7908839

SPENCE, JAMES MELVIN
EFFECTS OF TEACHING STYLE, LECTURE CONTENT AND STUDENT ACADEMIC MAJOR UPON FACULTY EVALUATIONS.

THE UNIVERSITY OF OKLAHOMA, PH.D., 1978

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

Effects of Teaching Style, Lecture Content and Student Academic Major upon Faculty Evaluations

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

by

JAMES M. SPENCE

Norman, Oklahoma

1978

Effects of Teaching Style, Lecture

Content and Student Academic

Major upon Faculty

Evaluations

APPROVED BY

DISSERTATION COMMITTEE

Abstract

EFFECTS OF TEACHING STYLE, LECTURE CONTENT AND STUDENT ACADEMIC MAJOR UPON FACULTY EVALUATIONS

BY: JAMES M. SPENCE

MAJOR PROFESSOR: WILLIAM H. GRAVES, JR., Ph. D.

This study was designed to investigate the impact of lecture content, teaching form and student academic major upon teacher evaluations and student achievement.

Sixty-two students viewed videotaped lectures which varied in degree of content (high, low) and teaching form (good, poor). Each of four groups viewed a different lecture. Each subject completed a 26-item achievement test, an 18-item teacher evaluation questionnaire and indicated an academic major.

Students who viewed the high content lectures obtained higher achievement scores than students who viewed the low content lectures. Achievement scores were also higher for students who viewed the good teaching than for those who viewed the poor teaching. Lecture content and teaching form interacted to affect lecturer evaluation. Although higher evaluations were given to the lecturer demonstrating good teaching form, the evaluations were not sensitive to differences in lecture content in that category. Differences in

content coverage were reflected by the evaluations given to the poor form lecturer; higher evaluations were given by students in the low content group.

Acknowledgements

The author wishes to express his appreciation to those individuals who have contributed to the completion of this dissertation.

I am especially indebted to the chairman of my doctoral committee, Dr. William H. Graves, for his direction and encouragement during the writing of this dissertation. His investment of time and interest in this study is greatly appreciated.

Appreciation is also expressed to the other members of my doctoral committee, Dr. Larry E. Toothaker, Dr. Michael Langenbach and Dr. George A. Letchworth, for their comments and constructive criticisms.

To my wife, Linda, goes my love and appreciation for her support and understanding during my graduate program.

This dissertation is dedicated to Jimmy, Andy and Blake.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

In recent years there has been an increasing demand to formalize student evaluation of university faculty. This demand has come from the public, school administrators and the student "consumer" of the teacher's wares. The interest of students in the evaluation of their teachers was compounded by the activism of the 1960's (see Berns, 1967).

Some faculty members believe that their students do not clearly distinguish between the teacher and the course content. In addition, some teachers feel students do not understand the role that "form" or style plays in teaching. Some teachers also feel that a student's academic major, a variable over which the teacher has no control, also affects the evaluations they receive from their students.

Good student evaluations are the result of more than just good teaching. There are basically three factors which may influence a student's rating of a teacher's effectiveness: (a) teacher characteristics, (b) student characteristics and (c) class characteristics. Many faculty members argue that since they have little or no control over personal and environmental variables, such as age of faculty, academic rank, student's academic major and class size, and since these factors do influence student ratings, then student evaluations do not adequately reflect the faculty's contribution to the learning process. The evaluations will either be inflated or deflated depending upon the pattern of the various factors. Before this argument can stand as a viable criticism, it must first be determined whether or not the factors do, in fact, influence student evaluations.

