

DIFFERENCES IN LEARNING AND ATTITUDES OBTAINED
FROM TWO METHODS FOR FEEDBACK OF MID - TERM
EXAMINATION RESULTS

Thesis for the Degree of Ph.D.

MICHIGAN STATE UNIVERSITY

JOHN B. BALCH

1969

This is to certify that the
thesis entitled
Differences in Learning and Attitudes Obtained from
Two Methods for Feedback of Mid-Term
Examination Results

presented by

John B. Balch

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Education

Walter F. Johnson
Major professor

Date April 17, 1969

~~JUN 9 1970~~ 1371

~~OCT 9 1970~~ RS4

OCT 25 1970 ENA9 6

~~Z 11 021~~

~~NOV 1 1970~~

~~JUN 9 1970~~ 2

57230
7-3-69

ACKNOWLEDGMENTS

The author wishes to express his appreciation to the members of his guidance committee, Dr. Walter F. Johnson, Chairman; Dr. Harold W. Grant; Dr. William E. Sweetland; Dr. Donald L. Grummon; and Dr. Donald J. Freeman. Particular respect and admiration is expressed for Dr. Walter F. Johnson, whose continued support, interest and friendship throughout the entire program of doctoral studies and related research proved to be a consistent source of motivation.

To my wife and parents this study is dedicated, without whose love and kind understanding it might never had been brought to completion.

ABSTRACT

DIFFERENCES IN LEARNING AND ATTITUDES OBTAINED FROM TWO METHODS FOR FEEDBACK OF MID-TERM EXAMINATION RESULTS

by John B. Balch

This study had three purposes.

First, using as a control group students who listened to a televised explanation of test results, it was to determine whether students who discussed their answers to an objective test (a mid-term) with each other in small groups scored better on a subsequent objective test (a final examination) than the television group. In the first two parts of the study, student scores on repeated test items and their scores on the total final examination were investigated.

The second purpose of the study was to determine whether there was any positive relationship between students' final exam scores and their tested tendency to participate in discussion.

Third, it was to determine whether there were any response differences on an exploratory attitude questionnaire relevant to the purposes and content of exam feedback when administered to different students assigned to the discussion or television group.

Design and Procedure of Study

The learning experiment was conducted with 166 students in the Summer Term and the attitude questionnaire was administered to 787 students in the Fall Term of 1967.

Analyses of covariance were conducted to evaluate the experimental treatment effects of students' discussing or watching TV.

Scores on students' tested tendency to discuss were classified as above or below the mean. The interaction of these levels of discussion tendency were analyzed with the treatment effect of having or not having the opportunity to discuss mid-term test items.

Students indicated their attitude responses on a five point scale of agreement. The Mann-Whitney U statistic was used in analyzing students' responses to each of the thirty-one questionnaire items. The problem of objectively classifying the questionnaire items was met by factor analyzing the attitude scale.

Findings

Although the mid-term exam scores were equated in analyses of covariance, no significant differences were found on final exam scores and their sub-parts between groups of students discussing mid-term exam results and students watching a televised presentation of these results. Of the thirty-one items exploring attitudes about examination feedback

listed in the questionnaire, eighteen items revealed significant differences between the groups. Twelve of the items favored the instructional effectiveness of the discussion treatment and six of the items favored the television treatment with respect to student agreement on the questionnaire.

When the questionnaire was factor analyzed, three of the items which had significantly favored the discussion group loaded on one factor. The discussion group agreed that (1) the feedback session stimulated them to interact with other students, (2) they felt personally involved, and (3) the session was not formal. This factor seemed relevant to the rationale of this study, namely that the discussion group more actively participated and was more personally involved.

DIFFERENCES IN LEARNING AND ATTITUDES OBTAINED FROM
TWO METHODS FOR FEEDBACK OF MID-TERM
EXAMINATION RESULTS

By

John B. Balch

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Administration and Higher Education

1969

TABLE OF CONTENTS

Chapter		Page
I	THE PROBLEM	1
	Introduction	1
	Purpose of the Study	2
	Statement of the Problem	2
	Need	3
	Assumptions	4
	Hypotheses	6
	Definition of Terms Used	7
	Limitations of the Study	8
	Organization of Study	9
II	REVIEW OF LITERATURE	11
	Classroom Communication	11
	Importance of Testing for Students . .	12
	Research Related to Student Discus- sion of Examination Results	13
	Identification of Possible Inter- actions Between Students and Method of Feedback	15
	Student Attitudes to Television Versus Conventional Instruction . .	16
	Summary	20
III	METHODOLOGY	21
	Population and Sample	21
	Instrumentation	22
	Reliability Coefficients in This Study . .	23
	Validity Procedures	24
	Factor Analysis Procedures	26

Chapter		Page
	Statistical Hypotheses	26
	Hypotheses About Achievement	26
	Attitude Hypotheses	30
	Data Collection	30
	Analysis of the Data	32
	Pre-Experiment Analyses of Variance	34
	Covariance Statistical Methodology of Experiment	35
	Covariance Assumptions	36
	Attitude Analysis	37
	Summary of Methodology	38
IV	ANALYSIS OF RESULTS	39
	Experiment I: The Effect of Method of Presenting Feedback and Student Discussion Tendency on Final Exam Performance	39
	Assumption of Homogeneity of Regression Uncontrolled Variables	39 42
	Hypotheses About Achievement	43
	Final Exam Means	45
	Analyses of Covariance Results	47
	Experiment II: Student Attitudes Toward Different Methods of Presenting Feedback	50
	Significant Attitude Items	53
	Non-Significant Attitude Items	56
	Factor Analysis	57
	Summary.	60
V	FINDINGS, CONCLUSION, AND SUMMARY	62
	The Problem	62
	Design and Procedures of the Study	62
	Results of Learning Experiment	66
	Results of Attitude Experiment	66
	Attitude Questionnaire Factor Analyzed: Consistent Results	67

Chapter	Page
Other Significant Differences Within Factors	69
Conclusions	73
Implications for Further Research	74
BIBLIOGRAPHY	78
APPENDICES	80

LIST OF TABLES

Table		Page
2.1	Summary of the Differences Between Television vs. Conventional Teaching	17
4.1	Analyses of Variance for Departure from Homogeneous Regression Using Final Examination as the Dependent Variable	40
4.2	Mean Scores on the Final Examination of Television and Discussion Groups, Students Above and Below the Average on Tested Tendency to Discuss, and Their Interactions* .	46
4.3	Analysis of Covariance Using the Total Final Exam as the Dependent Variable	47
4.4	Analysis of Covariance Using the Final Exam Items Repeated from the Mid-Term Exam as the Dependent Variable	48
4.5	Analysis of Covariance Using the Final Exam Items Modified from the Mid-Term Exam as the Dependent Variable	49
4.6	Analysis of Covariance Using the Final Exam Items Repeated and Modified from the Mid-Term Exam as the Dependent Variable	49
4.7	Differences in Student Attitude Toward Different Methods of Presenting Feedback on the Mid-Term Exam	51

LIST OF APPENDICES

Appendix		Page
A	Influence of Evaluating Procedures	80
B	Instructions to Teaching Assistants	81
C	Instructions to Subjects	83
D	TSIS	85
E	Course Questionnaire	87
F	Factors, Factor Loadings, Mann-Whitney U Values, and Frequency of Agreement on Questionnaire Responses of TV and Discus- sion Groups	90
G	Rotated Factor Loadings	106

CHAPTER I

THE PROBLEM

Introduction

In almost every college course, one obvious aim is to provide students with information. To improve the students' grasp of such information, some educators and psychologists recommend that instructors use test results for more than merely evaluating the students' current performance. That this is desirable is unquestionable, since it is relevant to the whole learning process.

The feasibility of using tests to teach as well as evaluate is suggested by two studies. Curtiss and Wood (3) and Sassenrath and Gaverick (20) found that when students discussed the results of their objective classroom examinations they learned more than if they merely viewed an authority's explanation of these answers. In these two studies neither test scores alone, nor referral to the chalkboard and textbook was as effective an aid to learning as students' discussion of the correct answers with their teacher.

This investigation follows these two studies closely, though it employs a different rationale, different variables, and in addition, explores students' attitudes about how they have received their examination results.

Purpose of the Study

The general purpose of this pilot study was to improve instructional effectiveness from the way students receive results of their performance on objective examinations.

This study had three specific purposes. First, using as a control group students who listened to a televised explanation of test results, it was to determine whether students who discussed their answers to an objective test (a mid-term) with each other in small groups scored better on a subsequent objective test (a final examination) than the television group. In the first two parts of the study, student scores on repeated test items and their scores on the total final examination were investigated. The second purpose of the study was to determine whether there was any positive relationship between students' final exam scores and their tested tendency to participate in discussion. Third, to determine whether there were any response differences on an exploratory attitude questionnaire relevant to the purposes and content of exam feedback when administered to different students assigned to the discussion or television group.

Statement of the Problem

The problem of this investigation was to offer three sets of experimental data relevant to the three purposes of the study and to interpret these data with respect to improving

instructional effectiveness. A subsidiary problem of the study was to explore why classroom discussion of examination results might improve instructional effectiveness in comparison to televising these results.

It is to be remembered from page one that Curtiss and Wood (3) and Sassenrath and Gaverick (20) found that instruction was more effective when students discussed their objective classroom examinations than if they merely viewed an authority's explanation of these answers. But in this study the effect of small group, student-led discussion of examination answers is compared to the effect of students viewing authorities' explanations of these answers via closed-circuit television.

Need

During the school year when this investigation was conducted, students in the educational psychology course for Teacher Education majors at Michigan State University had experienced televised presentations of the correct answers to their mid-term examinations. Some of the students and their instructors felt that students appeared to be so passive during these televised presentations that little learning was taking place. Hence this contention seemed to merit investigation by comparing the effects on learning and attitudes by students who received explanations to their mid-term examination by means of television with students who discussed their mid-term examination results among themselves.

Assumptions

Results stemming from this study may make some contribution to theory construction. The rationale upon which the study is based was to explore the feasibility of the idea that active involvement and participation in small groups might explain why classroom discussion could be an effective instructional aid.

A number of experiments have demonstrated that active learning is more effective than passive learning (McGeoch and Irion, 11). Both Bloom (2) and Krauskopf (9) have presented evidence that when discussions are effective, they promote active thinking about what went on in class and these discussions correlate with achievement measures relevant to these thoughts. Korn and Black (10) and Stern et. al. (18) have found a positive relationship between the amount of classroom participation and academic achievement.

It seems logical to assume that students who merely watch TV are less actively involved mentally when they receive examination results than students who listen to exam results followed by a discussion of these results. One way to explore the feasibility of whether there was active participation as a result of the feedback session involved determining whether students' attitudinal responses showed they were stimulated to interact with each other and were personally involved when discussing examination answers among themselves. In contrast the feasibility of the idea (that student discussion involves active participation as present in this study) will be

supported if the attitudes of students passively listening to TV indicated that they were not personally involved and stimulated to interact with each other.

By the rationale of this study, classroom discussion would seem to be a more effective instructional medium than televised instruction, especially for students whose test scores indicate they have a tendency to be active discussants in the classroom.

In a review of television instruction, McKeachie (12:42) has summarized the relevant research:

In a majority of experiments in which there were adequate controls, greater learning occurred in the 'live' classes than in those taught by television. Most of these differences were not statistically significant by themselves but their consistency is statistically significant. One can thus conclude that at the college level television is generally not as effective as face-to-face instruction.

In a different summary of research and by the same rationale, McKeachie (Gage, 7:1149) concludes that college students in television classes tend to rate courses and instructional television less favorably than students in conventionally taught classes.

McKeachie, then, has proposed that televised teaching will not eliminate the need for live classroom participation. He (12:22-25) also offers research results supporting the advantages for student-led discussion over teacher-led instruction with respect to both student achievement and how students rate courses.

However, from this investigator's (1) review of the literature, there does not seem to be any research evidence specifically evaluating TV techniques as a method of examination feedback. Nor has any experimenter compared TV and discussion conditions with individuals whose test scores indicate they have different tendencies for active discussion in the classroom.

It is, then, significant for this study that Giddan et. al. (8) have emphasized that different individuals have different discussion tendencies. Hence it seemed useful to try to determine whether students would benefit differentially from different methods of test feedback.

Hypotheses

The two hypotheses to be tested stem from the research reviewed in Chapter I and Chapter II and from the investigator's (1) previous research presented in Appendix A. The basis for these hypotheses, then, is that learning is affected by both the external elements of the situation, including the importance of testing to students and their activity during the discussion of their examination results, and by internal factors characteristic of the individual learner's tendency to participate actively in classroom discussions.

The two hypotheses to be tested in this study are:

1. There is a significant difference in achievement between students discussing mid-term examination scores in small, student-led groups and students listening to a closed-circuit TV presentation of

these results as measured on (a) final examination items presented previously on the mid-term examination and during a feedback session and (b) the total final examination.

2. There are significant interactions among individuals' student discussion tendencies, students' final examination scores and their sub-parts, and whether students have experienced discussion versus TV feedback of their mid-term examination scores.

Based on the results of the research of Davis and Johnson (4) and McKeachie's (12:43) review of research discussed in the next two chapters, it is also tentatively hypothesized that students' attitudinal responses will indicate that they prefer discussing their examination results to watching a televised presentation of these results. However as this is a pilot study, the purpose of analyzing specific attitudinal preferences is exploratory; it is possible that more definitive statements as to the operation of these variables can be offered subsequent to the examination of the data.

