides the student no basis for judging which outcomes are considered important.
    (c) it rewards behavior which is contrary to the objectives.

15. The level of aspiration of poor students
    (a) will be set cautiously.
    (b) will be set unreasonably high.
    (c) cannot be predicted.
    (d) will vary widely from trial to trial.

16. If a student is to gain the right sort of satisfaction from performing a good action, the situation should be set up so that before the action the pupil is
    (a) challenged.
    (b) relaxed.
    (c) threatened.
    (d) mildly irritated.

17. Transfer of a behavior pattern learned in one situation to a new situation will occur if
    (a) the situations are the same in all perceptible respects.
    (b) the learner thinks of the situations as similar even if they actually are not.
    (c) the situations are closely similar and the learner sees this similarity.

18. In attitudinal learning, the model alters
    (a) the learner's readiness.
    (b) the situation to which response is made.
    (c) the provisional try.
    (d) the consequence of the response.

19. The most significant function of subject matter learning is
(a) teaching specific responses to problems.
(b) clarifying situations to be encountered in future living.
(c) developing readiness for later subject matter.
(d) teaching technical vocabulary.

20. A democratic home is best described as one
(a) where the children have as much influence on decisions as the parents do.
(b) where the children are allowed to make decisions about their own activities.
(c) where the children make decisions about family affairs.
(d) where the children's desires are freely expressed and considered in making decisions.

21. Anecdotal records placed in the school file should be
(a) a random sample of the pupil's behavior.
(b) incidents representing unusual behavior by this pupil.
(c) incidents which raise special questions or give new evidence about the pupil.
(d) incidents in which the pupil's difficulties are exemplified.

22. Psychology contributes to education by
(a) defining what goals the teacher should strive for.
(b) testing commonly held beliefs about learning and educational procedures.
(c) showing what methods or techniques the teacher should use.
(d) stressing "new methods" of teaching.

23. Frank (age 20) is sour and rude. He has a chip on his shoulder. A trained counselor would view him in this way:
(a) If Frank would change his attitude, he could be helped.
(b) Frank evidently doesn't know how to satisfy his needs.
(c) Frank's unfortunate character was probably fixed in early childhood.
(d) Frank's behavior would change if people were pleasant to him.

24. Allowing the student to express his disagreements with authority is especially important in teaching subjects
(a) that students enter with definite ideas.
(b) which arouse emotional reactions.
(c) where answers are supported by reasoning.

25. The particular advantage of a standardized reading test is that
(a) it presents the same task to all pupils.
(b) it permits a comparison of the pupil's score with the score expected for his grade.
(c) it presents selections unlike those from everyday books.
(d) all reading tests give essentially the same information.

26. The teacher who obtains a sociogram by asking pupils whom they wish to work with should form groups
(a) to put together pupils who choose each other.
(b) to break up mutual-choice pairs and other cliques.
(c) on the basis of educational needs without regard to the sociometric result.
(d) to put together pupils of equal popularity.

27. Practice on a skill is particularly valuable for
(a) improving understanding.
(b) reducing time required for each movement.
(c) developing discrimination.

28. Of these self-concepts, the most desirable one is:
(a) whatever I do is good.
(b) if I fail at something, it isn't very important.
(c) I am capable of reaching my goals.
(d) I am constantly alert to my weaknesses.

29. Which of these habits, considered by itself, would be strong evidence of maladjustment to an adult?
(a) Driving hard to attain political prominence.
(b) Finding fault with the work of others.
(c) Enjoying wildly imaginative fiction.
(d) Taking poor care of one's property.

30. General scholastic ability is a term used to refer to
(a) the pupil's record on standard achievement tests.
(b) innate mental ability.
(c) performance on a variety of problem-solving tasks.
(d) a combination of special abilities and talents.

31. Which of these would be most likely to suggest faulty emotional development in a ninth-grade girl?
(a) Lack of interest in boys.
(b) Striving for perfection in all her work.
(c) Getting along smoothly with teachers.
(d) Strong interest in music, with only passive interest in other schoolwork.

32. The basic question in determining objectives is:
(a) What learning will be needed in later courses?
(b) What does the community want a young person to know?
(c) What behaviors are needed to perform successfully in this society?
(d) What can each of the school subjects contribute to development?

33. "Developmental tasks" are those learnings which
(a) the culture expects all members to master at a particular age.
(b) depend primarily on physical maturation.
(c) the child must complete before he is ready for school.
(d) are nearly complete by the time the child reaches puberty.

34. An adolescent boy is asked to climb a ten-foot ladder, and shows panic at the danger of falling. This is best interpreted as showing
(a) a specific fear conditioned in infancy.
(b) a general attitude of insecurity.
(c) negativism stemming from difficulty with adult authority.
(d) displaced emotion resulting from some incident earlier in the day.

35. By the definition in the text, attitudes differ from other meanings in this way:
(a) They are less clearly expressed.
(b) They are formed more indirectly.
(c) They involve acceptance or rejection.
(d) They apply to general classes of objects or events.

36. Practice material for typing courses will prove best, in terms of ability to type later on whatever job the student gets, if
(a) the material resembles closely the sort of copy used on most jobs.
(b) the material includes a complete coverage of varied words and sentences (for example, pharmaceutical terms like "polymyin") to include all letter combinations even though some are uncommon on jobs.
(c) the material is relatively easy, so that there is a high degree of over-learning of speed for the copy practiced.

37. To avoid the faults of tests as an evaluation procedure, tests should be
(a) used chiefly for the less important objectives.
(b) used regularly throughout the course.
(c) reserved for final evaluation at the end of the course.
(d) given little weight in marking.

38. When a teacher invents a new teaching procedure, psychology's chief contribution is to provide
(a) a scientific method for determining its value.
(b) principles that the proposal should conform to.
(c) superior procedures which made the invention unnecessary.
(d) ways of studying individuals when the procedure fails.

39. A person adopts the standards of the group that
(a) he is with at the time.
(b) accepts him in a friendly fashion.
(c) he believes he is like.

40. A teacher's case study of a pupil begins with
(a) administration of a general series of readiness tests.
(b) observation in varied situations.
(c) a detailed interview.

41. Setting up an individualized program for each pupil
(a) is especially useful in developing definite, easily checked skills.
(b) leads to less thorough coverage than a program of uniform assignments.
(c) is practicable only where pupils have quite similar readiness.
(d) provides for social development as well as intellectual growth.

42. Inattention ordinarily means
(a) lack of readiness.
(b) lack of interest.
(c) poor study skills.
(d) low mental ability.

43. Motion pictures are especially useful in teaching about
(a) events involving motion of change.
(b) experiences familiar to the pupil, but in which he has been emotionally involved.
(c) events which the pupil has not experienced.

44. When pupils are graded on the basis of their ability,
(a) good grades become an attainable goals.
(b) students no longer see grades as rewarding.
(c) the better students are doomed to frustration.

45. Socialization is concerned with developing
(a) social-emotional adjustment.
(b) attitudes, likes and dislikes, and beliefs.
(c) skills needed for keeping alive and well.
(d) patterns of action expected by one's group.