A number of teacher characteristics including age, sex, academic rank, academic degree attained, experience and teaching style have been investigated. Heilman and Armentrout (1936) found no significant differences between student ratings and teaching experience, age, or sex. This study, conducted at the Colorado State College of Education, involved 46 teachers and 50 classes. The lack of a relationship of student rating to sex of teacher has also been reported by others (Remmers, 1930; Remmers & Elliot, 1949). Downie (1952) reported no difference between the ratings of 237 faculty members under age forty and 169 aged forty and over. On the other hand, Lasher and Vogt (1974) collected data at Bowling Green State University over six quarters from 120 faculty members and 1,072 courses. They reported significant relationships between student evaluations of teaching effectiveness and age, faculty rank, academic degree and experience. At Brooklyn College, Riley, Ryan and Lifshitz (1950) found, for the most part, inverse relationships between student ratings and age, academic rank and academic degree. The elimination of age as an influence altered the relationships pertaining to rank and degree only slightly. Rayder (1968) reported "older, more educated and more experienced instructors were rated lower than younger and less experienced instructors" (p. 80). Other researchers also have found significant relationships existing between student ratings and several of the teacher characteristics, but they have identified the relationships as positive. Downie (1952) reported full professors tended to receive higher student evaluations than did other ranks. Clark and Keller (1954) and Gage (1961) indicated that associate professors and full professors were rated higher than were instructors or assistant professors. Centra (1973) demonstrated that "students rated women more favorably than men teachers" (p. 13).

In a study of lecture style, Naftulin, Ware and Donnelly (1973) hired an actor to portray an authority on the application of mathematics of human behavior. The actor, who knew nothing of the topic, was instructed to teach "charismatically and nonsubstantively" and to respond to questions with "double talk, neologisms, non sequiturs and contradictory statements." The subjects in this study, who were psychiatrists, psychologists and social workers, made significantly more favorable than unfavorable responses to a satisfaction questionnaire. These authors have demonstrated that students can be "seduced" into believing that they have learned, and consequently, giving a teacher a high satisfaction rating. A weakness of this study is the lack of a measure of achievement. Researchers have dealt with this weakness in their studies of the effect of teaching style on student evaluations. These investigators found that seductive behavior (friendliness, expressiveness, enthusiasm, humor, etc.) in a lecture presentation had a greater influence on student ratings and student achievement than did lecture content (Ware & Williams, 1975). In another study, they reported that even after the students experienced a second lecture of the same style and content, ratings of instruction still did not accurately reflect lecture content and student performance (Williams & Ware, 1977). In a third study, Williams and Ware (1976) manipulated expressiveness, content and incentive. Their results from the manipulation of expressiveness were not in total agreement with their earlier findings; student achievement was not affected by expressive manipulation. Their data were, however, congruent with their previous research relative to student ratings; that is, student ratings were not sensitive to differences in lecture content under the high-expressiveness conditions, but under the low-expressiveness conditions they indicated variations in lecture content.

Most student characteristics make little difference in the evaluations that teachers receive from them. A number of studies found no significant differences in the rating of a teacher by male and female students (Goodhartz, 1948; Rayder, 1968; Remmers, 1930; Remmers & Elliot, 1949). Rayder (1968) reported no substantial relationship of student's age, college class or GPA. Goodhartz (1948) reported similar results; he concluded "there is no conclusive evidence for believing that the ratings given to an instructor are affected ... by such factors as the student's sex or college class" (p. 348). The research on the effect of the student's college class on the ratings of teachers is not conclusive, however. Downie (1952) and Rosenshine, Cohen and Furst (1973) found that upper division students tend to rate teachers more favorably than do lower division students. Also, graduate students are more inclined to give higher ratings than are undergraduates (Remmers & Elliot, 1949).

The effect of the student's academic major upon teacher evaluation has received only minimal attention. An incidental finding in a study by Centra (1973) was that natural science courses as opposed to courses in education, humanities and social sciences were viewed by students "as having a faster pace, as being more difficult and as being less likely to stimulate student interest" (p. 13). A more substantive study was done by Null and Walter (1972) at Purdue University. "Major field of study" was one of ten variables they were interested in studying. They asked 192 students enrolled in an introductory biology course to evaluate ten behavioral characteristics of their teacher. They found that academic major did not have a significant impact on any of the teacher characteristics evaluated; patterns were revealed, however, by an examination of mean scores. Biology majors (89 of the 192 students) preceived nine of the ten dimensions of instructor behavior somewhat better than non-biology majors. In another study involving academic major as a comparison variable, Rayder (1968) reported that "student ratings of instructors were not substantially related to student's ... major area ... " (p. 78). In view of the cited research, very little can be said, with any degree of certainty, concerning the role that a student's academic major plays in how he will evaluate his instructor's teaching effectiveness.