Definition of Terms Used

To avoid semantic difficulties, the following terms were defined. These definitions as used in the context of this study were as follows:

Discussion Group. Students who discussed the results of their performance on their objective mid-term examination in small student-led groups.

Television Groups. Students who received results of their objective mid-term examination by television.

Feedback. Receiving results on objective mid-term examination.

Active Participation. Feedback by discussion.

Student Discussion Tendency. Student score on a scale employed to evaluate discussion tendency. This scale is labeled the Research Form B of the Academic Behavior Inventory, but it is also referred to in this study by the author's abbreviation TSIS. See Giddan et. al. (8).

Learning and/or Achievement. Students' scores on the final examination and sub-parts of the final examination which consisted of repeated and modified items from the mid-term examination.

Attitudes. Students' reactions to a questionnaire designed to assess their reactions to feedback.

Limitations of the Study

The students in this study were enrolled in the Summer and Fall Terms of Education 200, a course in educational psychology for Teacher Education majors at Michigan State University. Since the learning or achievement of these students must take precedence over other considerations, the investigators were not permitted to conduct experimental studies which would in any way interfere with normal classroom procedures such as drastically varying the method of television feedback or altering the number and kind of tests and test items students were to experience.

In accordance with an agreement that experimental testing would not take too much classroom time, it was agreed that students would receive discussion tendency tests only during the Summer Term and attitudinal questionnaires once in the Fall Term. These tests were exploratory instruments and may not sample adequately the population of possible items

with respect to student discussion tendency and student attitudes about feedback. Further there was no attempt to control variables between the Summer and Fall Term courses. It was assumed that students were exposed to the same course and that experimental techniques from one term could be applied in the next term to explore their effect.

As the application of theory in this study was also explorative, it is assumed that the results may make some contributions to theory construction. The limitation of this applied study, then, is in the particular way evidence will be accumulated on the conceptual variable of active personal involvement and in the more important conceptual matter itself.

Organization of Study

Chapter I presents an introduction of the study and to the purposes of the study, a statement of the problem, definition of terms, limitations of the study, and the hypotheses.

Chapter II reviews the literature on classroom participation and television instruction, the importance of discussion during feedback of examination results, and the advantages to students who are more active during classroom discussions. Chapter II also reviews the scales which evaluate individual differences in a) student discussion tendency and b) attitudes towards televised versus conventional instruction.

Chapter III describes the selection of the experimental groups that discussed their answers to tests during feedback and the selection of those students who listened to a televised explanation of test results. This chapter presents the procedures for testing the students' tendency to participate in discussion in small groups, the method for selecting the items from the mid-term feedback session that could reappear on the final examination, and the statistical methods used to analyze the data. The criteria of learning and students' attitudes are also described.

Chapter IV presents the analyses of (a) final examination scores obtained from comparing the scores of students who actively participated during discussion of test results to the scores of students who listened passively to a television presentation of test results and (b) how these experimental conditions of activity and passivity relate to students' tested tendency to discuss and (c) attitudinal differences between experimental and control groups.

Chapter V summarizes the problem, methodology, results, and implications for further study.

CHAPTER II

REVIEW OF LITERATURE

The nature of this study necessitated a general review of research pertinent to classroom communication, the relationship of classroom testing to learning and discussion, and the relevance of the scales used in this study.

Classroom Communication

The importance of classroom communication can be emphasized in several ways. A number of experiments have demonstrated that active learning is more effective than passive learning (McGeoch and Irion, 11). Korn and Black (10) and Stern, et. al. (18) have found a positive relationship between the amount of participation and several criteria of academic achievement, such as the development of critical thinking and logical expression of ideas. Opportunity for discussion also leads to the achievement of other higher educational goals, such as active thinking (Bloom, 2), problem-solving and a scientific attitude (McKeachie, 13), and increased academic motivation (Thistlethwaite, 19).

Nevertheless many students do not participate in class discussions and many institutions have no active student discussion in their classrooms. Pace (17:82) has stated: "In almost none of the colleges and universities was it said by

the bulk of the students that class discussions were typically vigorous and intense."

Likewise, Korn and Black (10) reported that one-third of the students they studied participated in class below the minimal expected amount.

Importance of Testing for Students

It is the rationale of this study that students are apt to participate in discussions about topics concerning them. McKeachie (15) has suggested that students are more concerned about how they are to be examined than about any other aspect of their instruction.

In a previous study, the present investigator (1) surveyed the literature and gathered strong support for the assumption that classroom testing procedures influence the nature and consequences of student learning (see Appendix A). McKeachie (Gage, 7:1154) analyzed studies related to feedback from examinations and summarized them as follows:

While we usually think of testing procedures in terms of their validity as measures of student achievement, their function as instruments prompting learning may be even more important. After dismal recitals of nonsignificant differences between differing teaching methods, it is refreshing to find positive results from variations in testing procedures.

These conclusions suggest that the activity or passivity of students during a discussion of examination results may be educationally significant factors and hence support the need for conducting the present experimental study.

Research Related to Student Discussion of Examination Results

The following studies support the effectiveness of group discussion as a method of active participation during feedback of examination results.

Curtiss and Wood (3) assessed the effect of having students actively check the incorrect items on their papers during a discussion of objective test items. This active student involvement resulted in greater pupil learning than having the teacher correct and return the papers without having student discussion. Sassenrath and Garverick (20) reported that students who discussed examination results learned more than (a) students who reread sections of their textbook after checking their answers and the appropriate page references from a correct list on the chalkboard or (b) students who had no opportunity to see their examinations but merely were given their total scores on the examination and their letter grades. In these two studies on the effects of discussion during feedback, the critical factor appears to be verbal interaction with others.

It has been said that it is a principle of learning that active learning is more effective than passive learning, and Bloom (2) found that students in discussion classes tend to spend more time in problem-solving types of thought than students in lecture classes. Both Bloom (2) and Karuskopf (9) have presented evidence that when discussions are effective, they promote active thinking about what went on in class and these discussions correlate with achievement measures relevant

to these thoughts.

There is also evidence that active writing of reactions to test materials promotes learning. McKeachie (14) found that students who were permitted to write whatever comments they wanted about their test items as they took their objective test did better on the second half of that test. These results suggest that allowing students to discuss their answers orally after taking a test might also improve their performance on a succeeding test.

Another advantage of discussing test results derives from social and clinical studies of attitudinal and personality change. Failure of some students to achieve goals of learning may not be due to deficiencies in the materials presented but rather to emotional barriers in the learner. Social and clinical psychological theory can be interpreted to suggest that if a student expresses his attitudes in a non-threatening situation it may help "unfreeze" the attitude. Small, student-led group discussions may provide such opportunities for expression as well as give opportunities to express other attitudes about testing which may be instrumental to meeting students' instructional needs. Consequently, the social influence of the small group may facilitate change during discussion of test results. McKeachie (12:19) may support this view when he interprets classic experiments on small group decision-making to show that it is sometimes easier for a group meeting as a group than for an individual, who may not have the

feeling he is in a group to change attitudes.

McKeachie (12:24) also interprets the many experiments on student-led discussions within this framework. For example, he states that Webb and Grib report that students report as a major advantage of student-led discussions their feeling of freedom to ask questions and express their own opinions. McKeachie (12:22-23) quotes Leuba of Antioch College to say:

A student is likely to understand a concept, principle, or other idea as he questions himself about it, looks for its implications and applications, puts it in his own words, and integrates it with previously acquired knowledge . . . Leuba reports satisfying results (from instructorless student discussion groups) in psychology courses and Webb and Grib report superiority of this technique in achievement in statistics and philosophy courses. . .

McKeachie (12:22-23) interprets the experiments in educational psychology, and general psychology by Gruber and Weitman to have established that students taught in small discussion groups without a teacher do at least as well on a final examination as students who heard the teacher lecture, but they were also superior in curiosity (as measured by question-asking behavior) and in interest in educational psychology.

Identification of Possible Interactions Between Students and Method of Feedback

As it might relate to learning in this study, there was an attempt to identify the student characteristics that interacted with the two feedback variables of the experiment: Passively listening to a TV explanation as opposed to small

group discussion of examination results.

The Research Form B of the Academic Behavior Inventory or the TSIS, a scale developed by Giddan et. al. (8) has twenty-four true-false items which predict the interaction between students and their discussion leaders in discussion sections. These investigators state that: "Where the effects of different teaching methods are being studied [as in the present study] this scale could assess the role of discussion tendencies on response to differing methods of instruction." Hence in the present study it was thought that students who scored higher on this student interaction scale would have the potential to score higher on repeated test items and the whole final examination, under conditions of discussion feedback, as compared to similar students experiencing TV feedback.

Student Attitudes to Television Versus Conventional Instruction

The exploratory question raised in this study is whether the attitudes of students receiving examination results by closed-circuit television are better than the attitudes of students who discuss these results.

Student attitudes toward closed-circuit television in general tend to be influenced by the interaction of a number of variables. All other things being equal, college students in television classes tend to rate courses and instructional television less favorably than students in conventionally taught classes: Support for this view is presented in Table 2.1 adopted from McKeachie's (12:43-44) review of research

Table 2.1. Summary of the Differences Between Television vs. Conventional Teaching (Reproduced from McKeachie's Table 4 in New Developments in Teaching).

Reference	Factual Knowledge	Retention Higher Level Cognitive	Attitude Motivation Personality
<u>Social Science</u>			
Carpenter and Greenhill Psych. Sociol.	C,C C	C	C
Macomber and Siegel Psych. Sociol. Econ.	TV,C,C*,C,C,C TV C,C*	C C*,C*,C	C
LePore and Wilson (Psych)	TV	TV	C
<u>Humanities</u>			
Carpenter and Greenhill (Music Apprec.)	C		
LePore and Wilson (English)	?		C
Klapper (English)	C,TV	C,C	
Seibert (English)	C*	C	
<u>Natural Science and Engineering</u>			
Carpenter and Greenhill Chemistry Meteorology	C,TV C		
Macomber and Siegel Physiol. Biol. Zoology	TV,C,C TV*	TV C	

Table 2.1. (Continued)

Reference	Criteria		
	Factual Knowledge	Retention Higher Level Cognitive	Attitude Motivation Personality
Seibert Chem. Mech. Engin.	C C	C	
Martin et. al. Chem. Graphics	TV* TV*		
LePore and Wilson (Science)	C,TV	TV,C	C,C
<u>Miscellaneous</u>			
Macomber and Siegel (Air Science)	TV,C,TV,C*C TV,C,C		
Seibert Math. Calculus	TV C		
Kasten and Seibert (Military Sci.)	C		
Grossman et. al. (Dentistry)	TV,C,C		
McDaniel and Filiatreau (Educ.)			C*

* Significant at beyond the .05 level of confidence

TV = TV superior

C = Conventional class superior

Differences are simply the actual direction of results of the experiment; when more than two measures were used, the Table reports the direction of the majority of the measures.

with college students.

Davis and Johnson (4) had administered a questionnaire to Michigan State students similar in construction to the one used in this study. They had factor analyzed their questionnaire to validate their a priori grouping of seventeen questionnaire items. Their final grouping was based on a priori judgment and factor analysis of the relevant attitudinal factors of students at Michigan State viewing television versus conventional classroom instruction. The final groupings were:

1. Stimulation of the students
2. Student-(Teacher) interaction
3. Clarity of presentation
4. Attention

Their questionnaire also contained items concerning optimum viewing conditions and the opportunities for additional viewing from which they were able to draw overall conclusions about the effectiveness of television instruction. With the exception of these questions, the kinds of problems Davis and Johnson (4) were trying to solve covered four aspects of instruction at Michigan State University (interest and stimulation, student-(teacher) interaction, clarity of presentation, and attention). Responses of students viewing television were compared with students viewing more conventionally taught courses. Differences between the two groups were tested and nine statistically significant results were obtained to the seventeen factor analyzed questions. Only three of these

differences favored their television groups.

Summary

Leaders in the field of education are concerned about the lack of discussion in classrooms. The advantages of discussing test results were reviewed in this chapter. There was also a review of scales which evaluate individual differences in a) student discussion tendency and b) student attitudes toward television versus conventional instruction.

CHAPTER III

METHODOLOGY

Population and Sample

This investigation was conducted at Michigan State University in the discussion sections of beginning educational psychology students. Most of these students were sophomores and juniors who planned to teach in elementary and secondary education after graduation. One hundred and sixty students were tested in the summer learning experiment: Eighty-one students composed the discussion groups and seventy-nine students the TV groups. Students from MSU Education 200 sections seemed to be appropriate subjects as they are accustomed to interacting together and to viewing the closed-circuit TV sets which are left throughout the term in their discussion sections. During Summer Term, each of the eight sections of students in the course was randomly assigned to one or the other of the experimental treatment conditions.

In Fall Term, an exploratory attitude questionnaire was administered to seven hundred eighty-seven students in thirty of the thirty-five sections of beginning educational psychology. Each section was assigned randomly to the experimental treatment conditions.

Instrumentation

A differential student discussion effect was investigated with the Research Form B of the Academic Behavior Inventory or, as it is also referred to, the Teacher-Student Interaction Scale (TSIS). Employing the Kudar-Richardson Formula 20, Lovell and Giddan found that it produced a reliability coefficient of .77 (Giddan et. al., 8). Across their eight samples of undergraduate discussion sections, the Research Form B of the Academic Behavior Inventory had been found to correlate with instructor ratings for quality and quantity of student discussion. Giddan, et. al. state "all but two (of sixteen) validity coefficients reached acceptable levels of statistical significance, with a majority of them having a $p < .001$. The median correlation for quantity of discussion was .61 and for quality of discussion .41."