46. Governments are often classified into such groups as democracy, dictatorship, monarchy, etc. In advanced subjects, thinking is improved by
(a) stating generalizations in terms of these concepts.
(b) reclassifying on some other basis before generalizing.
(c) subdividing the group before stating generalizations.

47. The advisable relation between early practice and explanation is this:
(a) Explanation should be introduced only after the learner has passed the random trial-and-error stage.
(b) Practice should be allowed before explanation only if it is needed to make the explanation intelligible.
(c) Practice and explanation should be alternated, in roughly equal proportion.
(d) Explanation should be offered only when the learner realizes that his form is poor.

48. Among five superior fourth-graders who can do arithmetic at the sixth-grade level, we would expect that
(a) four of the five can read at the sixth-grade level.
(b) the average reading ability will be at the fifth-grade level.
(c) four of the five will have fourth-grade reading ability.

49. Evaluation of school learning refers to
(a) preparation and administration of objective tests.
(b) assignment of marks on the basis of performance.
(c) judging the adequacy of responses.
(d) making provision for individual differences.

50. Which of these traits of teachers is most liked by pupils?
(a) Maintains an objective, impersonal attitude.
(d) Has high intellectual ability.
(c) Is sympathetic.
(d) Sets standards which are easy to attain.

Scoring Key

Pre And Post Achievement Test (PAT)

Item	Correct Response	Item	Correct Response
1	b	26	a
2	c	27	c
3	c	28	c
4	c	29	b
5	a	30	c
6	c	31	b
7	d	32	c
8	b	33	a
9	a	34	b
10	b	35	c
11	c	36	a
12	d	37	b
13	b	38	b
14	c	39	c
15	c	40	b
16	a	41	a
17	b	42	a
18	c	43	c
19	b	44	c
20	d	45	d
21	c	46	b
22	b	47	b
23	b	48	b
24	b	49	c
25	b	50	c

M E S - F.E. 200

Please read each statement carefully and then indicate whether you believe it is TRUE or FALSE by marking the appropriate space on the answer sheet.

1. The number of man's senses is five.

2. Girls are not more moralistic and better behaved by nature than boys.

3. Long slender hands indicate an artistic nature.

4. Most stealing offenses by Junior High students are not serious.

5. The limits for the development of any student are set more by his desire and persistance than by his heredity.

6. There are characteristic racial differences in intelligence.

7. Outstandingly beautiful girls are usually not as intelligent as girls who are not so beautiful.

8. The basic needs of adolescents are not different from the needs of people of other ages.

9. Smoking among young people is a sign of maladjustment.

10. Students who are impertinent and defiant are usually exhibiting signs of personality problems.

11. The speed of forgetting is slowest immediately after material (such as a poem) has been learned.

12. Social acceptance and the need for recognition are not necessarily motivational forces.

13. Fearfulness is not an indication of underlying behavior problems in children but is common among young children.

14. Subject-matter must have personal meaning to become interesting to the learner.

15. Thoughtlessness in students should be looked for by the teacher as a sign of serious emotional problems.

16. The student who constantly interrupts class proceedings is showing signs of maladjustment.

17. One cannot determine what his life vocation should be by the use of tests.

18. Pupils should be given heavy study loads so as to hasten maturation.

19. Many eminent men were feeble-minded when they were young.

20. You can estimate a person's intelligence pretty closely by his facial characteristics.

21. Pupils who are consistently restless are usually displaying symptoms of an emotional problem.

22. Women (as a group) are inferior to men (as a group) in intelligence.

23. Grades are important for forecasting a student's vocational success.

24. Carelessness in work is usually an indication of an underlying emotional problem.

25. Unsocialness is not a serious problem but is usually characteristic of the "brainy" or intellectual student.

26. Students who are usually disorderly are in need of help for some emotional problem.

27. Intelligence test scores cannot be increased by training.

28. Practice makes perfect.

29. Cruelty and bullying are not serious signs of maladjustment.

30. It is usually helpful in teaching to classify people into types or categories.

31. A person who has a square jaw will usually have a great deal of will power.

32. Children who are large for their age are usually not as intelligent as children who more nearly approach the size norms for their age.

33. Dishonesty is indicated when a person does not look you in the eye.

34. A high forehead usually indicates intellectual superiority.

35. If you have enough will power you can compensate for most defects of body or mind.

36. The teacher need not concern himself with mental health as there are specialists to do this.

37. Students who are easily discouraged need extra help and confidence but their problems are not usually serious enough to be referred to a counselor or psychologist.

38. The use of profanity by young people does not constitute a behavior problem.

39. Since the major concern of teachers is intellectual growth, the growth of the personality is of only minor importance to teachers.

40. One need not understand the culture in order to understand the behavior of young people.

41. Over-study and over-work can cause feeblemindedness.

42. Adolescence is a biological phenomenon and has no relation to the culture.

43. In most cases, disobedience is not a sign of serious maladjustment.

44. Continued unhappiness is usually not a sign of a serious emotional problem but indicates a lack of interest or the need for friends.

45. Children who are weak and retarded physically are usually especially intelligent.

46. In general, gifted children require less direction and guidance than average or dull children.

47. Social values of teachers are usually the same as the social values of the upper classes.

48. Grades in school subjects are a reliable and valid means for measuring a student's total growth.

49. Working class people are eager to learn and accept "middle-class" values.

50. The study of mathematics is valuable because it helps develop logical thinking.

51. Usually, adolescents need help with problems related to sex maturation.

52. Adolescents are basically negativistic.

53. There are innate differences in intelligence between different races.

54. All men are not created equal in the capacity for accomplishment.

55. Failing a student is a very good way to motivate him to improve his work.

56. Teachers usually have risen into the middle classes from the upper-lower classes.

57. Tardiness is usually an indication of an emotional problem.

58. I.Q. and intelligence are the same thing.

59. Man is an intellectual being and his emotions play an insignificant part in his behavior.

60. Being overly-sensitive is typical behavior of adolescents.

61. An expectant mother can influence the character of her unborn child.

Scoring Key

Misconceptions About Education Scale (MES)

Item	Correct Response	Item	Correct Response
1	F	31	F
2	T	32	F
3	F	33	F
4	T	34	F
5	F	35	F
6	F	36	F
7	F	37	F
8	T	38	F
9	F	39	F
10	F	40	F
11	F	41	F
12	F	42	F
13	F	43	T
14	T	44	F
15	F	45	F
16	F	46	F
17	T	47	F
18	F	48	F
19	F	49	F
20	F	50	F
21	F	51	T
22	F	52	F
23	T	53	F
24	F	54	T
25	F	55	F
26	F	56	T
27	F	57	F
28	F	58	F
29	F	59	F
30	F	60	F
		61	F

Room No._____ Seat No._____

This scale is designed to help your instructors in evaluating the effectiveness of instruction in this course. Your responses will have no bearing on your grade. In order to be of maximum value, you are asked to be sincere, honest and objective in your answers.

Please choose the alternative which best completes each statement for you and mark the appropriate space on the answer sheet.