The relationship between grades, expected or received, and teacher evaluations is not clear, even though it has been extensively researched. Many studies have found no relationship between these variables (Bendig, 1953; Blum, 1936; Clark & Keller, 1954; Holmes, 1972; Rayder, 1968; Remmers, 1930). On the other hand, studies by Anikeeff (1953), Kennedy (1975), Schuk and Crivelli (1973), Snyder and Clair (1976), Weaver (1960) and Walker (1974) have found that students' grades are significantly related to their evaluations of instructors. These relationships do not necessarily hold for all aspects of the evaluation. For example, in Holmes' (1971) study, "relationships existed between the grades students expected and the degree to which they reported they were stimulated by the instructor and the degree to which they felt the grading system was fair" (p. 956), but expected grades were not related to items assessing the instructor's presentation.

Many faculty members who teach large classes suggest that they receive lower evaluations because students feel isolated from the teacher, that they are viewed as simply a number or an unknown face in the class. To confound the influence of class size on the evaluations, most large classes are also required, as opposed to elective courses. And this is viewed by the faculty as another variable over which they have no control and one which negatively influences their ratings. Research on these two class characteristics offers very little support for their beliefs. While Lovell and Haner (1955) reported that instructors of classes over 30 received lower ratings, Mirus (1973) and others (Danielsen & White, 1976; Heilman & Armentrout, 1936) have reported that class size and teacher evaluations are positively related. Gage (1961), on the other hand, reported a curvilinear relationship; "teachers in courses with 30 to 39 students consistently received lower ratings than did those in courses with more or fewer students" (p. 17). Other researchers (Goodhartz, 1948; Guthrie, 1949; Solomon, 1966) have obtained results which indicate that there is no relationship between class size and teacher ratings.

The research on the influence of required vs elective courses on teacher evaluations is as inconclusive as most of the other research on teacher, student and class characteristics. Some investigators (Gage, 1961; Lovell & Haner, 1955) have found that student ratings of teachers of required courses are significantly lower than those of teachers of elective courses. In contrast, Goodhartz (1948), Heilman and Armentrout (1936) and Mirus (1973) have reported no significant difference in teacher evaluations between the required and elective courses.

From these teacher, student and class characteristics, three were selected for investigation - teaching form, student proclivity toward particular academic majors and degree of course content. The theoretical underpinning for this study lies within the Psychological Differentiation Theory developed by Witkin and his associates (Witkin, Dyk, Paterson, Goodenough & Karp, 1962). Witkin refers to the manner in which individuals perceive, think, problem solve, relate to others, etc. as their cognitive style. Cognitive styles relate to the tendencies that occur in a person's perception of stimulus configurations, encounters with symbolic representations and interactions with other people. Several cognitive styles have been discovered (see Ragan, Back, Stansell, Ausburn, Ausburn, Butler, Huckabay, & Burkett, Note 1), but the one that has had the greatest application to education has been field-dependence-independence (Witkin, 1973; Witkin et al., 1962; Witkin, Moore, Goodenough, & Cox, 1975). This cognitive style is a function of the extent to which the individual makes use of or depends upon an external frame of reference. Field-dependence-independence refers to the degree to which an individual perceives or conceptualizes part of a field as discrete from, rather than embedded in, the surrounding field as a whole. At the field-dependent end of the continuum, an individuals cognitive and social behavior is dominated by the prevailing field; at the field-independent end, individual's are able to experience a part separate from the whole. Manifestations of field-dependent cognitive behavior centers around passive acceptance of the field as it is found with little analysis or structuring of it. Field-dependent individuals tend to leave the organization of the field unmodified. Field-independent behavior, on the other hand, manifests itself after the field has been analyzed and structured, or after overcoming any inherent structure. When it is necessary to separate an item from the field, the field-independent individual will overcome any inherent structure and reorganize the field to facilitate the separation.