Across their eight samples of undergraduate sections, Giddan, et. al. (8) demonstrated that the TSIS did correlate with instructor ratings for quality and quantity of student discussion. (The two four-point rating scales were developed and used originally by Stern, et. al., 18:169). Giddan et. al. (8) concluded that "where the effects of different teaching methods are being studied [as in the present study] this scale could assess the role of discussion tendencies on response to differing methods of instruction."

Reliability Coefficients in This Study

In this study the Kudar-Richardson formula 20 reliability coefficients were 1) for the TSIS,.79, 2) for the mid-term examination,.70, 3) for the final examination,.82, 4) for the fourteen items on the final examination repeated from the mid-term examination in modified form,.54, 5) for the ten items on the final examination repeated in the exact form as they appeared on the mid-term examination,.55, 6) for the combined total of the twenty-four items repeated on the final examination from the mid-term examination,.69.

Reliability coefficients were not computed for the exploratative attitude questionnaire used in this study because members of the educational psychology department believed that they would not be valid. Davis and Johnson (4) did not report reliability coefficients to their similar questionnaire but felt that such five point Likert Scales were reliable when testing large numbers of students. For the purposes of exploring attitudes in the present study, reliability was assumed because 787 students were tested on a five point agreement scale.

In this study it was possible to compare the impact of students' discussing exam results to the impact of televising them rather than having students only evaluate television and discussion against some idealized standard. This procedure avoided invalid student judgments based on entirely

different standards.

Validity Procedures

Validity was established with the help of Drs. William E. Sweetland and Donald T. Freeman, experts in the content and the purposes of the course being scaled. They approved each five point Likert item used in the questionnaire in terms of the content and purposes of the course being scaled. Dr. Robert H. Davis advised on the technical construction of each scale item.

Like Davis and Johnson (4), this study used students from Michigan State University in order to permit more meaningful interpretation of both sets of similar questions. The limitations of interpretation should be noted: Different populations of students at MSU, the method of instruction and the exposure time to method of instruction, and the necessary rewording of questions created some real differences in the two studies. But it was thought that these very differences could help explain obtained results better than a set of questions which had not been administered to Michigan State students and which had not been related to television versus more conventional methods of instruction.

There is some evidence that the questions asked by Davis and Johnson (4) would elicit some comparable replies in the present study. They had concluded that most students (eighty-seven per cent) would not object to taking television

instruction under certain conditions. To a Winter Term exploratory opinion questionnaire eighty-seven per cent of the students in the course presently being studied had similarly indicated that television feedback had been "personally profitable." Hence it was thought that the results of the present study would be more interpretable by constructing a similar questionnaire to that of Davis and Johnson (4).

The rationale of the present study was not to follow the rationale of Davis and Johnson (4) but to modify their questions when they fitted the purposes of this study. Davis and Johnson (4) emphasized four aspects of instruction in seventeen of their questions, but this study had over half-again as many additional questions designed to assess other attitudes relevant to this course.

For example, it was the opinion of a course evaluation committee consisting of teaching assistants that students watching television as compared to students discussing examination answers would respond as if the hour were relatively more "boring," "formal," and "oversimplified" and that they would not become as personally involved as they would when they are stimulated by student interaction.

All questionnaire items (e.g., items related to personal involvement and stimulation of students to interact with one another) were factor analyzed to objectively classify the items on the five point attitude scale.

Factor Analysis Procedures

The procedures for factor analyzing the students' responses to the thirty-one items on the attitude questionnaire were:

1. A thirty-one by thirty-one response matrix was built and the intercorrelations among the items were obtained.
2. The principal axis solution was employed to factorize the matrix.
3. The factors were rotated by the varimax method of factor rotation.

Statistical Hypotheses

Hypotheses About Achievement

The research hypotheses of this study were stated in Chapter I. To facilitate statistical testing, these statements were transformed into null or operational form. The two major null hypotheses of this study were as follows:

- I. No differences will be found between groups of six students spending a a classroom hour discussing right and wrong answers of their objective mid-term examination and students who listen to a closed circuit TV presentation of these results as measured on a) final examination items presented previously on the mid-term examination and during a feedback session and b) the total final examination.
- II. No interactions will be found among individual students' discussion tendencies as measured on the Research Form B of the Academic Behavior Inventory, students' final examination scores and their sub-parts, and whether students have experienced discussion versus TV feedback of their mid-term examination scores.

Null Hypotheses I:

Symbolically: $H_0: M_1 = M_2$

Legend: M_1 = Group final examination means of students discussing their initial examination results

M_2 = Group final examination means of students experiencing televised explanations of their initial examination results

Both sets of means are adjusted to equate students' initial examination results.

Alternative Hypothesis I: The discussion group mean scores will exceed the television group mean scores on the final exam and their sub-parts.

Symbolically: $H_{1a}: M_1 > M_2$

Legend: Group final examination means of students discussing their initial examination results

M_2 = Group final examination means of students experiencing televised explanations of their initial examination results

Both sets of means are adjusted in order to equate students' initial examination results.

Alternative Hypothesis II: There will be significant interactions in the means of final exam scores among (1) the two experimental treatments and (2) the two levels of discussion tendency test scores.

Alternative Hypothesis II (a): Within the discussion treatment, the group scoring above the mean on the discussion tendency test will have a higher final exam mean than the group scoring below the mean on the discussion tendency test.

Alternative Hypotheses II (b): Within the TV treatment, there will be no difference on final exam scores between the group scoring above the mean (on the discussion tendency test) and the group scoring below the mean (on the discussion tendency test).

The final exam means are adjusted to equate students' initial exam results.

		FEEDBACK CONDITION	
		TV SCORES	DISCUSSION SCORES
DISCUSSION TENDENCY SCORES	above mean		
	below mean		

An assumption of hypothesis II, then, is that the exploratory discussion tendency test is a valid and reliable measuring instrument. Supportive data for this assumption have been presented.

There are no hypothesized differences between the students in the discussion treatment who score lower on the discussion tendency test and students in the television treatment regardless of their discussion tendency test score.

Support for statistical assumptions used in analyzing the data for both hypotheses will be presented. In interpreting the statistical data, the .05 level of confidence will be used to assess statistical significance.

The above alternative hypotheses are based on the conceptual assumption of the influence of active participation on student learning.

The first alternative hypothesis is based on the assumption that if students are given the opportunity to participate actively by interacting with each other during small group discussions their scores can be interpreted as if they have learned more than if they are not given this opportunity but instead watch television.

The second alternative hypothesis is based on the assumption that scores of students with a tested tendency to participate actively in discussion can be interpreted as if they have done so when given the opportunity to discuss mid-term exam results. Scores of students with less of this tested tendency to discuss are assumed to be interpretable as if they have not actively participated in discussion.

It is assumed that students with a below the mean discussion tendency score will have lower final exam scores if asked to participate in discussions than those with above the mean discussion tendency scores. It is presumed that a low discussion tendency score is related to the lack of skill in the ability to profit from discussion.

As the television groups were not given the opportunity to participate in classroom discussions during feedback, their tested tendency to do so is presumed less relevant to their learning. Nevertheless, the scores of students watching television are assumed to serve as a control or comparison condition between two extremes: Specifically, the comparison is between the hypothesized highest final exam scores of students assumed able to participate actively in discussion

(because of their high discussion tendency scores) and the hypothesized lowest final exam scores of students presumed not able to participate actively in discussions (because of their low discussion tendency scores).

Attitude Hypotheses

An investigation of students' specific attitudinal preferences towards either discussion or televised presentation of examination results was conducted. Although based on the research of Davis and Johnson (4) and McKeachie (22) previously discussed, these attitude tests are of an explanatory nature and are therefore two-tailed.

In null hypotheses form to facilitate statistical testing:

no differences will be found in attitudinal responses between students spending a classroom hour discussing mid-term exam scores and students listening to a closed-circuit TV presentation of these results as measured on each of thirty-one questionnaire items relevant to the purpose and content of examination feedback.

Data Collection

During the Summer and Fall Terms, students in different discussion sections of beginning Education 200 at MSU were presented correct answers to their mid-term examination in one of the two randomly assigned conditions:

- 1) In sections presented with correct test alternatives over closed-circuit TV with a description of why the correct alternatives were correct together with a blown-up picture of the mid-term

examination. These students watched a televised team of two teachers on one tape.

- 2) In discussion sections of small groups of six students, with one graduate assistant giving each group one copy of the correct answers to the test during the feedback session and asking them as a group to participate in discussing why the correct answers were correct.

Graduate assistants assessed the adequacy of the TV presentation and the discussion coverage of test items during the feedback sessions. These raters judged both of the above conditions to be adequate along the dimensions referred to in Appendix B.

During the feedback session, each student in this experiment was given his answer sheet and informed of the correct mid-term answers. In the discussion groups one member of the six-man discussion group was given the correct mid-term answers and asked to present them to his group for discussion. The groups were instructed to discuss each test question. (The instructions are presented in Appendices B and C.) (Pre-tests indicated students had sufficient knowledge of course materials to discuss test questions effectively.)

To facilitate discussion of pertinent course materials, each student in the discussion groups was also given an unmarked examination to be returned at the end of the feedback hour; one student in each discussion group was given the correct answers to the mid-term and the other members of each discussion group had to refer actively to him (see Appendices B and C).

During the Fall Term, attitudes of students who received examination results by television were compared with attitudes of students who discussed these results. The method used was an exploratory student attitude questionnaire.

The questionnaire was administered five days after the examination feedback session to 787 students in thirty sections of educational psychology. Each teacher of a different section in this course had one of his sections of students watch the presentation over television and had the other section of students discuss their examination results (see Appendices B and C). Sections of students were assigned randomly to these experimental conditions so there were equivalent numbers of sections assigned to each condition at each hour the sections were meeting. Generally it took students about ten minutes to fill out the questionnaires and no more than fifteen minutes to administer them.

Analysis of the Data

Evaluation of all summer sections consisted of determining the relative effectiveness of the discussion and television groups on the final examination. For separate statistical analysis and comparison the final exam included fourteen items from the mid-term examination which were modified and ten items which were in the exact form as they appeared on the mid-term test.¹ These repeated items, selected by a panel of

¹The modified items measured the same concept or principle as the corresponding mid-term items, but were presented under different stimulus conditions on the final examination.

judges after the feedback sessions, served to assess whether students had learned the content of these items (and/or the concept related to these items) as a result of one of the two feedback conditions of the experiments.

There were four final examination test score criteria or dependent variables in the summer experiment:

- 1) Number of modified recurring test items answered correctly. (These modified recurring items were chosen at random from adequate items of the item pool in advance of the mid-term. Some of these items had equivalent statistical results from previous test administrations.)
- 2) Number of recurring items answered correctly. (These items were again chosen at random from the adequate items of the item pool in advance of the mid-term.)
- 3) The above modified and recurring test scores combined.
- 4) Total scores on the final exam.

Independent variables included:

- 1) The two treatment conditions outlined above.
- 2) Student discussion tendency scores as assessed on the TSIS. (The TSIS is also referred to as the Research Form B of the Academic Behavior Inventory.)
- 3) Test scores on the mid-term examination.
- 4) Breakdown of test scores on the mid-term examination into sex differences.

As my hypothesis has indicated, small group discussion during feedback of examination results was thought to improve learning as measured by an evaluative test: in the TV groups student discussion was not possible; in the six man discussion groups, discussion was possible and encouraged but not

controlled. It was, therefore, necessary to investigate for a differential discussion effect among the students experiencing television and among those experiencing discussion feedback.

Pre-Experiment Analyses of Variance

A one-way analysis of variance: TSIS by classroom sections was initially conducted to investigate whether there were any initial differences between sections on the TSIS discussion tendency test. Similarly the mid-term scores by classroom sections were analyzed in a one-way analysis of variance to investigate whether there were any initial section differences in the students' scores on the mid-term examination. Further one-way analyses of variance were conducted to insure that there were no initial differences between those four sections of students drawn at random to experience TV during feedback, using student TSIS scores as the criteria, and few initial differences, using mid-term examination achievement scores as the criteria. These four analyses of variance were conducted to investigate whether there were any initially significant differences among sections or between sections that had experienced discussion feedback versus those sections that had experienced TV feedback.²

² There were no significant differences among the eight sections of students or between the two treatment groupings of students on either the discussion tendency test or the mid-term examination.

The section means on the mid-term examination were

Covariance Statistical Methodology of Experiment

McNemar (16:373) states that an analysis of covariance is more sensitive to the potential significant differences of dependent variables than more ordinary techniques. Therefore, in this study, initial student mid-term examination scores were adjusted in analyses of covariance using the final examination scores as the dependent variable.

Analyses of covariance were conducted to evaluate the experimental treatment effects of students discussing or watching TV on their final examination scores.

Scores on the discussion tendency scale (the test is found in Appendix D) were classified as above or below the mean. The interactions of the levels of discussion tendency were analyzed with the treatment effects as they together might interact and influence final exam scores.

Because two of the dependent variables were short tests of ten and fourteen items, these tests also were

32.00, 33.05, 33.50, 33.30, 35.85, 33.85, 32.50 and 32.43. The treatment means on this examination were 33.15 and 33.47. The F-ratio for section differences was 1.17 which was within the .32 level of confidence. The F-ratio for treatment means was .17 which was within the .69 level of confidence.