1. As far as the information in this course is concerned,
   a. I think it will be very useful in my future teaching.
   b. I think it will be somewhat useful in my future teaching.
   c. I am not sure whether or not it will be useful in my future teaching.
   d. I think it will be of little use in my future teaching.
   e. I think it will be of no use in my future teaching.

2. In the future whenever there is an opportunity to find out more about the subject matter in this course,
   a. I will positively do so.
   b. I will probably do so.
   c. I don't know what I will do.
   d. I will probably not do so.
   e. I will positively not do so.

3. The information in this course is
   a. very interesting to me.
   b. somewhat interesting to me.
   c. neither interesting nor uninteresting to me.
   d. somewhat uninteresting to me.
   e. very uninteresting to me.

4. Now that I have had some experience in this course,
   a. I am very glad I took the course.
   b. I am glad I took the course.
   c. I am neither glad nor sorry I took the course.
   d. I am sorry that I took the course.
   e. I am very sorry I took the course.

5. If I have the opportunity to take an advanced course similar to this one,
   a. I will certainly do so.
   b. I will probably do so.
   c. I don't know what I will do.
   d. I will probably not do so.
   e. I will certainly not do so.

6. The subject-matter in this course is
   a. very practical.
   b. somewhat practical.
   c. neither practical nor impractical.
   d. somewhat impractical.
   e. very impractical.

7. In comparing the lecture and discussion portions of this course, I feel that lec
   a. is much more informative than the discussion.
   b. is somewhat more informative than the discussion.
   c. and discussion are equally informative.
   d. is somewhat less informative than the discussion.
   e. is much less informative than the discussion.

8. Do you think that you have learned as much over television as you would have learned in a similar course, but taught without the use of television?
   a. Much more.
   b. A little more.
   c. About the same.
   d. A little less.
   e. Much less.

9. Do you think this course is more or less valuable than it would have been if television had not been used?
   a. Much more valuable.
   b. Somewhat more valuable.
   c. About the same.
   d. Somewhat less valuable.
   e. Much less valuable.

10. As far as class participation is concerned, I would have liked to participate
   a. much more.
   b. a little more.
   c. no more than at present.
   d. a little less.
   e. much less.

11. In comparing the panel discussion groups and 6 x 6 discussion groups, I feel that the 6 x 6 discussion groups
   a. were much more helpful than the panel discussion groups.
   b. were somewhat more helpful than the panel discussion groups.
   c. were neither more nor less helpful than the panel discussion groups.
   d. were somewhat less helpful than the panel discussion groups.
   e. were much less helpful than the panel discussion groups.

12. As far as the time devoted to discussion in the course, I would have liked to have spent
   a. much more time on discussion.
   b. a little more time on discussion.
   c. no more nor no less time in discussion.
   d. a little less time in discussion.
   e. much less time in discussion.

13. Aside from the technical difficulties involved in presenting movies over television, I feel that the movies
   a. were very useful.
   b. were somewhat useful.
   c. were neither useful nor useless.
   d. were somewhat useless.
   e. were very useless.

14. As far as the reading and reporting of journal articles once weekly is concerned
   a. I think it was very helpful.
   b. I think it was somewhat helpful.
   c. I am not sure if it was helpful.
   d. I think it was of little help.
   e. I think it was of no help.

15. As far as the participation in panels was concerned, I would have liked
   a. to participate much more in panel discussion.
   b. to participate somewhat more in panel discussion
   c. to participate neither more nor less in panel discussion.
   d. to participate somewhat less in panel discussion.
   e. to participate much less in panel discussion.

16. As far as the subject-matter of the lecture in comparison to textbook material was concerned,
    a. they were too much alike.
    b. they were somewhat alike.
    c. they were about right.
    d. they were somewhat unlike.
    e. they were too much unlike.

17. As far as television instruction is concerned, I think it is
    a. much less disturbing than ordinary courses.
    b. somewhat less disturbing than ordinary courses.
    c. no more disturbing than ordinary courses.
    d. somewhat more disturbing than ordinary courses.
    e. much more disturbing than ordinary courses.

18. As far as television instruction is concerned, I think that I can concentrate
    a. much better than in ordinary classes.
    b. somewhat better than in ordinary classes.
    c. no more nor less than in ordinary classes.
    d. somewhat poorer than in ordinary classes.
    e. much poorer than in ordinary classes.

19. As far as the instruction is concerned, I feel that it is
    a. well above average.
    b. above average.
    c. about average.
    d. below average.
    e. well below average.

20. If I had it to do again, knowing what I now know, I would
    a. positively choose a large television section again.
    b. probably choose a large television section again.
    c. have no preference.
    d. probably would not choose a large television section again.
    e. positively would not choose a large television section again.

21. I feel that the objectives of the course have been
    a. very well met.
    b. well met.
    c. satisfactorily met.
    d. poorly met.
    e. very poorly met.

22. As far as the instructors in this course are concerned, I think they are
    a. among the best.
    b. better than most.
    c. about average.
    d. poorer than most.
    e. among the poorest.

23. Use the space on the back of this sheet to make any comments or suggestions that you feel would be helpful to the instructors in improving the course.

Item Weights

Course Evaluation Scale (CES)

	a	b	c	d	e
1.	5	4	3	2	1
2.	5	4	3	2	1
3.	5	4	3	2	1
4.	5	4	3	2	1
5.	5	4	3	2	1
6.	5	4	3	2	1
7.	1	2	3	4	5
8.	5	4	3	2	1
9.	5	4	3	2	1
10.	1	2	3	4	5
11.	1	2	3	4	5
12.	1	2	3	4	5
13.	5	4	3	2	1
14.	5	4	3	2	1
15.	1	2	3	4	5
16.	5	4	3	2	1
17.	5	4	3	2	1
18.	5	4	3	2	1
19.	5	4	3	2	1
20.	5	4	3	2	1
21.	5	4	3	2	1
22.	5	4	3	2	1

APPENDIX B

ORIGINAL DATA

APPENDIX B

ORIGINAL DATA

Code Detail

Group:
TV - Television
NTV - Non-television

Subgroup:
The numbers refer to the 16 cells of the experimental design as shown on Page

Sex:
M - Male
F - Female

Class:
1 - Freshman
2 - Sophomore
3 - Junior
4 - Senior
5 - Special

Curriculum:
-- - No major
10 - Agriculture
20 - Business
22 - Bus. Educ. & Sec. Studies
25 - Political Science
26 - Social Work

Curriculum (cont'd)
30 - Engineering
40 - Home Economics
50 - Fine Arts
52 - Language & Literature
55 - Biological Science
56 - Physical Science
57 - Social Science
59 - Pre-Medical
71 - Education
73 - Physical Education
96 - Speech

Measuring Instruments:
ACE - American Council on Education Psychological Examination for College Freshmen
PICS - Preferred Instructor Characteristics Scale
PRE PAT - Pre-Achievement Test
POST PAT - Post-Achievement Test
PRE MTAI - Pre-Minnesota Teacher Attitude Inventory
POST MTAI - Post-Minnesota Teacher Attitude Inventory
PRE MES - Pre-Misconceptions about Education Scale
POST MES - Post-Misconceptions about Education Scale
CES - Course Evaluation Scale

APPENDIX B - ORIGINAL DATA (cont'd.)