In terms of social orientation, field-dependent behavior depends upon the prevailing social frame of reference. They are described as people who are interested in other people and like to be around them, are warm, affectionate and accepting of others. Field-independent social behavior occurs after social information from the environment is restructured and analyzed. They are described as cold and distant in their interpersonal relationships, inconsiderate, demanding and rude, and concerned with ideas and principles rather than people.

These differences in cognitive and social abilities manifest themselves in educational interests and choices. Witkin and his associates (1975) state that

> As a general principle, relatively field-independent persons, taken as a group, are likely to show interest in domains where their cognitive skill ... is called for and where relations with people are not particularly involved. In contrast, relatively field-dependent persons, as a group, are likely to show interest in domains with a "people" emphasis ... and for which analytical/structuring competence does not particularly matter. (p. 63)

Field-independent individuals are less inclined to be influenced by the social skills or form of a lecturer, since they are less sensitive to social cues than individuals who are field-dependent. In terms of the evaluation of the lecturer, a field-independent audience is influenced more by the content of a lecture than by the lecturer's form of delivery. A field-dependent audience, on the other hand, is influenced more by the form of the lecturer than by the content of his presentation.

The present experiment investigated the three variables, form, content and academic major, in order to determine the effect these variables may have upon student-teacher evaluation and student achievement. Specifically, the following hypotheses were tested: (1) There is a statistically significant interaction between student academic major and course content to affect student achievement scores and teacher evaluation; (2) there is a statistically significant interaction between student academic major and teaching form to affect student achievement scores and teacher evaluations.

Method

Design

A 2 (Natural Science, Social Science) x 2 (High Content, Low Content) x 2 (Good Form, Poor Form) fixed effects factorial design was developed to determine whether or not academic major, lecture content and teaching form influence teacher evaluations and student achievement. A multivariate analysis of variance technique was used to test the hypotheses. The .95 level of confidence was selected as the critical value to determine significance. The strength of association, omega squared, was determined for the multivariate (Tatsuoka, 1970) and the univariate (Hays, 1973) models. The sample size provided .92 power against a one standard deviation difference at the .05 level of significance for a main effect with two levels in a 3-way analysis of variance (Kirk, 1968).

Subjects

Thirty-two (32) undergraduate students majoring in either a natural science (astronomy, biology, botany, chemistry, geology, mathematics, physics or zoology) and thirty

(30) undergraduate students majoring in a social science (anthropology, economics, history, philosophy, political science, psychology or sociology) volunteered to participate in this study. The students, enrolled in the introductory psychology course at the University of Oklahoma, received course credit for their participation.

Materials

The materials for the videotaped lectures, the evaluation questionnaire and the test instrument were used in previous research (Ware, 1975) and were used here with the author's permission (see Appendix B). The nature of the lecture content is such that it should be of interest to both field-dependent and field-independent students. It is unlikely, however, that students in an introductory psychology course would have any prior knowledge of the lecture topic. Field-independent subjects should find the material concerning the biochemical modification of memory interesting; this material should be consistent with the general areas of interest which require cognitive skills and demands minimal involvement with others. The lecture content centering around the educational implications of memory modification should be more appealing to field-dependent individuals. They seem to be attracted to the material which explains the impact that memory changes have in terms of facilitating the well being of others.

Videotape equipment. Color videotape equipment was used in the production of four videotape recordings (see Stimulus tapes). The videotapes were played on Panasonic color monitors for the presentation of the stimuli to the subjects.