The section means on the discussion tendency test were 14.10, 14.90, 16.60, 14.15, 13.65, 12.55, 15.61, and 13.62. The treatment means on this test were 14.37 and 14.39. The F-ratio for section differences was 1.52 which was within the .17 level of confidence. The F-ratio for treatment means was .001 which was within the .98 level of confidence. (At the .02 level of confidence it can be concluded, therefore, that there were no initial differences between treatment groups as to discussion tendency scores.)

then run under the special consideration of combining these two final examination sub-parts of repeated items from the mid-term examination to increase criterion reliability.

McNemar (16:372) also states that if uncontrolled multiple variables correlate near zero with dependent variables, they need not be controlled experimentally or statistically. Simple correlations between the uncontrolled student mid-term examination scores and final examination scores were expected to exceed a .6 correlation as has happened on previous Education 200 tests. However, breakdowns of the students on the mid-term examination as to sex will be investigated to establish correlation near zero with final examination scores and their sub-parts. Following McNemar's (16:372) reasoning, the sex variable can be eliminated from covariance analyses in this study if it contributes very little to the prediction of final examination scores and their sub-parts.

Covariance Assumptions

Edwards (6:292) emphasized that "it is important to stress that the application of the analysis of covariance does assume that the regression lines for the various treatment groups all can be assumed to have a common slope." Consequently, analyses for homogeneity of regression were conducted by assessing the extent that the regression lines for the two experimental treatments and two levels of student discussion were all four parallel. More specifically,

different analyses for homogeneity of regression were conducted with the students classified by (a) experimental treatment and (b) above or below the mean on tested level of discussion tendency. These analyses were repeated for the dependent variables of this experiment. If the assumption of homogeneous regression for these treatments and levels is tenable for the dependent variables, then following Edwards (6:292), analyses of covariance can be performed.

The uncontrolled multiple variables in this study were the mid-term examination scores of the students and the breakdown of these scores into sex differences. By using the highly sensitive covariance analyses statistic, the uncontrolled multiple variables were to be controlled statistically if there was evidence they might influence final examination scores and their sub-parts.

Attitude Analysis

The explorative attitude questionnaire was analyzed with the Mann-Whitney U statistic. Each of the thirty-one questions was compared with respect to student responses grouped by whether they had listened to television or discussed examination answers. Ties between groups on each point of agreement on the scale were adjusted in the statistical analysis by the method recommended by Siegel (22:124-125). Siegel (22:157) recommends the use of the Mann-Whitney U statistic because it is the most powerful test "of location" appropriate to large samples meeting the assumption of ordinal data.

Summary of Methodology

Analyses of covariance were conducted to evaluate the experimental treatment effects of students discussing or watching TV on their final examination scores.

Scores on the discussion tendency scale (the test is found in Appendix D) were classified as above or below the mean. The interactions of the levels of discussion tendency were analyzed with the treatment effects as they together might interact and influence final examination scores.

By using the highly sensitive covariance analyses statistic, the uncontrolled multiple variables were to be controlled statistically if there was evidence they might influence final examination scores and their sub-parts.

Mann-Whitney U tests for independence of experimental and control groups were computed for each Likert scale questionnaire item. The attitude scale is found in Appendix E. A factor analysis of the whole scale assessed the groupings of items.

CHAPTER IV

ANALYSIS OF RESULTS

In this chapter a report of the analysis of data is presented in two parts. The results of the learning experiment are presented in the first part. The results of the attitude experiment are presented in the second part.

Experiment I:

The Effect of Method of Presenting Feedback and Student Discussion Tendency on Final Exam Performance

The analysis of covariance statistic used in testing Experiment I, assumes (1) homogeneity of regression and (2) all significant variables have been controlled. It was therefore necessary to test for these assumptions prior to completing the final analyses of covariance.

Assumption of Homogeneity of Regression

As a test of homogeneity of regression, an analysis of variance for departure from homogeneity of regression was completed for each dependent variable. The results of these tests are presented in Table 4.1.

Table 4.1. Analyses of Variance for Departure from Homogeneous Regression Using Final Examination as the Dependent Variable

Total Final Examination					Items Modified on the Final Exam from the Mid-Term Exam				
Source	d.F.	S.S.	M.S.	F	Source	d.F.	S.S.	M.S.	F
Regression	4	6251.19			Regression	4	98.70		
Departure from homogeneity	3	191.78	63.93	1.18*	Departure from homogeneity	3	1.00	.33	.17*
Total regression	7	6442.97			Total regression	7	99.71		
Error	152	8242.93	54.23		Error	152	295.79	1.95	
Total		14685.90			Total		395.50		

*Not Significant

*Not Significant

Table 4.1. Continued.

Items Modified on the Final Exam from the Mid-term Exam					Combined Items From Mid-Term Exam Appear- in Exact or Modified Form on the Final Exam				
Source	d.F.	S.S.	M.S.	F	Source	d.F.	S.S.	M.S.	F
Regression	4	229.41			Regression	4	615.71		
Departure from homogeneity	3	6.80	2.27	.78*	Departure from homogeneity	3	7.87		.41*
Total regression	7	236.21			Total regression	7	623.58		
Error	152	442.73	2.91		Error	152	970.32	6.38	
Total		678.94			Total		1593.90		
*Not Significant					*Not Significant				

Since none of the observed F-ratios approached a significant level (when alpha was conservatively set at .20), it is reasonable to conclude that the regression lines are parallel and that the assumption of homogeneity of regression is therefore tenable.

Uncontrolled Variables

From Chapter III, it can be recalled that the uncontrolled variables in this study were the mid-term examination scores of the students and the break-downs of these scores by sex differences.

If sex differences correlate near zero with the final examination and its sub-parts, then it need not be controlled experimentally or statistically. Sex differences did not contribute demonstrably to the over-all variance. Multiple correlation coefficients using sex differences correlated .04 with the final examination, .07 with the ten items repeated on the final examination in exact form as on the mid-term examination, and .04 with the fourteen items modified on the final examination from the mid-term examination. Covariance analyses were, therefore, performed without using sex differences.

From Chapter III, it can be recalled that mid-term exam scores were to be statistically controlled in covariance analyses. This decision was reached because mid-term exam scores had correlated .60 with final examination scores in a pilot study prior to this investigation. During the term

when this experiment was conducted, mid-term examination scores correlated .65 with final examination scores.

Mid-term examination scores correlated .62 with the combined twenty-four items modified and repeated in exact form on the final examination from the mid-term examination. Mid-term examination scores correlated .57 with the fourteen items modified on the final examination from the mid-term examination.

However mid-term exam scores only correlated .50 with the ten items repeated in identical form on the final exam from the mid-term. Hence, corollary null hypotheses 1b and 2b (to be presented next) are to be viewed within this limitation (mentioned in Chapter I).

Hypotheses About Achievement (Experiment I)

The research hypotheses of Experiment I were stated in Chapter I. The null hypotheses of Experiment I and their alternatives were also specified in Chapter III. The research hypotheses, and their corollary null hypotheses, are stated as follows:

Hypothesis I: The performance on the final exam of students who receive televised feedback will differ from the corresponding performance of students who receive discussion feedback. The corollary hypotheses stated in null form are as follows:

Corollary Ia: The mean total final exam score of students receiving TV feedback will not differ from the corresponding mean for students receiving discussion feedback.

Corollary Ib: The mean score on repeated items for students receiving TV feedback will not differ from the corresponding mean for students receiving discussion feedback.

Corollary Ic: The mean score on modified items for students receiving TV feedback will not differ from the corresponding mean for students receiving discussion feedback.

Corollary Id: The mean combined score (modified and repeated items) for students receiving TV feedback will not differ from the corresponding mean for students receiving discussion feedback.

Hypothesis II: There will be a significant interaction between the method of presenting feedback and student discussion tendency in terms of their effect on final exam performance.¹ The corollary hypotheses stated in null form are as follows:

Corollary IIa: There will not be a significant interaction between the method of presenting feedback and student discussion tendency in terms of the effect of these two variables on total scores on the final exam.

Corollary IIb: There will not be a significant interaction between method of presenting feedback and student discussion tendency in terms of the effect of these two variables on repeated items on the final exam.

Corollary IIc: There will not be a significant interaction between method of presenting feedback and student discussion tendency in terms of the effect of these two variables on modified items on the final exam.

¹Giddan's (8) Research Form B of the Academic Behavior Inventory was used to test students' discussion tendency.

Corollary IIId: There will not be a significant interaction between method of presenting feedback and student discussion tendency in terms of the effect of these two variables on combined items (modified and repeated) on the final exam.

Final Exam Means

The means on the final examination of (1) the television and discussion groups, (2) students above and below the average on tested tendency to discuss, and (3) their interactions are presented in Table 4.2. The means are adjusted to equate students mid-term examination scores. The means are presented in Table 4.2 according to students' average scores on the:

1. total final exam,
2. ten items repeated in identical form on the final exam from the mid-term exam,
3. fourteen items modified on the final exam from the mid-term exam, and
4. combined total of twenty-four items from the mid-term exam appearing in identical or modified form from the final exam.

There appear to be two consistent relationships and a third nearly consistent relationship among the means presented in the four groupings of Table 4.2. As the total final exam is made up in part by identical and modified items from the mid-term exam, the dependent variables were not independent of one another: These relationships are:

1. Students with an above-average discussion tendency test score consistently have higher mean scores on the final exam and its sub-parts than do students with a below-average discussion tendency test score.

Table 4.2. Mean Scores on the Final Examination of Television and Discussion Groups, Students Above and Below the Average on Tested Tendency to Discuss, and Their Interactions*

Means on the Total Final Exam			Means of Ten Items Repeated in Exact Form on the Final Exam from Mid-term Exam		
<u>TV</u> <u>Discussion</u>			<u>TV</u> <u>Discussion</u>		
Above average	61.16	62.22	61.69	Above average	8.33 8.68 8.51
Below average	60.03	59.53	59.78	Below average	8.32 8.42 8.38
	60.60	60.88			8.33 8.56
Means of Fourteen Items Modified on the Final Examination from the Mid-Term Exam			Means of Combined Twenty-Four Items From the Mid-Term Exam Appearing in Identical or Modified Form on the Final Exam		
<u>TV</u> <u>Discussion</u>			<u>TV</u> <u>Discussion</u>		
Above average	10.70	10.81	10.75	Above average	19.02 19.49 19.26
Below average	10.78	10.02	10.40	Below average	19.11 18.50 18.78
	10.74	10.42			19.06 18.97

*All of the above final exam means are adjusted to equate students' mid-term exam scores.

2. Students with an above-average discussion tendency test score who discussed their mid-term exam answers consistently have the highest mean scores on the final exam and its sub-parts.
3. A less consistent relationship was that students with a below-average discussion tendency test score who discuss their mid-term examinations have the lowest mean scores on the final exam and its sub-parts than students who did not. The one exception was on the ten items repeated in the identical form on the final examination as they were on the mid-term exam.

Although none of the predicted difference and/or interactions were extensive, an analysis of covariance was nevertheless completed to determine whether or not they were statistically significant for any of the dependent variables. This analysis seemed reasonable in view of the large sample size.

Analyses of Covariance Results

The analysis of covariance test was used as a test of Hypotheses I and II. The results of this test was repeated for each of the dependent variables. The results, when the total final exam served as the dependent variable, are presented in Table 4.3.

Table 4.3. Analysis of Covariance Using the Total Final Exam as the Dependent Variable

Source	d.F.	S.S.	M.S.	F
Treatment	1	3.10	3.10	.06*
Level of Discussion	1	145.84	145.84	2.68*
Interaction	1	24.45	24.45	.45*
Error	155	8434.71	54.42	

*No Significant Differences

The F value needed for (hypothesized) significance ($P < .05$) is 2.67. Since none of the obtained ratios exceeded this value, it can be concluded that the means did not differ significantly.

The results of the analysis of covariance for the testing of Null Hypotheses I and II, using the ten items repeated on the final exam in the identical form as they appeared on the mid-term exam as the dependent variable, are presented in Table 4.4.

Table 4.4. Analysis of Covariance Using the Final Exam Items Repeated from the Mid-Term Exam as the Dependent Variable

Source	d.F.	S.S.	M.S.	F
Treatment	1	1.97	1.97	1.03*
Level of Discussion	1	.66	.66	.34*
Interaction	1	.61	.61	.32*
Error	155	296.80	1.91	

*No Significant Differences

The F value needed for significance ($P < .05$) is 2.67. Since none of the obtained ratios exceeded this value, it can be concluded that the means did not differ significantly.

The results of the analysis of covariance for the testing of Null Hypotheses I and II, using the fourteen items modified from the mid-term exam for the final exam as the dependent variable are presented in Table 4.5.

Table 4.5. Analysis of Covariance Using the Final Exam Items Modified from the Mid-Term Exam as the Dependent Variable

Source	d.F.	S.S.	M.S.	F
Treatment	1	4.04	4.04	1.39*
Level of Discussion	1	4.93	4.93	1.70*
Interaction	1	7.68	7.68	2.65*
Error	155	449.53	2.90	

*No Significant Differences

The F value needed for significance ($P < .05$) is 2.67. Since none of the obtained ratios exceeded this value, it can be concluded that the means did not differ significantly.

The results of the analysis of covariance for the testing of Null Hypotheses I and II, using the combined repeated and modified items from the mid-term exam on the final exam as the dependent variable are presented in Table 4.6.