STUDENT	GROUP	SUB-GROUP	SEX	CLASS	CURRI-CULUM	ACE	PICS	PRE PAT	POST PAT	PRE MTAI*	POST MTAI*	PRE MES	POST MES	CES
1	TV	1	M	3	57	109	19	21	23	063	081	46	50	69
2	TV	1	M	3	57	144	25	26	29	085	128	41	47	49
3	TV	1	M	2	55	109	20	21	22	077	098	47	45	49
4	NTV	2	M	4	55	150	09	18	25	019	045	39	41	56
5	NTV	2	M	3	71	105	21	18	22	061	125	41	40	57
6	NTV	2	M	2	50	105	10	20	27	013	096	43	43	55
7	TV	3	F	2	—	120	26	17	22	043	095	35	42	62
8	TV	3	F	2	40	114	12	20	29	131	138	42	49	57
9	TV	3	F	2	52	123	23	20	34	088	123	51	47	56
10	TV	3	F	2	57	108	21	14	31	039	069	39	48	57
11	TV	3	F	2	57	129	23	22	29	099	114	44	43	44
12	TV	3	F	2	71	117	11	16	30	091	108	44	42	63
13	TV	3	F	2	71	112	12	17	23	104	114	38	46	39
14	TV	3	F	2	71	120	22	16	23	066	052	38	50	57
15	TV	3	F	3	22	121	10	21	35	037	067	39	41	51
16	TV	3	F	3	52	130	11	21	33	033	102	36	48	65
17	TV	3	F	4	22	111	10	19	31	114	141	38	45	64
18	NTV	4	F	3	71	106	18	15	24	155	146	41	43	65
19	NTV	4	F	3	71	105	09	21	22	043	109	42	41	65
20	NTV	4	F	3	57	120	19	22	39	070	127	43	52	62
21	NTV	4	F	2	10	116	27	16	26	087	136	37	44	66
22	NTV	4	F	2	22	114	17	22	35	060	115	39	43	47
23	NTV	4	F	2	40	116	12	18	27	079	106	38	36	68
24	NTV	4	F	2	52	119	29	17	29	094	106	42	51	50
25	NTV	4	F	2	57	130	11	18	22	065	090	37	33	51
26	NTV	4	F	2	71	120	19	17	25	071	104	41	44	53
27	NTV	4	F	1	57	131	14	21	25	072	120	39	40	55
28	NTV	4	F	1	52	120	21	20	28	082	112	47	48	69
29	TV	5	M	4	57	123	04	21	32	139	141	52	53	61
30	TV	5	M	3	57	119	00	15	24	032	032	41	44	46
31	TV	5	M	2	57	117	01	30	33	046	091	42	49	38

*A constant of 50 has been added to MTAI scores.

APPENDIX B - ORIGINAL DATA (cont'd.)*

STUDENT	GROUP	SUB-GROUP	SEX	CLASS	CURRI-CULUM	ACE	PICS	PRE PAT	POST PAT	PRE MTAI**	POST MTAI**	PRE MES	POST MES	CES
32	NTV	6	M	3	56	129	05	17	27	120	091	39	43	63
33	NTV	6	M	2	—	122	00	19	19	071	090	33	41	59
34	NTV	6	M	3	71	127	07	22	26	106	120	47	49	69
35	TV	7	F	2	—	116	00	18	28	055	084	32	43	65
36	TV	7	F	2	—	122	00	16	25	059	118	38	46	43
37	TV	7	F	2	—	117	05	25	25	060	107	39	33	64
38	TV	7	F	2	22	113	03	21	27	025	102	38	46	55
39	TV	7	F	2	71	110	01	14	21	035	087	38	35	48
40	TV	7	F	2	71	122	03	18	23	064	093	33	31	53
41	TV	7	F	2	71	112	06	20	27	079	076	37	40	63
42	TV	7	F	2	71	110	07	20	25	058	093	35	44	65
43	TV	7	F	3	22	103	07	17	22	055	079	34	38	45
44	TV	7	F	3	40	126	03	19	32	088	145	43	47	70
45	TV	7	F	3	40	115	06	21	25	059	085	40	41	56
46	NTV	8	F	1	71	114	03	17	28	055	085	31	45	73
47	NTV	8	F	2	40	114	01	22	27	073	100	40	38	48
48	NTV	8	F	2	50	112	07	22	21	114	108	46	42	60
49	NTV	8	F	2	71	121	02	19	27	078	094	45	47	65
50	NTV	8	F	2	71	145	00	19	23	144	120	42	47	62
51	NTV	8	F	2	71	139	08	17	28	031	059	37	37	65
52	NTV	8	F	3	56	118	08	19	23	044	119	37	47	70
53	NTV	8	F	3	52	130	05	22	29	107	129	42	44	57
54	NTV	8	F	3	52	108	00	25	23	048	090	43	48	50
55	NTV	8	F	3	22	109	01	28	37	151	148	41	46	66
56	NTV	8	F	2	—	141	08	15	26	142	142	47	51	52
57	TV	9	M	4	55	098	15	15	27	078	059	42	44	61
58	TV	9	M	2	73	099	23	15	26	095	110	46	49	57
59	TV	9	M	2	73	103	12	13	35	073	085	35	35	44
60	NTV	10	M	3	57	097	32	16	39	126	143	48	36	63
61	NTV	10	M	3	22	094	18	14	18	065	073	33	37	48

*Code detail will be found on page

**A constant of 50 has been added to MTAI scores.

APPENDIX B - ORIGINAL DATA (cont'd.)*

STUDENT	GROUP	SUB-GROUP	SEX	CLASS	CURRI-CULUM	ACE	PICS	PRE PAT	POST PAT	PRE MTAI**	POST MTAI**	PRE MES	POST MES	CES
62	NTV	10	M	2	73	097	28	21	23	033	107	39	41	66
63	TV	11	F	3	52	099	17	21	29	090	120	40	47	46
64	TV	11	F	3	71	096	15	21	26	087	108	33	40	60
65	TV	11	F	4	50	096	15	14	27	018	050	30	31	35
66	TV	11	F	2	73	104	13	16	26	080	074	38	46	61
67	TV	11	F	2	71	101	23	26	25	087	126	38	48	53
68	TV	11	F	2	71	104	20	22	28	047	095	34	40	56
69	TV	11	F	2	71	085	18	22	30	086	092	39	40	21
70	TV	11	F	2	71	093	13	19	30	099	112	44	48	66
71	TV	11	F	2	71	089	13	27	27	090	133	37	47	48
72	TV	11	F	2	71	104	10	17	23	094	095	39	46	65
73	TV	11	F	5	—	096	12	22	31	064	117	39	45	65
74	NTV	12	F	3	71	089	09	19	27	060	074	35	43	58
75	NTV	12	F	3	22	083	10	18	32	113	129	47	48	61
76	NTV	12	F	2	71	101	10	18	35	087	101	42	49	64
77	NTV	12	F	2	71	100	18	18	21	068	096	35	43	53
78	NTV	12	F	2	71	098	16	16	38	090	113	34	40	68
79	NTV	12	F	2	71	097	18	17	31	102	115	33	45	60
80	NTV	12	F	2	71	091	20	19	31	047	129	41	39	64
81	NTV	12	F	2	71	090	23	13	20	022	089	31	39	58
82	NTV	12	F	4	71	080	14	18	27	082	116	42	42	73
83	NTV	12	F	4	50	065	14	14	31	058	104	40	52	72
84	NTV	12	F	3	71	098	35	17	26	093	117	46	44	62
85	TV	13	M	2	—	097	07	26	25	071	071	49	43	51
86	TV	13	M	3	25	103	00	18	27	061	116	44	45	60
87	TV	13	M	3	59	094	00	17	26	031	097	39	42	48
88	NTV	14	M	2	73	071	02	22	19	017	044	37	44	71
89	NTV	14	M	2	73	060	03	17	22	024	033	35	40	49
90	NTV	14	M	2	—	074	01	15	27	040	054	42	46	54
91	TV	15	F	1	—	086	00	18	27	0333	067	41	44	44