Stimulus tapes. Four color videotape recordings were made of a male "teacher" lecturing on the subject of the biochemistry of memory. The presentations ranged from 17 minutes, 35 seconds to 20 minutes, 57 seconds; the mean length of the recordings was 19 minutes, 9 seconds. The high content lectures contained 26 facts concerning the subject matter, while the low content lectures related four such facts. The recordings in which good form was demonstrated presented a teacher who was friendly, lacked distracting mannerisms and used a well-modulated voice. The teacher who presented poor form appeared aloof, spoke in a monotone and exhibited several distracting mannerisms.

Four (4) judges (doctoral students in Educational Psychology or Counseling Psychology) rated each of the four videotapes using a six-item questionnaire (see Appendix D). The items were concerned with the lecturer's mannerisms, tone of voice and interest in his students. Each item was rated on a five-point scale. When the ratings of the two good form videotapes ($X = 23.75$) were compared to the ratings of the two poor form videotapes ($X = 11.50$), they were found to be significantly different (F (1, 3) = 242.00, p <

.001). The Spearman-Brown prediction formula (Winer, 1971) was used in conjunction with a 4 (videotape) x 4 (judge) analysis of variance to establish interrater reliability (r = .938).

Evaluation questionnaire. Evaluation of the lecturer was made using an 18-item, 5-point questionnaire (see Appendix E). The items concerned the student's satisfaction with the lecturer, and pertained to various aspects of lecturer behavior including manner of presentation, sense of humor and knowledge of subject matter. A factor analysis of the questionnaire had revealed two factors (Ware, 1975). Items loading the highest on Factor I tended to refer to the effect the lecturer has on the students. These items pertain to several self-rated effects of the lecturer on students. Factor II tended to refer to student perceptions of lecturer characteristics. A total evaluation score was computed from the sum of the 18 items. The internal consistency (Cronbach's alpha) for the questionnaire was .93 which compares favorably to the .96 reported by Ware (1975).

Test instrument. A 26-item cognitive test was given to the subjects (see Appendix F). In terms of Bloom's taxonomy (Bloom, 1954), the 26 items were analyzed by the researcher and found to be directed primarily toward knowledge and comprehension. The length of time given to the lectures precluded any attempts at teaching to any higher levels of the taxonomy. There was one four-choice item for each of the facts presented in the high content lectures. Using the total number of correct items, a total cognitive test score was computed for each subject. The internal consistency (KR-20) for the cognitive test was .73; Ware (1975) reported a KR-20 of .61.

Procedure

The natural science subjects and the social science subjects were randomly assigned to one of four conditions: (1) high content-good form, (2) high content-poor form, (3) low content-good form and (4) low content-poor form. The "academic major" variable was ignored for data collection purposes. This allowed subjects in the natural and social sciences to view the videotapes at the same time for each of the four conditions. The random assignment of subjects resulted in 18, 14, 16 and 14 subjects being placed in conditions 1, 2, 3 and 4, respectively.

After viewing the lecture, the subjects were instructed to complete the evaluation questionnaire. The cognitive test was then completed.

Results

A 2 (Natural Science, Social Science) x 2 (High Content, Low Content) x 2 (Good Form, Poor Form) fixed effects multivariate analysis of variance was performed on the evaluation (Eval) scores and the cognitive (Cog) scores. The *F* approximations of the Wilks' lambda values resulting from this analysis are presented in Table 1. A linear combination of Eval and Cog was found to be statistically significant for the main effects of content and form. The interaction between content and form was also found to be significant. The proportions of variance in the combined Eval and Cog variables that are accounted for by the content and form main effects was 12.2 and 1.9 percent, respectively. The main effect for major accounted for less than one percent of the variance. The interaction between content and form accounted for only 1.1 percent of the variance. The remaining interactions each accounted for less than one percent of the variance. The multivariate omega squared values are presented in Table 2. In order to determine the source of these significant effects, additional analyses were performed. Each of the dependent variables, Eval and Cog, was analyzed using a univariate analysis of variance technique.

Cognitive Test Scores

A 2 (Academic Major) x 2 (Lecture Content) x 2 (Teaching Form) analysis of variance was performed on the Cog scores obtained by the students (see Table 3). The mean cognitive test scores for each of the eight lecture groups are presented in Table 4.