Table 4.6. Analysis of Covariance Using the Final Exam Items Repeated and Modified from the Mid-Term Exam as the Dependent Variable

Source	d.F.	S.S.	M.S.	F
Treatment	1	.37	.37	.06*
Level of Discussion	1	9.19	9.19	1.46*
Interaction	1	12.63	12.63	2.00*
Error	155	978.19	6.31	

*No Significant Differences

The F value needed for significance ($P < .05$) is 2.67. Since none of the obtained ratios in this study exceeded this value, it can be concluded that the means did not differ significantly.

Experiment II:

Student Attitudes Toward Different Methods of Presenting Feedback

A summary of student responses to the exploratory attitude questionnaire are presented in this section of Chapter IV. The null hypothesis tested for each of the thirty-one questions was:

There are no attitude response differences between students who spend a classroom hour discussing right and wrong answers of their objective mid-term examinations and students who listen to a closed-circuit TV presentation of these results.

The Mann-Whitney U statistic was employed in determining independence of relationship because of the reasons presented in Chapter III. (Siegel (22:157) recommends the use of the Mann-Whitney U statistic because it is the most powerful and valid test "of location" appropriate to large samples meeting the assumption of ordinal data.)

The problem of objectively classifying questionnaire items was met by factor analyzing the attitude scale. All items will be reported with respect to the factor in which they had their highest loadings in Table 4.7.

Table 4.7. Differences in Student Attitude Toward Different Methods of Presenting Feedback on the Mid-Term Exam²

Item No.	Abbreviated Description	Percent of		Direc- tion	Mann- Whitney U	Probability Level
		Positive Discussion	Ratings TV			
Factor I: Understanding of the Course						
2.	Help study for final exam	41%	40%		72688	.62 N.S.
3.	Understood interrelationships course	21%	39%	+TV	59433***	.001***
6.	Better grade on final exam	24%	29%		69456	.11 N.S.
10.	Did not learn a great deal	42%	41%		73156	.76 N.S.
12.	Did not understand at own rate	43%	61%	-TV	55758***	.001***
13.	Did not think about it afterwards	74%	77%		72218	.51 N.S.
15.	Do differently if I were teaching	43%	50%	-TV	66643*	.05*
18.	Not desirable feature course	38%	43%		68967	.13 N.S.
20.	Not understand purpose teacher	48%	52%		69701	.18 N.S.
26.	Not certain of correct answers	51%	45%	+TV	67912*	.05*
27.	Purposes of course not clear	48%	40%	+TV	63947***	.001***
28.	Understand course content	35%	37%		70769	.26 N.S.
Factor II: Held Attention						
8.	Enough to hold attention	65%	78%	+TV	64100***	.001***
9.	Materials held attention	68%	63%	-TV	68209*	.05*
14.	Attention did not wander	58%	59%		73595	.92 N.S.
21.	Not too much irrelevant material	66%	77%	+TV	63142***	.001***
23.	Not boring	60%	52%	-TV	66574*	.05*
24.	Welcome feedback	86%	87%			
31.	Over-simplified	52%	59%	-TV	64821*	.05*

Table 4.7. Continued

Item No.	Abbreviated Description	Percent of		Direc- tion	Mann- Whitney U	Probability Level
		Positive Ratings Discussion	TV			
Factor III: Involved and Interacted						
22.	Didn't stimulate interaction students	22%	74%	-TV	27519***	.001***
25.	Not too formal	86%	53%	-TV	43309***	.001***
30.	Not personally involved	16%	71%	-TV	25355***	.001***
Factor IV: Clear Presentation						
1.	No difficulty seeing materials	61%	52%	-TV	63346***	.001***
4.	Wanted to ask questions	53%	60%		69456	.11 N.S.
5.	Needed time to absorb materials	60%	75%	-TV	60055***	.001***
11.	Wanted to talk with teacher	66%	66%		69443	.17 N.S.
19.	Several improvements needed	55%	68%	-TV	61989***	.001***
29.	Questions no longer seem ambiguous	72%	55%	+TV	58068***	.001***
Factor V: Noise Interfered with Instruction						
7.	Can see teacher later	80%	80%		73556	.61 N.S.
16.	Always hear what is going on	78%	69%	-TV	64508***	.001***
17.	Few disturbances during feedback	77%	79%		72984	.64 N.S.

²Items have been grouped according to results of the factor analysis, see Appendix F

*P < .05

**P < .01

***P < .001

N.S. not significant

To facilitate the interpretation of the meaning of the experimental results, some positively worded items have been changed to negative statements. In addition some negatively worded items have been changed to positive statements. Further all items will be reported in percentages of students' acceptance by either pooling "Strongly agree" and "Agree" responses or "Strongly disagree" and "Disagree" responses.³

These percentages are depicted in columns three and four of Table 4.7. In view of the magnitude of these differences, a Mann-Whitney U test was computed for each item on the questionnaire. The results of these tests are depicted in columns five and six of Table 4.7.

Significant Attitude Items

Of the thirty-one items covering attitudes about examination feedback listed in the questionnaire, eighteen items revealed statistically significant differences between groups ($P < .05$). In Table 4.7 the results are summarized for those items that were statistically significant. The percentage of student acceptance of the item is given and whether the acceptance was or was not in favor of the instructional effectiveness of TV (see footnote 3).

As indicated in Table 4.7 there were 12 significant items in favor of the discussion treatment and six

³In Appendix F items are reported with respect to exact distribution of responses on all five points of the attitude scale. Items are also reported with their exact wording as they appear in Appendix E.

significant items in favor of the television treatment with respect to student agreement. These items can be summarized as follows:

Thirty-nine per cent of the students watching television agreed that the feedback session helped them understand interrelationships among various parts of the course, but only 21% of the students discussing their answers responded as if it had helped them; 31% of the television group and 54% of the discussion group said they were not helped (item 3). However 61% of the students in the discussion treatment reported that the feedback hour helped them to understand course materials at their own rate of comprehension and only 43% of the television group acknowledged this help (item 12). One half of the students who saw TV said they would not give exam results this way if they were teaching the course; 43% of the discussion group agreed that they would not have students discuss exam results if they were teaching (item 15). Nevertheless as a result of feedback more students who discussed their exam results than those who saw them on TV still felt they were uncertain of why the correct answers were correct (51% and 45%, respectively, for item 26). Forty-eight per cent of the discussion group and 40% of the TV group agreed that the purposes of the course were not more clear as a result of the feedback hour (item 27).

Seventy-eight per cent of the TV group and 65% of the discussion group replied that there often was not enough going on to hold their attention when they received their exam

results (item 8). In contrast 68% of the TV students and 63% of the discussion students did reply that the materials they saw during feedback held their attention (item 9). But 77% of the students watching TV and 66% of the students discussing exam answers disagreed with the statement that there was too much irrelevant material covered during the experimental session (item 21). Sixty per cent of the discussion group and 52% of the TV group did not think the way they received examination feedback was boring (item 23). Whereas 52% of the students in small discussion groups thought the presentation over simplified, 59% in TV groups gave this response (item 31).

There was a consistency in response to items related to whether students said they were stimulated by other students, personally involved, and whether the feedback session seemed too formal. Students watching TV said they were not stimulated to interact with other students (74%) and not personally involved (71%) whereas only 22% and 16% of the students discussing exam answers gave these replies (items 22 and 30, respectively). Similarly 86% of the students discussing answers but only 53% of students watching TV disagreed with item 25 related to whether the feedback hour was too formal.

Thirty-four per cent of the TV group and 20% of the discussion group replied they had difficulty seeing materials; 52% of the TV students and 60% of the discussion students reported no difficulty (item 1). Three quarters of the TV

group and 60% of the discussion group said they needed more time to absorb and assimilate materials (item 5). Sixty-eight per cent of the TV students and 55% of the discussion students reported that several improvements needed to be made in the method by which they received their examination results (item 19). However 55% of the TV group and 72% of the discussion group stated that many exam questions still seemed ambiguous after feedback (item 29).

Seventy-eight per cent of the students listening to TV and 60% of the students in small group discussions reported that they could always hear what was going on during the feedback hour (item 16).

Non-Significant Attitude Items

Of the thirty-one items listed in the attitude questionnaire, the response to thirteen items revealed no statistically significant differences between groups and consequently the null hypothesis was accepted. In Table 4.7 the results are summarized for those items that were not statistically significant.

An examination of Table 4.7 shows that seven of the items related to learning and understanding the course were not significant. Further inspection of the six remaining non-significant items reveals no differences between the TV and discussion students in terms of items having their highest loadings in three different factors.

Factor Analysis

The problem of objectively classifying questionnaire items was met by factor analyzing the attitude scale. All items have been reported with respect to the factor in which they had their highest loadings. The loadings of each item within its factor are reported in Appendix F.

Three of the items which had significantly favored the discussion group loaded on one factor. The discussion group agreed that (1) the feedback session stimulated them to interact with other students, (2) they felt personally involved, and (3) the session was not formal. Items 22 and 30 both loaded highest on this factor related to whether there was interaction and involvement as a result of feedback. Item 25 related to whether students agreed that the exam presentation was too formal.

In Table 4.7 the significant differences within this factor were reported along with the percentage endorsement of students indicating that the direction of significance favored the instructional effectiveness of discussing exam answers. Discussion students agreed more than TV students with statements about whether they were stimulated to interact with each other and were personally involved during what the TV students agreed was too formal an examination feedback hour. On item 22, 74% of the TV students disagreed and 62% of the discussion students agreed. On item 25, 53% of the TV students and 86% of the discussion students agreed. On item 30, 71% of the TV students disagreed and 64% of the

discussion students agreed.

Significant differences also occurred on items involving what the students felt they had learned and understood. Three of the items with loadings on this factor suggested that students who discussed exam answers were more apt to agree that they were not helped to understand the course more than students who received the TV feedback. These items involved understanding: interrelationships and purposes of the course and why the correct answers were correct. However the reverse was true for two other items in this group. Students who discussed the exam were more apt to agree that they were understanding at their own rate of comprehension and that they would give students exam results the same way they received them if they were teaching the course.

Of the seven items related to the factor of holding students' attention, five were also significant. Three of these items favored the discussion group: Sixty-eight per cent of the discussion group agreed that the materials given them to look at during the feedback hour held their attention while only 63% of the TV group agreed. Sixty per cent of the discussion group disagreed with the statement asking them whether the feedback session was boring while only 52% of the TV group agreed. Fifty-nine per cent of the TV group agreed that feedback was oversimplified, while only 52% in the discussion group agreed. On the other hand two items in this group favored the TV group. Over three-fourths of this group agreed that during the feedback hour there was:

enough going on to hold their attention and that not too much irrelevant material was covered during the feedback hour. The corresponding percentages for the discussion group were 65% and 66%.

The only other item favoring TV loaded on a different factor: namely, clarity of presentation. Seventy-two per cent of the discussion group but only 55% of the TV group agreed that even after feedback, many questions were ambiguous. But 61% of the discussion group and only 52% of the TV group agreed that they had no difficulty seeing materials. In addition 75% of the TV group and only 60% of the discussion group said they needed time to absorb and assimilate materials. Sixty-eight per cent of the TV group and only 55% of the discussion group thought several improvements needed to be made in the method by which they received their examination results. Of the six items in this factor, four were significant and three favored discussion feedback.

The fifth factor which was noise interfering with instruction had one significant item favoring discussion feedback: Seventy-eight per cent of the discussion students and 69% of the TV students agreed that they were always able to hear what was going on. The two other items loading in this group were not significant.

It should be pointed out that to facilitate the interpretation of meaning in this study, all items have been reported in percentages of students' acceptance by either pooling "Strongly agree" and "Agree" responses or "Strongly

disagree" and "Disagree" responses. In addition some positively worded items have been changed to negative statements. Some negatively worded items have been changed to positive statements. It is to be remembered that the purpose of the attitude questionnaire was explorative and hypothesis generating.

In Appendix F all items listed in the questionnaire are summarized:

1. by factor loadings,
2. Mann-Whitney U values,
3. level of probable difference and direction of significance,
4. in percentages as well as raw scores for student agreement to each of the five alternatives on the questionnaire and
5. by the content of the item as listed in Appendix E.

Summary

An analysis of the data was presented in this chapter. Although mid-term examination scores were statistically equated through use of the analysis of covariance procedure, no significant differences were found on final exam scores and their sub-parts between groups of students discussing mid-term exam results and students watching a TV presentation of these results. Of the thirty-one items exploring attitudes about examination feedback listed in the questionnaire, eighteen items revealed significant differences

between the groups. Twelve of the items favored the instructional effectiveness of the discussion treatment and six of the items favored the television treatment with respect to student agreement on the questionnaire. A discussion of each significant item was presented. When the questionnaire was factor analyzed, three of the items which had significantly favored the discussion group loaded on one factor. The discussion group agreed that (1) the feedback session stimulated them to interact with other students, (2) they felt personally involved, and (3) the session was not formal.

CHAPTER V

FINDINGS, CONCLUSION, AND SUMMARY

The Problem

The problem of this pilot study was to offer two sets of experimental data relevant to the instructional effectiveness of presenting examination results. Two experiments were conducted to compare the effects of discussion and television feedback. The first experiment compared these effects on students' learning. The second experiment compared these effects on students' attitudes.

For statistical purposes and to assess instructional effectiveness it was hypothesized that students watching television would not differ from those discussing examination results.

Design and Procedures of the Study

The learning experiment was conducted with 166 students in the Summer Term and an attitude questionnaire was administered to 787 students in the Fall Term of 1967.