*Code detail will be found on page

**A constant of 50 has been added to MTAI scores.

APPENDIX B - ORIGINAL DATA (cont'd.)*

STUDENT	GROUP	SUB-GROUP	SEX	CLASS	CURRI-CULUM	ACE	PICS	PRE PAT	POST PAT	PRE MTAI**	POST MTAI**	PRE MES	POST MES	CES
92	TV	15	F	1	71	101	02	20	21	060	059	32	45	54
93	TV	15	F	3	22	090	07	17	30	036	121	31	34	58
94	TV	15	F	3	22	096	07	17	24	080	112	42	46	51
95	TV	15	F	1	57	097	00	14	32	068	068	43	43	48
96	TV	15	F	3	71	085	06	19	30	125	135	45	47	62
97	TV	15	F	2	71	084	03	20	27	087	113	38	41	67
98	TV	15	F	2	71	097	02	22	34	068	105	32	44	59
99	TV	15	F	2	71	089	00	20	19	030	058	34	36	39
100	TV	15	F	2	40	090	08	14	21	075	074	35	42	47
101	TV	15	F	2	—	082	07	18	26	041	096	39	45	59
102	NTV	16	F	2	71	088	05	17	34	061	095	34	49	56
103	NTV	16	F	2	71	094	01	23	32	059	077	36	39	49
104	NTV	16	F	2	71	076	06	20	20	088	083	43	36	54
105	NTV	16	F	2	71	076	05	17	23	094	122	35	44	58
106	NTV	16	F	2	57	104	00	20	24	020	042	38	35	71
107	NTV	16	F	2	—	099	00	16	31	095	108	42	42	39
108	NTV	16	F	1	71	086	03	18	23	071	089	32	44	61
109	NTV	16	F	1	—	073	06	14	26	017	035	16	29	75
110	NTV	16	F	3	50	039	00	19	29	084	127	39	41	58
111	NTV	16	F	3	40	075	03	19	26	052	115	34	43	43
112	NTV	16	F	3	22	104	01	21	25	080	115	41	45	67
113	TV		M	1	71	115	11	23	32	046	102	34	44	63
114	TV		M	2	10	109	24	14	27	037	068	29	34	68
115	TV		M	2	10	104	22	21	28	032	073	38	39	50
116	TV		M	2	50	134	14	14	23	057	084	33	36	49
117	TV		M	2	57	149	09	25	35	073	100	51	52	41
118	TV		M	2	73	112	14	21	26	080	131	38	48	53
119	TV		M	3	57	128	11	13	27	036	119	39	47	65
120	TV		M	3	57	106	11	16	19	092	124	46	44	56
121	TV		M	4	30	117	35	13	31	093	111	39	43	61

*Code detail will be found on page

**A constant of 50 has been added to MTAI scores.

APPENDIX B - ORIGINAL DATA (cont'd.)*

STUDENT	GROUP	SUB-GROUP	SEX	CLASS	CURRI-CULUM	ACE	PICS	PRE PAT	POST PAT	PRE MTAI**	POST MTAI**	PRE MES	POST MES	CES
122	TV		F	1	57	134	18	21	27	121	158	41	49	41
123	TV		F	1	57	136	06	25	32	086	110	43	49	49
124	TV		F	1	73	129	19	20	22	098	120	38	40	37
125	TV		F	2	26	099	03	22	30	087	073	43	45	53
126	TV		F	2	40	131	11	22	34	035	072	41	48	54
127	TV		F	2	40	143	03	21	24	097	084	35	40	38
128	TV		F	2	57	090	07	15	28	054	106	34	41	61
129	TV		F	2	71	136	10	15	25	074	090	41	39	44
130	TV		F	2	71	107	34	18	30	046	043	41	45	72
131	TV		F	2	71	114	17	18	28	066	106	38	48	49
132	TV		F	2	71	117	13	19	24	110	145	41	47	64
133	TV		F	2	71	144	06	25	26	120	132	39	47	54
134	TV		F	2	71	121	00	17	29	042	085	45	48	63
135	TV		F	2	71	101	00	14	22	134	124	37	43	59
136	TV		F	3	71	092	05	20	26	090	141	38	41	66
137	TV		F	2	96	105	10	19	26	093	089	42	39	52
138	TV		F	3	22	105	00	15	28	062	109	32	44	52
139	TV		F	3	56	140	08	21	20	038	077	42	35	49
140	TV		F	3	57	116	17	26	36	064	089	46	46	58
141	TV		F	3	71	125	15	16	28	109	131	42	46	55
142	TV		F	4	57	148	19	18	30	068	109	45	49	52
143	NTV		M	-	—	066	09	16	20	113	112	44	49	58
144	NTV		M	2	—	111	00	23	24	051	114	34	41	60
145	NTV		M	2	10	115	16	14	17	006	047	30	42	61
146	NTV		M	2	10	119	07	18	23	044	073	26	38	68
147	NTV		M	2	56	097	22	14	26	067	077	42	53	56
148	NTV		M	2	57	133	18	21	28	022	102	41	41	65
149	NTV		M	2	57	096	13	17	21	097	089	44	44	57
150	NTV		M	2	71	086	10	18	33	087	092	35	40	60
151	NTV		M	3	—	117	00	22	26	083	106	47	48	70

*Code Detail will be found on page

**A constant of 50 has been added to MTAI scores.

APPENDIX B - ORIGINAL DATA (cont'd.)*

STUDENT	GROUP	SUB-GROUP	SEX	CLASS	CURRI-CULUM	ACE	PICS	PRE PAT	POST PAT	PRE MTAI	POST MTAI	PRE MES	POST MES	CES
152	NTV		M	3	50	086	10	12	30	163	121	37	42	70
153	NTV		M	3	55	093	15	15	26	060	082	43	47	55
154	NTV		M	3	96	148	06	22	21	025	029	38	29	49
155	NTV		M	4	20	103	02	17	24	118	127	40	48	57
156	NTV		F	2	—	109	03	22	32	092	113	40	44	71
157	NTV		F	2	40	120	10	21	17	047	129	30	36	72
158	NTV		F	2	57	041	20	12	19	073	119	37	39	57
159	NTV		F	2	71	141	23	18	29	080	135	36	55	65
160	NTV		F	2	71	110	05	22	22	044	070	40	42	72
161	NTV		F	2	71	098	11	14	22	036	070	30	37	44
162	NTV		F	2	71	097	01	19	27	048	097	41	40	55
163	NTV		F	3	22	115	06	18	33	065	127	40	44	52
164	NTV		F	3	40	122	06	23	24	095	091	42	44	65
165	NTV		F	3	71	089	16	15	22	066	106	43	44	61
166	NTV		F	3	71	098	07	18	20	104	137	48	43	67
167	NTV		F	5	25	174	00	27	39	118	149	48	50	42