Table 1

Multivariate Analysis of Variance of the Effects of Major, Content and Form upon Cognitive and Evaluation Scores

Source	df	SSCP		F[1] (2, 53)
Major (M)	1	26.819		2.59
		118.618	524.642	
Content (C)	1	321.615		11.87 **
		-112.774	39.544	
Form (M)	1	61.582		13.51 **
		458.110	3407.899	
M x C	1	34.937		1.24
		3.720	0.396	
M x F	1	0.074		1.39
		5.321	385.115	
C x F	1	39.422		3.62 *
		168.012	716.047	
M x C x F	1	19.056		0.94
		43.530	99.440	

Within	54	750.163	
		232.282	7387.502
Total	61	1253.667	
		916.819	12560.581

[1]The approximate *F* values are based on Wilks' lambda (Rao, 1973).

* $p < .05$

** $p < .001$

Table 2

Strength of Relationship for Cog and Eval

Group	Univariate Omega Squared	Multivariate Omega Squared
Major (M)		
Cog	0.010	0.006
Eval	0.030	
Content (C)		
Cog	0.234	0.122
Eval	0.000	
Form (M)		
Cog	0.036	0.019
Eval	0.253	
M x C		
Cog	0.016	0.009
Eval	0.000	
M x F		
Cog	0.000	0.000
Eval	0.019	
C x F		
Cog	0.019	0.011
Eval	0.045	
M x C x F		
Cog	0.039	0.003
Eval	0.000	

Table 3

Analysis of Variance of the Effects of Major, Content and Form upon Cognitive Scores

Source	*df*	*SS*	*MS*	*F*
Major (M)	1	26.819	26.819	1.93
Content (C)	1	321.615	321.615	23.15 **
Form (F)	1	61.582	61.582	4.43 *
M x C	1	34.937	34.937	2.51
M x F	1	0.074	0.074	0.01
C x F	1	39.422	39.422	2.84
M x C x F	1	19.056	19.056	1.37
Within	54	750.163	13.892	
Total	61	1253.666		

* $p < .05$

** $p < .001$

Table 4

Means and Standard Deviations of Cognitive Scores

Group	N	Mean	Standard Deviation
Natural Science			
High Content			
Good Form	9	17.333	3.742
Poor Form	6	12.667	4.926
Low Content			
Good Form	8	8.500	1.773
Poor Form	7	9.286	3.039
Social Science			
High Content			
Good Form	9	13.444	5.270
Poor Form	8	10.875	4.155
Low Content			
Good Form	8	9.875	3.182
Poor Form	7	8.286	2.059
Total	62	11.451	4.619

The main effect for content was significant. Students who viewed the high content lectures scored higher on the cognitive test than those who viewed the low content lectures. The main effect for teaching form was also statistically significant. Students exposed to the good form lectures performed better on the cognitive test than the students exposed to the poor form lectures.

The main effect for academic major was nonsignificant; the interaction effects which included academic major were also nonsignificant.

The most important effect in influencing cognitive test scores was the content main effect. This accounted for 23.4 percent of the variance of the obtained cognitive scores. The academic major and form main effects accounted for 1.0 and 3.6 percent of the variance, respectively. The major-content, major-form and content-form interaction effects accounted for, in order, 1.6, 0.0 and 1.9 percent of the variance. The second-order interaction accounted for less than one percent of the variance (see Table 2.).

Evaluation scores

A 2 (Academic Major) x 2 (Lecture Content) x 2 (Teaching Form) analysis of variance was performed on the Eval scores (see Table 5). The mean evaluation scores for each of the eight lecture groups are presented in Table 6.