During the Summer and Fall Terms, students in different discussion sections of beginning Education 200 at MSU were presented correct answers to their mid-term examination in one of the two randomly assigned conditions:

- 1) Students were presented with correct test alternatives over closed-circuit TV with a description of why the correct alternatives were correct together with a blown-up picture of the mid-term examination. These students watched a televised team of two teachers on one tape. This was also the method of feedback used during the previous three terms in Education 200.
- 2) Students were assigned to discussion sections consisting of small groups of six students. A graduate assistant gave each group a copy of the correct answers to the test during the feedback session and asked them as a group to participate in discussing why the correct answers were correct.

Evaluation of all sections during the summer learning experiment consisted of determining the relative effectiveness of these conditions on the final examination. For separate statistical analysis and comparison, the final exam included 14 items from the mid-term examination which were modified and 10 items which were in the exact form as they appeared on the mid-term test.¹

These repeated items, selected by a panel of judges after the feedback sessions, served to assess whether students had learned the content of these items (and/or the concept related to these items) as a result of one of the two feedback conditions of the experiment.

There were four final examination performance criteria or dependent variables in this experiment:

1. Number of modified recurring items answered correctly. (These items were adequate items of the item pool in advance of the mid-term. Some of these items had equivalent statistical results from previous test administrations.)

¹The modified items tested the same concept or principle as the corresponding mid-term items but were presented under different stimulus conditions on the final examination.

2. Number of recurring items answered correctly.
(These items were again chosen at random from the mid-term.)
3. The above modified and recurring items combined.
4. The total scores on the final examination.

Independent variables included:

1. The two treatment conditions outlined above.
2. Student discussion tendency as assessed on the TSIS.
3. Test scores on the mid-term examination.

The treatment groups were found to have very similar scores as to their tested tendency to be active in classroom discussions as measured on the TSIS. Final examination scores of individual students within treatment groups were, therefore, classified as above or below the mean of students tested tendency to discuss. These levels of discussion tendency could, then, be analyzed for an interaction effect with the treatment effect of having or not having the opportunity to discuss mid-term test items.

FINAL EXAM SCORES

Grouped by

		Discussion Group	TV Group
Grouped by TSIS SCORES	Above Average		
	Below Average		

Analyses of covariance were, then, conducted to evaluate the possible interaction of tested tendency to discuss with the experimental treatment effects of students' discussing or watching TV, as these effects might interact and influence performance on the final examination.

Analyses of covariance were also conducted to evaluate the experimental treatment effects of students' discussing or watching TV on their final examination scores. By using the highly sensitive covariance analyses statistic, mid-term exam scores were to be controlled statistically as they might attribute to the variance of final examination scores.

During the Fall Quarter an exploratory thirty-one item attitude questionnaire was constructed. This instrument was based on the M.S.U. research of Davis and Johnson (4). It was designed to meet the same objectives of the Davis and Johnson instrument as well as provide answers to applied questions raised by the faculty of Educational Psychology 200 course and this investigator.

During the Fall Quarter, 787 students in this course completed the questionnaire. Responses to each item were indicated on a five point scale of agreement. The Mann-Whitney U statistic was used to determine whether or not observed differences between students in the TV and discussion groups were statistically significant. The problem of objectively classifying the questionnaire items was met by factor analyzing the attitude scale.

Results of Learning Experiment

Although mid-term examination scores were statistically equated through use of the analysis of covariance procedure, no significant interactions were found between the experimental treatments and levels of student discussion tendency on the final exams and their subscores. In addition no significant differences were found on final exam scores and their subscores between groups of students discussing mid-term exam results and students watching a TV presentation of these results. Given the obtained consistent but non significant differences reported in the last chapter, it cannot be concluded that no differences in learning occur when students receive exam results by the different methods, but can only be concluded that no significant differences were found in this investigation by the methodology used in this study.

Results of Attitude Experiment

Of the thirty-one items exploring attitudes about examination feedback listed in the questionnaire, the Mann-Whitney U statistic revealed eighteen of these items to significantly differentiate the two treatment groups. Twelve of the items favored the instructional effectiveness of the discussion treatment and six of the items favored the television treatment with respect to student

agreement on the questionnaire. The greater number of significant differences in favor of the attitudes of the discussion group over the television group implies:

1. Students might prefer discussing exam answers to having them televised, or
2. the choice of questionnaire items may not be representative of all possible relevant attitudes about exam feedback.

Attitude Questionnaire Factor Analyzed: Consistent Results

Personal Involvement and Interaction When the questionnaire was factor analyzed, three of the items which had significantly favored the discussion group loaded on one factor. This factor related to the discussion students agreeing more than the TV students that they were stimulated to interact and to be personally involved as a result of what they agreed was not too formal a feedback hour. This factor seemed relevant to the rationale of this study, namely that the discussion group was more actively involved as a result of feedback than the TV group. All the items were significant within the factor; however, the meaningfulness of this fact is offset by a number of considerations.

These differences in reported active involvement may be an attribute of the wording of the questions. Students might have been inclined to respond to questions about discussion as if it had stimulated them to personal involvement even if they had not been exposed to the discussion method used in this study. The statistical dif-

ferences may be less inherent in the method than in superficial response sets to formal televised presentations and to what students felt they should have been reporting about student-led discussion.

On the other hand, Davis and Johnson (4) concluded:

One of the most significant differences between the lecture hall and the TV classroom lies in the fact that the lecture hall allows two-way communication whereas TV does not . . . [nevertheless] questions dealing with interest and stimulation revealed no significant differences between those students who saw the lectures live and those who viewed the lectures in a television classroom.

It is a most tentative conclusion to suggest that students in the present study were personally involved and stimulated to interact moreso as a result of discussing exam answers than by watching a more formal presentation in a televised classroom. There was no attempt to assess attitudes towards exam feedback before exposure to the methods and to compare these differences after exposure. These were exploratory studies without the intent or methodological controls to experiment specifically with the factor of active student involvement.

The research reviewed in earlier chapters and the factorized results of this experiment, then, only tend to support the conclusion that students can respond as if they are more stimulated to active personal involvement when discussing exam results among themselves than when viewing them over the more formal TV media.

Other Significant Differences Within Factors

For the practical purposes of this study there were four other sets of questionnaire items factor analyzed. Individual questionnaire items loading within factors will be identified with respect to significant differences found between treatment groups.

Learning and Understanding Course Of the twelve items loading within the factor of learning and understanding the course, five were statistically significant. Three of these items favored the television group, two the discussion group:

Favoring TV were 3 items: "understanding interrelationships of the course, its purposes, and why the correct answers were correct."

Favoring discussion were 2 items: "understanding at students own rate of comprehension" and "would not offer students exam answers differently if I were teaching course."

Differences in what students learned and how they would teach the course did not reflect any clear advantage for either method used in this study. These results concur with the results of the learning experiment which reflected no statistical differences between methods.

Beyond these conclusions, there is a nagging inconsistency with the simplified generalization that the obtained attitudinal results relate solely to the personal involvement factor. All of the attitude responses simply do not correlate with this factor. There

is another educational implication found weaving through other factors.

Television has the advantage of bringing together interrelationships of the course and holding students attention as it does so. Discussion allows students to absorb materials at their own rate of comprehension.

It could be that for some students, discussion has the greater advantage but for other students, TV is more advantageous. In this study there was no measurement of those variables that might explain such heterogeneous mixtures of student personality types and cognitive needs as is possibly found within both experimental groups. Furthermore, such a mixture might possibly account for not obtaining statistical differences in the learning experiment.

Holding Attention Of the seven items having their highest loadings within the factor of holding students' attention, five were also statistically significant. But three of these items favored the discussion treatment and two the TV treatment:

Favoring TV were 2 items: there was "not too much irrelevant material covered during feed-back"; there was "enough going on to hold students' attention."

Favoring discussion were 3 items: "the feed-back session was not boring," nor was it "over-simplified"; "the feedback materials held students' attention."

Apart from these obtained differences, there were

again no over-all advantages indicated from exposure to either method. Davis and Johnson (4) concluded the "phenomenon [of attention] requires further study."

Beyond this obvious conclusion, influences in the present experiment should be considered. It is understandable students would be attentive to those aspects of feedback informing them how they have been and will be tested. Hence, students reported they were relatively more attentive to the materials given them by the test makers in the discussion condition of the experiment. In comparison the TV group reported relatively greater attention during the whole of the TV presentation when they could attend directly to what the test makers were presenting rather than to their hand-out materials. Further, students did not feel there was as much irrelevant material covered by the test makers over television as during student-led discussions. However, this difference as well as the contradictory complaints students had that TV was both more boring and over-simplified than discussion might have been inherent in the nature of student predispositions.

Clarity of Presentation and Noise Interfering with Instruction

Of the nine items related to the two factors of clarity of presentation and noise interfering with instruction, five were again statistically significant. Only one of these items favored the TV

treatment:

Clarity of presentation was composed of six items:

Favoring TV was 1 item: "as a result of feedback, exam questions no longer seemed ambiguous."

Favoring discussion were 3 items: "students reported no difficulty seeing materials," that not so many students needed time to absorb materials," nor did they agree that many "improvements were needed in the feedback session."

Noise interfering with instruction had 1 of 3 items favoring discussion: students reported "relatively less trouble hearing what was going on during feedback."

It is difficult to generalize from the statistical significance to the educational significance of items related to clarity of presentation. The advantage of TV may be in the clarification of ambiguous test questions. However, students viewing television said they needed more time to absorb the significance of these same questions. Perhaps the request for improvements in the presentation of test answers suggests there are inherent visual and audial problems in the TV media or at least that is what students are prone to report. Davis and Johnson (4) also mentioned:

students in television classrooms had difficulty seeing lecture materials . . . in some TV sections the disturbances created by inattentive students distracted others.

Only 12% of the students in both experimental conditions of the present study reported "disturbances

interfering with instruction." However, when the question was reworded, the TV group reported relatively more "trouble hearing what was going on during feedback." These results offer no clear generalization about how the television media should be improved. Basing tentative conclusions on all the factorized results, it may be the need for improvements in exam feedback relates as much to feelings about active participation and personal involvement as to actual noise interfering with instruction.

Conclusions

The following conclusions were therefore reached in this study:

1. Students discussing exam results in small student-led groups responded to a questionnaire as if they had more actively participated as a result of feedback than did students watching a televised presentation of these results. Specifically, discussion students agreed more than TV students that they had been stimulated to interact with each other and were personally involved as a result of what they agreed was not too formal a feedback hour. (This is an admittedly very tentative conclusion in view of obvious

methodological problems.)

2. No significant differences on final exam scores and various subscores of final exam performance were found between groups of students discussing mid-term exam results among themselves and students watching a TV presentation of these results.
3. No significant interactions were found among individual student discussion tendencies as measured on the Research Form B of the Academic Behavior Inventory, students' final exam scores and their subscores, and whether students experienced discussion or television feedback of their objective mid-term exam scores.

Implications for Further Research

The present study is exploratory; consequently, an issue raised by the findings is the need to carry out additional studies to verify the present results. For example, in addition to further research on the present attitude scale, there is a need to assess the effects of adding and deleting attitude items. Further, the attitude scale should be studied more fully under different conditions.

It is to be remembered from page one that Sassen-

rath and Gaverick (20) (and Curtiss and Wood, 3) found that learning was more effective when students discussed several objective classroom examinations with their teacher than if they merely viewed an authority's explanation of these answers. But in this study the effect of small group, student-led discussion of one objective examination was compared to the effect of students viewing authorities' explanation of their answers via closed-circuit television.

Specific questions needing further research relate to the methodological differences of these studies: It would seem that a longer time to discuss examination results might be helpful in demonstrating the effectiveness of discussion. In addition choosing more items and more difficult items to repeat from one test to another increases the reliability and the chance for obtained differences to be significant.

It may be that when the constructors of course tests televise exam answers, they also cue students to what to look for in subsequent course exams. This effect may offset any advantages students might feel about discussing answers among themselves. This might especially be true when students have the advantage of discussing course materials during half their class hours as in the present study.

Certainly there are many local influences in any experiment. What students learn from exam feedback and

student attitudes about exam feedback will only be clarified by a number of research studies under more controlled conditions.

It is to be emphasized that the purpose of analyzing the attitude questionnaire in this study was exploratory. The following observations might contain the basis for further study, possible extension, and/or replication of this study: When students watch television as compared to discussing examination answers, they respond as if they feel:

1. more bored with a "formal" and "over simplified" TV presentation,
2. less "personally involved" and that they are not stimulated to interact with other students,
3. they had difficulty seeing and hearing what is going on during feedback,
4. they are not understanding course materials at their rate of comprehension, and need more time to absorb and assimilate the course materials.
5. improvements need to be made in the way they received their examination results and that they would give students examination results differently if they were teaching the course, and
6. they were not attentive to materials presented them.

The following observations might also be the basis for further study of the positive benefits students will feel from listening to TV feedback: They, more than students who discuss these results, will feel as if:

7. They are now more certain of why the correct answers are correct and not feel that so many test questions are ambiguous,
8. they better understand the interrelationships of the course; that not too much irrelevant material had been presented; and that the purposes of the course are clearer,
9. enough was going on over TV to hold their attention.

Determination of cause and effect relationships on both the attitude questionnaire and the Research Form B of the Academic Behavior Inventory used in this study would prove helpful. Regular use of these scales over a period of years or periodically would be helpful in at least three ways. First, it would provide evidence of the effectiveness of varying the method of feedback. Second, it would provide an index of reliability. And third, the effect of course content, teaching staff, and student differences on a number of variables could be determined.