*Code detail will be found on page

**A constant of 50 has been added to MTAI scores.

APPENDIX C

EXPLANATION OF COMPUTATIONAL PROCEDURES

APPENDIX C

EXPLANATION OF COMPUTATIONAL PROCEDURES

The following series of tables are included to show the manner in which the data from the criterion measures for the sixteen subgroups was arranged to test the null hypotheses by the analysis of variance and covariance. A table is shown for each of the main and interaction effects, in which appears the identifying number of the subgroups in each cell of the table, the number of students in each subgroup, and the total numbers. The original data was processed as follows:

1. The sums of scores, sums of squares and sums of cross-products were computed for each individual.

2. The results for individuals from step one were combined by subgroups and subgroup totals computed.

3. The subgroup summations were placed in their appropriate cells in the tables shown below, and the summation of scores, sums of squares, and sums of cross-products was completed for each cell of the table and for the total of the table.

4. The summarized data for each table--the summation for each of the main and interaction effects--was tabulated and the adjusted post-test sums of squares, variances and F ratios were computed. The results of the final computations are presented in Chapter VI, "Analysis of the Results."

By referring to the tables in this Appendix and using the origi-

nal data from Appendix B, the sums of squares, adjusted sums of squares, variances and F ratios can be computed.

Main Effects

TABLE C-1

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS OF ABILITY LEVEL (ACE) EFFECT

High ACE		Low ACE		
Subgroup Number	N	Subgroup Number	N	Total
1	3	9	3	
2	3	10	3	
3	11	11	11	
4	11	12	11	
5	3	13	3	
6	3	14	3	
7	11	15	11	
8	11	16	11	
Total	56		56	112

TABLE C-2

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS PREFERENCE FOR INSTRUCTION (PICS).

High PICS		Low PICS		
Subgroup Number	N	Subgroup Number	N	Total
1	3	5	3	
2	3	6	3	
3	11	7	11	
4	11	8	11	
9	3	13	3	
10	3	14	3	
11	11	15	11	
12	11	16	11	
Total	56		56	112

TABLE C-3

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS OF SEX EFFECT

Male		Female		
Subgroup Number	N	Subgroup Number	N	Total
1	3	3	11	
2	3	4	11	
5	3	7	11	
6	3	8	11	
9	3	11	11	
10	3	12	11	
13	3	15	11	
14	3	16	11	
Total	24		88	112

TABLE C-4

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS OF DISCUSSION TECHNIQUES (TECH)

Television		Non-Television		
Subgroup Number	N	Subgroup Number	N	Total
1	3	2	3	
3	11	4	11	
5	3	6	3	
7	11	8	11	
9	3	10	3	
11	11	12	11	
13	3	14	3	
15	11	16	11	
Total	56		56	112

Interaction Effects

TABLE C-5

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS OF ACE x PICS INTERACTION

	High ACE		Low ACE		
	Subgroup Number	N	Subgroup Number	N	Total
High PICS	1	3	9	3	
	2	3	10	3	
	3	11	11	11	
	4	11	12	11	
	Total	28	Total	28	56
Low PICS	5	3	13	3	
	6	3	14	3	
	7	11	15	11	
	8	11	16	11	
	Total	28	Total	28	56
Total		56		56	112

TABLE C-6

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS OF ACE x SEX INTERACTION

	High ACE		Low ACE		
	Subgroup Number	N	Subgroup Number	N	Total
Male	1	3	9	3	
	2	3	10	3	
	5	3	13	3	
	6	3	14	3	
	Total	12	Total	12	24
Female	3	11	11	11	
	4	11	12	11	
	7	11	15	11	
	8	11	16	11	
	Total	44	Total	44	88
Total		56		56	112

TABLE C-7

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS
OF ACE x TECH INTERACTION

	High ACE		Low ACE		
	Subgroup Number	N	Subgroup Number	N	Total
TV	1	3	9	3	
	3	11	11	11	
	5	3	13	3	
	7	11	15	11	
	Total	28	Total	28	56
NTV	2	3	10	3	
	4	11	12	11	
	6	3	14	3	
	8	11	16	11	
	Total	28	Total	28	56
Total		56		56	112

TABLE C-8

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS
PICS x SEX INTERACTION

	High PICS		Low PICS		
	Subgroup Number	N	Subgroup Number	N	Total
Male	1	3	5	3	
	2	3	6	3	
	9	3	13	3	
	10	3	14	3	
	Total	12	Total	12	24
Female	3	11	7	11	
	4	11	8	11	
	11	11	15	11	
	12	11	16	11	
	Total	44	Total	44	88
Total		56		56	112

TABLE C-9

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS
OF PICS x TECH INTERACTION

	High PICS		Low PICS		
	Subgroup Number	N	Subgroup Number	N	Total
	1	3	5	3	
	3	11	7	11	
TV	9	3	13	3	
	11	11	15	11	
	Total	28	Total	28	56
	2	3	6	3	
	4	11	8	11	
NTV	10	3	14	3	
	12	11	16	11	
	Total	28	Total	28	56
Total		56		56	112

TABLE C-10

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS
OF SEX x TECH INTERACTION

	Male		Female		
	Subgroup Number	N	Subgroup Number	N	Total
	1	3	3	11	
	5	3	7	11	
TV	9	3	11	11	
	13	3	15	11	
	Total	12	Total	44	56
	2	3	4	11	
	6	3	8	11	
NTV	10	3	12	11	
	14	3	16	11	
	Total	12	Total	44	56
Total		24		88	112

TABLE C-11

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS
OF ACE x PICS x SEX INTERACTION

	High ACE				Low ACE				
	High PICS		Low PICS		High PICS		Low PICS		
	Subgroup Number	N	Subgroup Number	N	Subgroup Number	N	Subgroup Number	N	Total
	1	3	5	3	9	3	13	3	
Male	2	3	6	3	10	3	14	3	
	Total	6	Total	6	Total	6	Total	6	24
	3	11	7	11	11	11	15	11	
Female	4	11	8	11	12	11	16	11	
	Total	22	Total	22	Total	22	Total	22	88
Total		28		28		28		28	112

TABLE C-12

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS
OF ACE x PICS x TECH INTERACTION