Table 5

Analysis of Variance of the Effects of Major, Content and Form upon Evaluation Scores

Source	df	SS	MS	F
Major (M)	1	524.642	524.642	3.83
Content (C)	1	39.544	39.544	0.29
Form (F)	1	3407.899	3407.899	24.91 **
M x C	1	0.396	0.396	0.00
M x F	1	385.115	385.115	2.82
C x F	1	716.047	716.047	5.23 *
M x C x F	1	99.440	99.440	0.73
Within	54	7387.502	136.806	
Total	61	12560.584		

* $p < .05$

** $p < .001$

Table 6

Means and Standard Deviations of Evaluation Scores

Group	N	Mean	Standard Deviation
Natural Science			
High Content			
Good Form	9	64.333	9.950
Poor Form	6	45.000	14.269
Low Content			
Good Form	8	56.375	9.665
Poor Form	7	55.857	18.819
Social Science			
High Content			
Good Form	9	60.778	5.585
Poor Form	8	36.500	8.452
Low Content			
Good Form	8	58.250	11.461
Poor Form	7	42.571	13.722
Total	62	53.129	14.474

The main effect for academic major and all interaction effects involving academic major were nonsignificant.

A significant interaction effect was found between lecture content and teaching form. Those students who viewed the high content lectures gave higher evaluations to the lecturer who used good form than the lecturer who used poor form (see Figure 1). In order to interpret the interaction effect, comparisons among the four groups were made. The group means were compared using Scheffe's procedure (Kirk, 1968). The results of these comparisons are presented in Table 7. While the group Eval means for the groups who viewed the lectures delivered by a lecturer using good teaching form did not differ significantly from one another, the contrast between the group means for the groups who viewed the lectures delivered by a lecturer using poor teaching form was found to be significant. It appears, then, that evaluation scores indicate significant mean differences in lecture content for lectures presented using poor form, but failed to reflect such a difference for the lectures presented using good form. These simple main effects accounted for the significant main effect for form.

In addition, contrasts between pairs of group means were made to determine whether or not significant differences in evaluations existed within each of the content groups, between the group that viewed the good teaching form and the group that viewed the poor teaching form. For the