For these instruments to be as useful as possible, continued efforts should be made to refine the items and categories used. New items peculiar to new problems may be added to facilitate evaluation and representativeness of the population of instructional concerns.

In general, use of these scales in other academic institutions and courses both similar to and different from those in the present sample would be useful further research.

BIBLIOGRAPHY

1. Balch, John. "The Influence of the Evaluating Instrument on Students' Learning." Am. Ed. Res. J., Vol I (3), May 1964, pp. 169-182.
2. Bloom, B. S. "Thought Processes in Lectures and Discussion," Journal of General Education, 1953, 7, pp. 160-169.
3. Curtiss, F. D., and Wood, G. G. "A Study of the Relative Teaching Value of Four Common Practices in Correcting Examination Papers," School Review, 1929, 37, pp. 615-623.
4. Davis, Robert H., and Johnson, F. Craig. Evaluation of Regular Classroom Lectures Distributed by CCTV to Campus Classrooms. Project Report No. 202, Educational Development Program, May 1966.
5. Dennis, L. E., and Kauffman, J. F. (eds.). The College and the Student. Washington, D. C.: American Council on Education, 1966, pp. 76-100.
6. Edwards, Allen L. Experimental Design in Psychological Research. New York-Chicago-San Francisco-Toronto-London: Holt, Rinehart and Winston, July 1963.
7. Gage, Nathaniel L. (ed.). Handbook of Research on Teaching. Chicago: Rand McNally and Co., 1963.
8. Giddan, Norman S., Lovell, Victor R., Haimson, Alvin I., and Hatton, John M. A Scale to Measure Teacher-Student Interaction. Stanford University Counseling Center, Unpublished paper.
9. Krauskopf, C. J. The Use of Written Responses in the Stimulated Recall Method. Unpublished doctoral dissertation, Ohio State University, 1960. (Dissertation Abstract, 1960, 21, 1953).
10. Korn, H. A., and Black, J. D. Factors Influencing Student Participation in Classroom Discussion. Unpublished paper. Stanford Counseling Center, 1962.
11. McGeoch, J. A., and Irion, A. L. The Psychology of Human Learning. New York: Longmans, Green, 1952, pp. 46.

12. McKeachie, Wilbert J. New Developments in Teaching. Unpublished paper. Prepared for Higher Education Literature Search Project, U. S. Office of Education under Contract to Duke University.
13. McKeachie, W. J. Procedures and Techniques of Teaching: A Survey of Experimental Studies. In N. Sanford (ed.) The American College. New York: Wiley and Sons, 1962.
14. McKeachie, W. J. "Anxiety in the College Classroom," Journal of Educational Research, 1951, 45, pp. 153-160.
15. McKeachie, W. J. Teaching Tips. Ann Arbor, Michigan: The George Wahr Publishing Company, 1960.
16. McNemar, Quinn. Psychological Statistics. New York-London-Sydney: John Wiley and Sons, Inc., 1962, pp. 451.
17. Pace, C. R. Perspectives on the Student and his College. In L. E. Dennis and J. F. Kauffman, The College and the Student, Washington, D. C.: American Council on Education, 1966, pp. 76-100.
18. Stern, G., Stein, M., and Bloom, B. Methods in Personality Assessment. Glencoe, Illinois: Free Press, 1956.
19. Thistlethwaite, D. L. College Press and Changes in Study Plans of Talented Students. Evanston, Ill.: National Merit Scholarship Corp., 1960.
20. Sassenrath, Julius M., and Garverick, Charles M. Effects of Differential Feedback from Examinations on Retention and Transfer. J. of Educational Psy., 1965, 56, 5, pp. 259-263.
21. Sanford, Nevitt, (ed.). The American College. New York: Wiley and Sons, 1962.
22. Seigel, Sidney. Non-Parametric Statistics. New York-Toronto-London: McGraw-Hill Book Company, Inc., 1956.

APPENDIX A

The strong influence of the evaluating procedures on students' learning can be summarized from this investigator's (1) 1964 research with respect to the following groups of findings:

- 1) The influence of the evaluating instrument (on students' learning).
- 2) The influence of essay, multiple choice, and other types of tests (on students' learning).
- 3) The influence of the complexity of the evaluating instrument.
- 4) Frequency of testing.
- 5) Time between studying and testing.
- 6) Knowledge of results.
- 7) Interaction of teaching methods and student characteristics.

The consensus of the studies over the past sixty years supports the conclusion that the evaluating procedures do influence students' learning. Learning is affected both by the external elements of the situation, including the amount, organization, complexity, and meaningfulness of the material to be learned, and by internal factors characteristic of the learner.

APPENDIX B

Instructions to Teaching Assistants

Pre-Tested on a Trial Mid-Term This Summer

TV Feedback Condition

Please hand out Review Formats and Answer Sheets to all students. Please read aloud:

The test feedback will be on closed-circuit TV. You have been given a Review Format with instructions not to take notes. You may take the Review Format home with you, but please return all other materials at the end of the TV feedback.

(Coincidentally, please do not review course materials preceding the mid-term after this date. Please take notes during the feedback session as to your estimate of a) percentage of students attentive to the feedback; b) how attentive you think they are. Please do not permit your biases about method of feedback or those of your students to be discussed in your classroom.)

Discussion Condition

Please "randomly" assign students to six-man groups, hand out Review Formats, Tests, and Answer Sheets. Please give one group member in each group a keyed copy of the mid-term, with instructions as follows:

You are to work as a group discussing alternatives to test items. These answers will be confirmed by the group member with the keyed test. You may take the Review Format home but please return all other materials at the end of the feedback session.

(Coincidentally, please do not review course materials proceeding the mid-term after this date. Please do not permit your biases or those of your student to be discussed in your classroom. Please take notes during the feedback session as to your estimate of a) percentage of students discussing; b) being attentive; and c) how attentive these students are during feedback.)

APPENDIX C

Instructions to Subjects

TV Feedback Condition

"The purpose of test feedback is to help you to clarify in your own mind the correct alternatives to test items. Some of the items on the mid-term will reappear on the final."

"You have been given a memoeed Review Format which you may take home to aid your recognition of test items. Please do not take any notes during feedback."

"The test feedback will be presented on closed-circuit TV, and will cover all items on the mid-term. PLEASE DO NOT MARK YOUR IBM ANSWER SHEET."

Discussion Condition (Pre-tested on a trial mid-term this summer)

"The purpose of test feedback is to help you to clarify in your own mind the correct alternatives to test items. Some of the items on the mid-term will reappear on the final."

"You will form six-man groups in your sections to discuss items. Each group member will be given a Review Format which you may take home to aid your recognition of test items. Feel free to comment about the test to each other in your group but do not take any notes this hour. You are being asked to systematically discuss each question permitting each

member of your group the opportunity to participate actively in the learning process. You will want to read each question on the test given to you before asking the member of your group with the correct answer to read it to you. Please do not mark your IBM answer sheet."

APPENDIX D

TSIS

Answer either 1 (true) or 2 (false) on the answer sheet.

1. I often talk to the teacher after a class.
2. I ask a lot of questions in class.
3. I discuss my test results personally with an instructor.
4. I always raise my hand in class when I know the answer to a question.
5. I am able to communicate more easily with a teacher in the privacy of his office than in the classroom.
6. I don't like to be called on by a teacher in class unless I have raised my hand or otherwise indicated I wish to speak.
7. I like competitive intellectual discussions with my teachers.
8. If a teacher critically evaluates something I've said in class, I tend to clam up.
9. When a teacher criticizes something I've said, it stimulates me to defend my ideas or logical position.
10. I much prefer classes where a teacher lectures and only infrequently encourages questions from the students.
11. I seldom talk in class.
12. I usually assume that teachers know their field of endeavor and hence seldom discuss or argue with them about interpretations, test results, etc.
13. I'd rather read more about something that interests me than have an intellectual discussion with a teacher.

14. I talk in class only when I'm absolutely sure of my facts and ideas.
15. I have a great deal of difficulty talking in classes where I know the instructor is carefully evaluating what I have to say in order to "grade" my discussion participation.
16. I like to mingle with teachers and other faculty at receptions, teas, etc.
17. I like to get invited to a professor's house so that I can get to know him better.
18. I would never think of inviting a teacher or professor to a party at my house.
19. When something I don't understand arises in class, I'd rather discuss it with a fellow student than the teacher.
20. When my classroom participation is not being graded, I talk much less.
21. I don't enjoy classes where the teacher has little time for conversation with the students.
22. I would like to be a college professor.
23. I much prefer classes where the teacher really talks with students, not just at them.
24. I tend to shy away from talking in classes where the teacher stirs up heated and intense discussions.

APPENDIX E

Course Questionnaire

Fall Quarter 1967

By answering this questionnaire you will be helping the University in a very important way to evaluate different methods of instruction. This evaluation will not be made until after your instructor has turned in your final grade, and your answers to these questions will not affect your grade.

Instructions

Fill in the top three lines on the answer sheet supplied. Print the information with the soft lead pencil given to you. In the box at the right top of your answer sheet, fill in the correct blanks for your student number.

Respond to the following statements by blackening the space on the answer sheet which, according to the key below, best describes your reaction to the statement.

- KEY: 1 - Strongly agree
2 - Agree
3 - Uncertain
4 - Disagree
5 - Strongly disagree

1. I had no difficulty seeing what was presented to me during the hour I received my examination results.
2. As a result of the feedback session I will be able to study more effectively for the final examination.
3. The feedback session helped me understand the interrelationships among various parts of the course.
4. I felt the need to ask questions which were not answered during the hour when I received my examination results.
5. I frequently needed more time to absorb or assimilate material presented during the feedback session.
6. As a result of the feedback session, I feel I can achieve a higher grade on the final examination than I did on the mid-term examination.
7. I feel perfectly free to discuss questions stemming from the mid-term examination with my instructor.
8. Often there wasn't enough going on during the classroom hour when I received my examination results to hold my attention.
9. The materials given to me to look at during this feedback hour held my attention.
10. I learned a great deal during this feedback hour.
11. I would have understood the material better if it had been possible to discuss test questions with my instructor during the feedback hour.
12. This feedback hour helped me to understand course materials at my own rate of comprehension.
13. The feedback session was so stimulating that I thought about the subject matter which was covered a great deal after the session was over.
14. I found my attention wandering frequently during the feedback hour.
15. If I were teaching this course I would give students examination results the same way I received them.
16. I could always hear what was going on during the feedback hour.
17. It was frequently difficult to follow the feedback presentation because of disturbances in the room.

18. A desirable feature of this course is the way examination results are given to students.
19. Several improvements need to be made in the method I received my examination feedback.
20. As a result of the feedback session I better understand what my instructor was trying to accomplish in my discussion section.
21. Too much irrelevant material was covered during the examination feedback session.
22. The feedback session stimulated my interaction with other students.
23. The way I received examination feedback was boring.
24. I welcome the opportunity to have examination feedback.
25. The feedback session seemed too formal.
26. As a result of feedback I felt certain of why the correct answers were correct.
27. The purposes of the course were more clear to me as a result of the feedback session.
28. I understand the content of the course better as a result of the examination feedback.
29. Even after feedback, I felt that many items on the mid-term examination were ambiguous (more than one correct answer).
30. I felt personally involved during the feedback session.
31. The examination feedback presentation was over simplified.

APPENDIX F

Factors, Factor Loadings, Mann-Whitney U Values, and Frequency of Agreement
on Questionnaire Responses of TV and Discussion Groups

Factor I: Understanding of the Course

Item 3: Factor loading +.658

Direction of Significance +TV

The feedback session helped me understand the interrelationships among various parts of the course.