	High ACE				Low ACE				
	High PICS		Low PICS		High PICS		Low PICS		
	Subgroup Number	N	Subgroup Number	N	Subgroup Number	N	Subgroup Number	N	Total
	1	3	5	3	9	3	13	3	
TV	3	11	7	11	11	11	15	11	
	Total	14	Total	14	Total	14	Total	14	56
	2	3	6	3	10	3	14	3	
NTV	4	11	8	11	12	11	16	11	
	Total	14	Total	14	Total	14	Total	14	56
Total		28		28		28		28	112

TABLE C-13

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS
OF ACE x SEX x TECH INTERACTION

	High ACE				Low ACE				
	Male		Female		Male		Female		
	Subgroup Number	N	Subgroup Number	N	Subgroup Number	N	Subgroup Number	N	Total
TV	1	3	3	11	9	3	11	11	
	5	3	7	11	13	3	15	11	
	Total	6	Total	22	Total	6	Total	22	56
NTV	2	3	4	11	10	3	12	11	
	6	3	8	11	14	3	16	11	
	Total	6	Total	22	Total	6	Total	22	56
Total		12		44		12		44	112

TABLE C-14

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS
OF PICS x SEX x TECH INTERACTION

	High PICS				Low PICS				
	Male		Female		Male		Female		
	Subgroup Number	N	Subgroup Number	N	Subgroup Number	N	Subgroup Number	N	Total
TV	1	3	3	11	5	3	7	11	
	9	3	11	11	13	3	15	11	
	Total	6	Total	22	Total	6	Total	22	56
NTV	2	3	4	11	6	3	8	11	
	10	3	12	11	14	3	16	11	
	Total	6	Total	22	Total	6	Total	22	56
Total		12		44		12		44	112

TABLE C-15

ARRANGEMENT OF SUBGROUPS FOR ANALYSIS
OF ACE x PICS x SEX x TECH INTERACTION

		High ACE				Low ACE				
		High PICS		Low PICS		High PICS		Low PICS		
		Subgroup Number	N	Subgroup Number	N	Subgroup Number	N	Subgroup Number	N	Total
TV	Male	1	3	5	3	9	3	13	3	12
	Female	3	11	7	11	11	11	15	11	44
	Total		14		14		14		14	56
NTV	Male	2	3	6	3	10	3	14	3	12
	Female	4	11	8	11	12	11	16	11	44
	Total		14		14		14		14	56
Total			28		28		28		28	112

APPENDIX D

COMPUTATIONS TO TEST ASSUMPTIONS OF ANALYSIS OF VARIANCE AND COVARIANCE

TABLE D-1

COMPUTATION TO TEST HOMOGENEITY OF REGRESSION AND LINEARITY OF REGRESSION OF SUBGROUPS ON PAT

Nature of Variation	Symbol	df	Sums of Squares	Variance
Subgroup regression coefficients about common coefficient	S_1	k-1 15	168.24	11.22
Scores about regression line for their own subgroup	S_2	N-2k 80	1604.18	20.05
Subgroup post-test means about regression line based on means	S_3	k-2 14	269.14	19.22
Difference between regression coefficient based on means and common regression coefficient within subgroups	S_4	1	37.28	37.28
Scores about regression line with common slope b_w	$S_w=S_1+S_2$	N-k-1 95	1772.42	18.66
Subgroup means about regression line with slope b_w	$S_b=S_3+S_4$	k-1 15	306.42	20.43
Scores about regression line for total group (all subgroups)	S_T	N-2 110	2078.84	18.90

Test for Homogeneity of Regression

$$F=\frac{S_1}{S_2}=\frac{11.22}{20.05}=0.56$$

$F_{1\%}(15,80)=2.28$ Accept

Test for Linearity of Regression

$$F=\frac{S_3}{S_4}=\frac{19.22}{18.66}=1.03$$

$F_{1\%}(14,95)=2.32$ Accept

TABLE D-2

COMPUTATION TO TEST HOMOGENEITY OF VARIANCE OF SUBGROUPS ON PAT

Subgroup	N	n	ns^2	s^2	log s^2	n log s^2	1/n
1	3	2	0.50	0.25	9.39794-10	18.79588-20	.50
2	3	2	4.50	2.25	0.35218	0.70436	.50
3	11	10	158.99	15.90	1.20140	12.01400	.10
4	11	10	203.91	20.39	1.30942	13.09420	.10
5	3	2	12.74	6.37	0.80414	1.60828	.50
6	3	2	38.00	19.00	1.27875	2.55750	.50
7	11	10	81.56	8.16	0.91169	9.11690	.10
8	11	10	130.02	13.00	1.11394	11.13940	.10
9	3	2	0.50	0.25	9.39794-10	18.79588-20	.50
10	3	2	240.63	120.32	2.08034	4.16068	.50
11	11	10	57.33	5.73	0.75815	7.58150	.10
12	11	10	282.99	28.30	1.45179	14.51790	.10
13	3	2	0.68	0.32	9.50515-10	19.01130-20	.50
14	3	2	29.83	14.92	1.17377	2.34754	.50
15	11	10	233.82	23.38	1.36884	13.68840	.10
16	11	10	128.18	12.82	1.10789	11.07890	.10
						160.92262-60	
	112	96	1604.18			100.92262	4.80

N = no. of subjects
n = N-1 (df)

ns^2 = adjusted sums of squares
s^2 = variance

$\log ns^2 = \log 1604.18 = 3.20526$
$\log n = \log 96 = \underline{1.98227}$
1.22299

$B' = 2.3026\ 96(1.22299) - 100.92262$
$= 2.3026(16.48442)$
37.9570

$$C = 1 - \frac{1}{3(16-1)}\left[4.80 - \frac{1}{96}\right] = 1.0643$$

$$B = \frac{B'}{6} = \frac{37.9570}{1.0643} = 35.66$$

$X^2 = 30.58 < 35.66$ Reject

TABLE D-3

COMPUTATION TO TEST HOMOGENEITY OF REGRESSION AND LINEARITY OF REGRESSION OF SUBGROUPS ON MTAI

Nature of Variation	Symbol	df	Sums of Squares	Variance
Subgroup regression coefficients about common coefficient	S_1	k-1 15	4752.85	316.86
Scores about regression line for their own subgroup	S_2	N-2k 80	31145.20	389.32
Subgroup post-test means about regression line based on means	S_3	k-2 14	8288.55	592.04
Difference between regression coefficient based on means and common regression coefficient within subgroups	S_4	1	58.61	58.61
Scores about regression line with common slope b_w	$S_w = S_1 + S_2$	N-k-1 95	35898.05	377.87
Subgroup means about regression line with slope b_w	$S_b = S_3 + S_4$	k-1 15	8347.16	556.48
Scores about regression line for total group (all subgroups)	S_T	N-2 110	44245.21	402.23