Scale	Strongly Agree		Agree		Uncertain		Disagree		Strongly Disagree		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
TV	17	4.2	143	35	122	29.8	99	24.2	28	6.9	409	100
Discussion	4	1.1	75	20.7	124	34.3	130	35.9	29	8	362	100

U 59433 Significant at .001 level

Item 12: Factor loading +.458

Direction of Significance -TV

This feedback hour helped me to understand course materials at my own rate of comprehension.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	7 1.7	62 15.2	89 21.8	189 46.2	62 15.2	409 100
Discussion	13 3.6	107 29.6	87 24.1	131 36.3	23 6.4	361 100

U 55758

Significant at .001 level

91

Item 15: Factor loading +.456

Direction of Significance -TV

If I were teaching this course I would give students examination results the same way I received them.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	24 5.9	62 15.2	119 29.1	127 31.1	77 18.8	409 100
Discussion	16 4.5	86 24	104 29	99 27.6	54 15	359 100

U 66643

Significant at .05 level

Item 26: Factor loading +.505

Direction of Significance +TV

As a result of feedback I felt certain of why the correct answers were correct.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	22 5.4	113 27.6	88 21.5	117 28.6	69 16.9	409 100
Discussion	15 4.2	85 23.6	76 21.1	108 29.9	77 21.3	361 100

U 67912

Significant at .05 level

92

Item 27: Factor loading +.719

Direction of Significance +TV

The purposes of the course were more clear to me as a result of the feedback session.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	8 2	109 26.7	129 31.5	134 32.8	29 7.1	409 100
Discussion	3 1	61 17	124 34.4	135 37.4	38 10.5	361 100

U 63947

Significant at .001 level

Item 2: Factor loading +.714

As a result of the feedback session I will be able to study more effectively for the final examination.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	32 7.8	132 32.1	145 35.3	72 17.5	30 7.3	411 100
Discussion	19 5.2	129 35.5	111 30.6	76 20.9	28 7.7	363 100

U 72688

Level of Probability .62

93

Item 6: Factor loading +.720

As a result of the feedback session, I feel I can achieve a higher grade on the final examination than I did on the mid-term examination.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	23 5.6	98 23.9	193 47.1	75 18.3	21 5.1	410 100
Discussion	18 4.9	72 19.7	177 48.5	74 20.3	24 6.6	365 100

U 70163

Level of Probability .11



Item 10: Factor loading +.725

I learned a great deal during this feedback hour.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	20 4.9	83 20.3	138 33.7	121 29.6	47 11.5	409 100
Discussion	9 2.5	78 21.6	121 33.4	121 33.4	33 9.1	362 100

U 73156 Level of Probability .76

94

Item 13: Factor loading +.562

The feedback session was so stimulating that I thought about the subject matter which was covered a great deal after the session was over.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	5 1.2	41 10.1	42 10.3	211 51.7	109 26.7	408 100
Discussion	3 1	39 10.7	52 14.3	171 47.1	98 27	363 100

U 72218 Level of Probability .51

Item 18: Factor loading +.533

A desirable feature of this course is the way examination results are given to students.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	17 4.2	95 23.3	119 29.1	121 29.6	57 13.9	409 100
Discussion	11 3.1	104 28.9	108 30	92 25.6	45 12.5	360 100

U 68967

Level of Probability .13

95

Item 20: Factor loading +.585

As a result of the feedback session I better understand what my instructor was trying to accomplish in my discussion section.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	6 1.5	66 16.1	123 30.1	143 35	71 17.4	409 100
Discussion	8 2.2	60 16.6	121 33.5	121 33.5	51 14.1	361 100

U 69701

Level of Probability .18

Item 28: Factor loading +.748

I understand the content of the course better as a result of the examination feedback.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	15 3.7	139 34	105 25.7	124 30.3	26 6.4	409 100
Discussion	2 1	124 34.4	100 27.8	104 28.9	30 8.3	360 100

U 70769

Level of Probability .26

96

Factor II: Held Attention

Item 8: Factor loading -.748

Direction of Significance +TV

Often there wasn't enough going on during the classroom hour when I received my examination results to hold my attention.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	10 2.4	32 7.8	47 11.5	239 58.3	82 20	410 100
Discussion	20 5.5	56 15.4	45 12.4	185 51	57 15.7	363 100

U 64109

Significant at .001 level

Item 9: Factor loading +.593

Direction of Significance -TV

The materials given to me to look at during this feedback hour held my attention.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	44 10.8	215 52.6	49 12	80 19.6	21 5.1	409 100
Discussion	50 13.8	197 54.4	49 13.5	48 16	8 2.2	362 100

U 68209

Significant at .05 level

Item 21: Factor loading -.556

Direction of Significance +TV

Too much irrelevant material was covered during the examination feedback session.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	3 1	18 4.4	73 17.9	245 60	70 17.1	409 100
Discussion	19 5.2	39 10.7	63 17.4	199 54.8	43 11.9	363 100

U 63142

Significant at .001 level

Item 23: Factor loading -.620 Direction of Significance -TV
 The way I received examination feedback was boring.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	35 8.6	90 22	69 16.9	193 47.2	22 5.4	409 100
Discussion	18 5	73 20.2	54 15	185 51.1	32 8.8	362 100

U 66374 Significant at .05 level

98

Item 31: Factor loading -.445 Direction of Significance -TV
 The examination feedback presentation was over simplified.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	15 3.7	55 13.7	96 23.9	199 49.5	37 9.2	402 100
Discussion	13 3.7	60 17	97 27.5	164 46.5	19 5.4	353 100

U 64821 Significant at .05 level

Item 14: Factor loading -.770

I found my attention wandering frequently during the feedback hour.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	27 6.6	100 24.5	41 10.1	198 48.5	42 10.3	408 100
Discussion	21 5.8	84 23.2	48 13.3	176 48.6	33 9.1	362 100

U 73595 Level of Probability .92

Item 24: Factor loading +.391

I welcome the opportunity to have examination feedback.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	174 42.7	180 44.1	36 8.8	15 3.7	3 1	408 100
Discussion	167 46.3	146 40.4	27 7.5	14 3.9	7 1.9	361 100

U 71250 Level of Probability .43

1

2

3

4

5

6

7

8

9

10

11

12

Factor III: Involved and Interacted

Item 22: Factor loading **+ .787** Direction of Significance **-TV**

The feedback session stimulated my interaction with other students.

Scale	Strongly Agree $\frac{\text{N}}{\text{N}} \%$	Agree $\frac{\text{N}}{\text{N}} \%$	Uncertain $\frac{\text{N}}{\text{N}} \%$	Disagree $\frac{\text{N}}{\text{N}} \%$	Strongly Disagree $\frac{\text{N}}{\text{N}} \%$	Total $\frac{\text{N}}{\text{N}} \%$
TV	4 1	39 9.6	63 15.4	232 56.9	70 17.2	408 100
Discussion	38 10.5	182 50.1	63 17.4	75 20.7	5 1.4	363 100

U 27519 Significant at .001 level

100

Item 25: Factor loading **- .581** Direction of Significance **-TV**

The feedback session seemed too formal.

Scale	Strongly Agree $\frac{\text{N}}{\text{N}} \%$	Agree $\frac{\text{N}}{\text{N}} \%$	Uncertain $\frac{\text{N}}{\text{N}} \%$	Disagree $\frac{\text{N}}{\text{N}} \%$	Strongly Disagree $\frac{\text{N}}{\text{N}} \%$	Total $\frac{\text{N}}{\text{N}} \%$
TV	19 4.7	87 21.3	87 21.3	198 48.5	17 4.2	408 100
Discussion	8 2.2	8 2.2	33 9.1	245 67.7	68 18.8	362 100

U 43309 Significant at .001 level

Item 30: Factor loading +.797 Direction of Significance -TV

I felt personally involved during the feedback session.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	5 1.2	54 13.3	58 14.3	192 47.2	98 24.1	407 100
Discussion	44 12.2	188 52.1	70 19.4	54 15	5 1.4	361 100

U 25355 Significant at .001 level

101

Factor IV: Clear Presentation

Item 1: Factor loading +.321 Direction of Significance -TV

I had no difficulty seeing what was presented to me during the hour I received my examination results.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	53 13.0	159 39.1	54 13.3	92 22.6	49 12.0	407 100
Discussion	58 16.1	160 44.4	68 18.9	56 15.6	18 5.0	360 100

U 63346 Significant at .001 level

Item 5: Factor loading -.481

Direction of Significance -TV

I frequently needed more time to absorb or assimilate material presented during the feedback session.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	142 34.6	165 40.2	30 7.3	65 15.9	8 2	410 100
Discussion	82 22.5	135 37.1	35 9.6	101 27.8	11 3.0	364 100

U 60055 Significant at .001 level

102

Item 19: Factor loading -.472

Direction of Significance -TV

Several improvements need to be made in the method I received my examination feedback.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	95 23.2	182 44.5	63 15.4	63 15.4	6 1.5	409 100
Discussion	51 14.2	148 41.1	81 22.5	72 20	8 2.2	360 100

U 61989 Significant at .001 level

Item 29: Factor loading -.514

Direction of Significance +TV

Even after feedback, I felt that many items on the mid-term examination were ambiguous (more than one correct answer).

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	97 23.8	127 31.1	67 16.4	99 24.3	18 4.4	408 100
Discussion	35 37.4	125 34.6	39 10.8	54 15	8 2.2	361 100

U 58068 Significant at .001 level

103

Item 4: Factor loading -.702

I felt the need to ask questions which were not answered during the hour when I received my examination results.

Scale	Strongly Agree $\frac{N}{\%}$	Agree $\frac{N}{\%}$	Uncertain $\frac{N}{\%}$	Disagree $\frac{N}{\%}$	Strongly Disagree $\frac{N}{\%}$	Total $\frac{N}{\%}$
TV	83 20.2	165 40.2	40 9.8	111 27.1	11 2.7	410 100
Discussion	70 19.3	123 34	40 11.1	118 32.6	11 3.0	362 100

U 69456 Level of Probability .11

Item 11: Factor loading -.738

I would have understood the material better if it had been possible to discuss test questions with my instructor during the feedback hour.

Scale	Strongly Agree $\frac{\text{N}}{\%}$	Agree $\frac{\text{N}}{\%}$	Uncertain $\frac{\text{N}}{\%}$	Disagree $\frac{\text{N}}{\%}$	Strongly Disagree $\frac{\text{N}}{\%}$	Total $\frac{\text{N}}{\%}$
TV	129 31.5	141 34.5	70 17.1	63 15.4	6 1.5	409 100
Discussion	80 22.3	157 43.7	69 19.2	48 13.4	5 1.4	359 100

U 69443 Level of Probability .17

Factor V: Noise Interfered with Instruction

Item 16: Factor loading +.715 Direction of Significance -TV

I could always hear what was going on during the feedback hour.

Scale	Strongly Agree $\frac{\text{N}}{\%}$	Agree $\frac{\text{N}}{\%}$	Uncertain $\frac{\text{N}}{\%}$	Disagree $\frac{\text{N}}{\%}$	Strongly Disagree $\frac{\text{N}}{\%}$	Total $\frac{\text{N}}{\%}$
TV	65 16	218 53.6	22 5.4	84 20.6	18 4.4	407 100
Discussion	75 20.8	208 57.6	24 6.7	46 12.7	8 2.2	361 100

U 64508 Significant at .001 level

Item 7: Factor loading +.453

I feel perfectly free to discuss questions stemming from the mid-term examination with my instructor.

Scale	Strongly Agree $\frac{\text{N}}{\text{N}} \%$	Agree $\frac{\text{N}}{\text{N}} \%$	Uncertain $\frac{\text{N}}{\text{N}} \%$	Disagree $\frac{\text{N}}{\text{N}} \%$	Strongly Disagree $\frac{\text{N}}{\text{N}} \%$	Total $\frac{\text{N}}{\text{N}} \%$
TV	139 33.8	191 46.5	50 12.2	25 6.1	6 1.5	411 100
Discussion	113 31	182 49.9	45 12.3	16 4.4	9 2.5	365 100

U 73556 Level of Probability .61

105

Item 17: Factor loading -.745

It was frequently difficult to follow the feedback presentation because of disturbances in the room.

Scale	Strongly Agree $\frac{\text{N}}{\text{N}} \%$	Agree $\frac{\text{N}}{\text{N}} \%$	Uncertain $\frac{\text{N}}{\text{N}} \%$	Disagree $\frac{\text{N}}{\text{N}} \%$	Strongly Disagree $\frac{\text{N}}{\text{N}} \%$	Total $\frac{\text{N}}{\text{N}} \%$
TV	7 1.7	46 11.3	33 8.1	262 64.1	61 14.9	409 100
Discussion	7 1.9	37 10.2	38 10.5	229 63.1	52 14.3	363 100

U 72984 Level of Probability .64

APPENDIX G

Rotated Factor Loadings

Item No.	Factor No. 1	Factor No. 2	Factor No. 3	Factor No. 4	Factor No. 5
1	0.2572	-0.721	0.2366	0.3209	0.3077
2	0.7139	0.1642	0.0630	0.1282	-0.0044
3	0.6582	0.1129	-0.1196	0.1605	0.1759
4	-0.0815	-0.0753	-0.0511	0.7024	-0.0282
5	-0.1077	0.2576	-0.1858	-0.4808	-0.2686
6	0.7191	0.0434	0.0676	0.0053	0.0105
7	0.1399	0.1511	0.0541	-0.0172	0.4531
8	-0.0979	-0.7480	0.0929	0.0411	-0.1780
9	0.2745	0.5930	0.2205	-0.0066	0.0069
10	0.7245	0.2934	0.1366	0.1318	0.0512
11	-0.0804	-0.0479	-0.1064	-0.7377	0.0900
12	0.4579	-0.0063	0.4076	0.3361	0.1148
13	0.5618	0.3979	0.1890	-0.0502	-0.0073
14	-0.1804	-0.7704	-0.1084	-0.0317	-0.0454
15	0.4564	0.2026	0.3125	0.4061	0.0939
16	0.0682	0.0541	0.1627	0.0867	0.7146
17	-0.0240	-0.1516	0.0288	-0.0445	-0.7452
18	0.5329	0.2124	0.3084	0.3035	0.0975
19	-0.3605	-0.2129	-0.3018	-0.4723	-0.1332

Item No.	Factor No. 1	Factor No. 2	Factor No. 3	Factor No. 4	Factor No. 5
20	0.5847	0.0844	0.1771	-0.0279	0.1971
21	-0.1107	-0.5560	0.1502	-0.2537	-0.1858
22	0.1566	0.0628	0.7865	-0.0675	0.0289
23	-0.2714	-0.6199	-0.3164	-0.1886	-0.1095
24	0.1373	0.3909	0.1426	-0.0365	0.0094
25	0.0405	-0.1689	-0.5809	-0.2694	-0.1847
26	0.5052	0.1184	0.0188	0.5009	0.0489
27	0.7187	0.1568	-0.0221	0.1291	0.0411
28	0.7480	0.1989	0.0385	0.1734	0.0727
29	-0.1027	-0.2361	0.2721	-0.5138	-0.0416
30	0.1286	0.1632	0.7971	0.0470	0.0470
31	-0.0890	-0.4452	-0.0445	-0.2503	-0.0540

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03061 5862