Test for Homogeneity of Regression

$$F = \frac{S_1}{S_2} = \frac{316.86}{389.32} = 0.81$$

$F_{1\%}(15,80) = 2.28$ Accept

Test for Linearity of Regression

$$F = \frac{S_3}{S_w} = \frac{592.04}{377.87} = 1.57$$

$F_{1\%}(14,95) = 2.32$ Accept

TABLE D-4

COMPUTATION TO TEST HOMOGENEITY OF VARIANCE OF SUBGROUPS ON MTAI

Subgroup	N	n	ns^2	s^2	log s^2	n log s^2	1/n
1	3	2	108.41	54.21	1.73408	3.46816	.50
2	3	2	1675.17	837.59	2.92303	5.84606	.50
3	11	10	3587.01	358.70	2.55473	25.54730	.10
4	11	10	1493.41	149.34	2.17418	21.74180	.10
5	3	2	1129.17	564.59	2.75173	5.50346	.50
6	3	2	530.78	265.39	2.42388	4.84776	.50
7	11	10	3632.24	363.22	2.56017	25.60170	.10
8	11	10	2313.00	231.30	2.36418	23.64180	.10
9	3	2	608.79	304.40	2.48344	4.96688	.50
10	3	2	1384.47	692.24	2.84025	5.68050	.50
11	11	10	3289.15	328.92	2.51709	25.17090	.10
12	11	10	2092.28	209.23	2.32062	23.20620	.10
13	3	2	912.15	456.08	2.65904	5.31808	.50
14	3	2	125.10	62.55	1.79623	3.59246	.50
15	11	10	5157.75	515.78	2.71246	27.12460	.10
16	11	10	3106.22	310.62	2.49223	24.92230	.10
	112	96	31145.10			236.18096	4.80

N = no. of subjects
n = N-1 (df)

ns^2 = adjusted sums of squares
s^2 = variance

$\log ns^2 = \log 31145.10 = 4.49339$
$\log n = \log 96 = 1.98227$
2.51112

$B' = 2.3026\ 96(2.51112) - 236.18096$
$= 2.3026(4.88656)$
11.2518

$$C = 1 + \frac{1}{3(16-1)}\left[4.80 - \frac{1}{96}\right] = 1.0643$$

$$B = \frac{B'}{C} = \frac{11.2518}{1.0643} = 10.57$$

$\chi^2_{.99} = 30.58 > 10.57$ Accept

TABLE D-5

COMPUTATION TO TEST HOMOGENEITY OF REGRESSION AND LINEARITY OF REGRESSION OF SUBGROUPS ON MES

Nature of Variation	Symbol	df	Sums of Squares	Variance
Subgroup regression coefficients about common coefficient	S_1	k-1 15	207.50	13.83
Scores about regression line for their own subgroup	S_2	N-2k 80	1294.33	16.18
Subgroup post-test means about regression line based on means	S_3	k-2 14	176.78	12.63
Difference between regression coefficient based on means and common regression coefficient within subgroups	S_4	1	10.70	10.70
Scores about regression line with common slope b_w	$S_w = S_1 + S_2$	N-k-1 95	1501.83	15.81
Subgroup means about regression line with slope b_w	$S_b = S_3 + S_4$	k-1 15	187.48	12.50
Scores about regression line for total group (all subgroups)	S_T	N-2 110	1689.31	15.36

Test for Homogeneity of Regression

$$F = \frac{S_1}{S_2} = \frac{13.83}{16.18} = 0.85$$

$F_{1\%(15,80)} = 2.28$ Accept

Test for Linearity of Regression

$$F = \frac{S_3}{S_w} = \frac{12.63}{15.81} = 0.80$$

$F_{1\%(14,95)} = 2.32$ Accept

TABLE D-6

COMPUTATION TO TEST HOMOGENEITY OF VARIANCE OF SUBGROUPS ON MES

Subgroup	N	n	ns^2	s^2	log s^2	n log s^2	1/n
1	3	2	12.65	6.33	0.80140	1.60280	.50
2	3	2	2.67	1.34	0.12710	0.25420	.50
3	11	10	94.09	9.41	0.97359	9.73590	.10
4	11	10	141.44	14.14	1.15045	11.50450	.10
5	3	2	9.53	4.77	0.67852	1.35704	.50
6	3	2	8.45	4.23	0.62634	1.25268	.50
7	11	10	211.29	21.13	1.32490	13.24900	.10
8	11	10	159.40	15.94	1.20249	12.02490	.10
9	3	2	0.01	0.01	0.00000	0.00000	.50
10	3	2	54.00	27.00	1.43136	2.86262	.50
11	11	10	90.77	9.08	0.95809	9.58090	.10
12	11	10	143.09	14.31	1.15564	11.55640	.10
13	3	2	4.17	2.04	0.30963	0.61926	.50
14	3	2	3.28	1.64	0.21484	0.42968	.50
15	11	10	109.63	10.96	1.03981	10.39810	.10
16	11	10	249.86	24.99	1.39777	13.97770	.10
	112	96	1294.33			100.40568	4.80

N=no. of subjects
n =N-1 (df)

ns^2 =adjusted sums of squares
s^2 =variance

log ns^2 =log 1294.33 = 3.11197
log n =log 96 = 1.98227
1.12970

B' = 2.3026 96(1.1297)-100.40568
= 2.3026 (8.04552)
18.5256

$$C = 1 - \frac{1}{3(16-1)}\left[4.80 - \frac{1}{96}\right] = 1.0643$$

$$B = \frac{B'}{C} = \frac{18.5256}{1.0643} = 17.41$$

$X^2_{.99} = 30.58 > 17.41$ Accept

TABLE D-7

COMPUTATION TO TEST HOMOGENEITY OF VARIANCE OF SUBGROUPS ON CES

Subgroup	N	n	ns^2	s^2	log s^2	n log s^2	1/n
1	3	2	266.67	133.33	2.12493	4.24986	.50
2	3	2	2.00	1.00	0.00000	0.00000	.50
3	11	10	690.99	69.10	1.83948	18.39480	.10
4	11	10	651.64	65.16	1.81398	18.13980	.10
5	3	2	272.67	136.33	2.13459	4.26918	.50
6	3	2	50.67	25.33	1.40364	2.81728	.50
7	11	10	824.00	82.40	1.91593	19.15930	.10
8	11	10	670.18	67.02	1.82620	18.26200	.10
9	3	2	158.00	79.00	1.89763	3.79526	.50
10	3	2	186.00	93.00	1.96848	3.93696	.50
11	11	10	1996.55	199.66	2.30246	23.02460	.10
12	11	10	372.00	37.20	1.57054	15.70540	.10
13	3	2	78.00	39.00	1.59106	3.18212	.50
14	3	2	266.00	133.00	2.12385	4.24770	.50
15	11	10	715.00	71.50	1.85431	18.54310	.10
16	11	10	230.55	23.06	1.36286	13.62860	.10
	112	96	7430.92			171.30650	4.80

N = no. of subjects
n = N-1 (df)

ns^2 = adjusted sums of squares
s^2 = variance

$\log ns^2 = \log 7430.92 = 3.87104$
$\log n = \log 96 = 1.98227$
1.88877

$B' = 2.3026\ 96(1.88877) - 171.30650$
$= 2.3026\ (10.01542)$
23.0615

$$C = 1 - \frac{1}{3(16-1)}\left[4.80 - \frac{1}{96}\right] = 1.0643$$

$$B = \frac{B'}{C} = \frac{23.0615}{1.0643} = 21.66$$

$\chi^2_{.99} = 30.58 > 21.66$ Accept