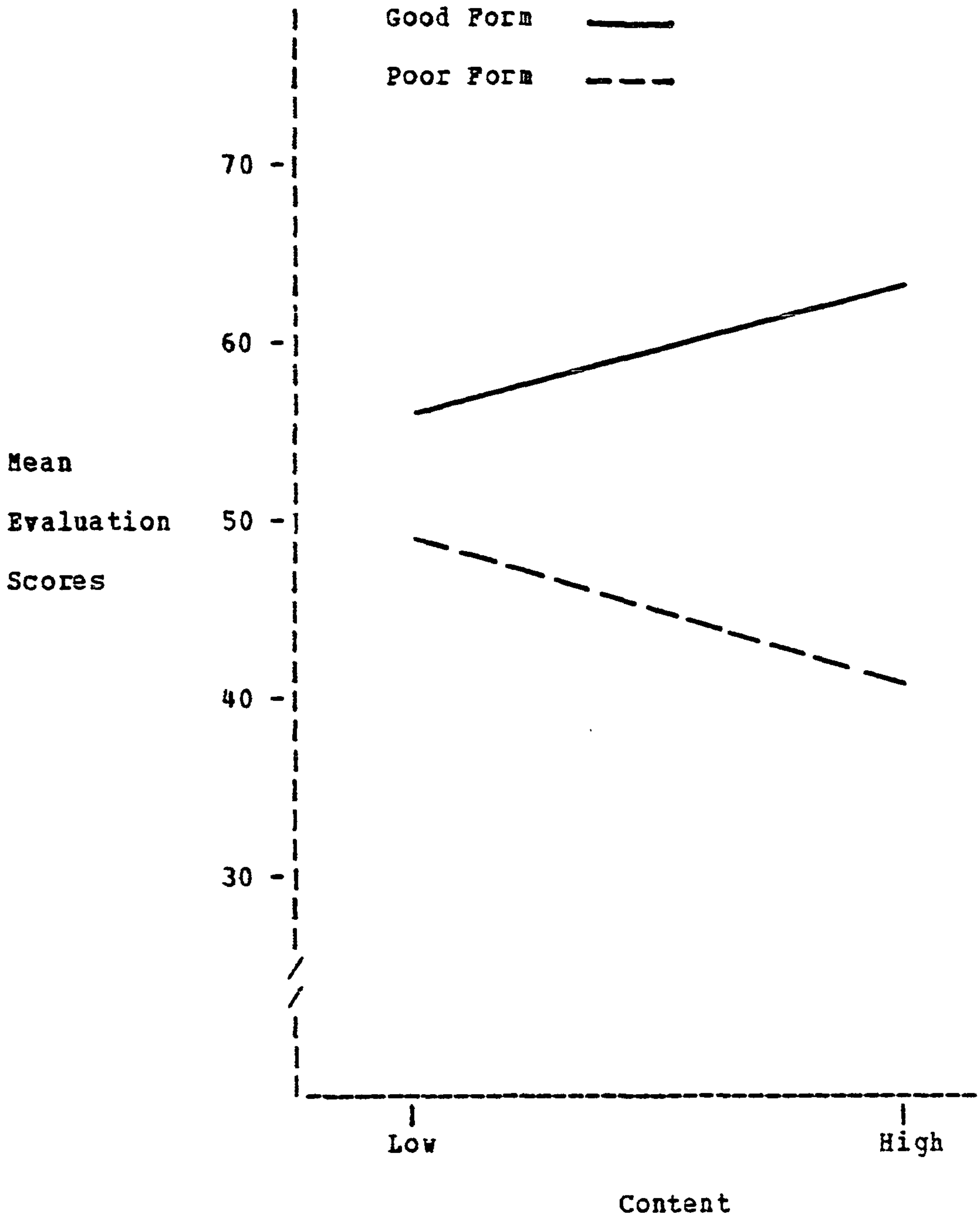

Figure 1: The interaction between Content and Form

Table 7

Results of Group Comparisons of Evaluation Scores using Scheffe's Procedure (Collapsed over Academic Major)

Comparison[1]	Group Means	Difference	*F*
HC/GF vs. HC/PF	62.56 and 40.14	22.42	28.92 *
HC/GF vs. LC/GF	62.56 and 57.31	5.24	1.70
HC/GF vs. LC/PF	62.56 and 49.21	13.34	10.25 *
HC/PF vs. LC/GF	40.14 and 57.31	-17.17	16.09 *
HC/PF vs. LC/PF	40.14 and 49.21	-9.07	4.21 *
LC/GF vs. LC/PF	57.31 and 49.21	8.10	3.58

[1]HC = high content, LC = low content, GF = good form, PF = poor form

* $p < .05$

low content lectures, there was no significant difference in the Eval means. The contrast between the group Eval means for the high content lectures, however, was significant. It seems, then, that students evaluate low content lectures about the same regardless of the teaching form used by the lecturer; on the other hand, students evaluate lectures delivered using good form significantly higher than lectures delivered using poor form within the high content lectures.

The effect most important in influencing evaluation scores was found to be the form main effect. This main effect accounted for 25.3 percent of the variance of the obtained evaluation scores. The academic major main effect and the content-form interaction effect accounted for 3.0 and 4.5 percent of the variance, respectively; each of the remaining effects was nonsignificant and accounted for less than one percent of the variance (see Table 2).

Discussion

The purpose of this study was to investigate the effects of two characteristics of lecture presentation and one student characteristic upon two measures of student outcome.

While academic major appeared to have no influence upon either faculty evaluations or student achievement, both lecture content and teaching form seemed to affect evaluations and achievement. The impact of content and form on test performance was more straightforward than their impact on faculty evaluations.

The data support the hypothesis that student test performance is influenced by the two lecture characteristics, lecture content and teaching form. Students tended to score higher on the cognitive test when the lecture was high in content than when there was relatively little content in the lecture. In view of the fact that the test instrument dealt with specific information presented in the high content lecture, this finding is logically consistent. Student performance on the test also was higher when the lecture was presented by a lecturer using good teaching form. Although the impact of these two characteristics is relative in terms of the total variance accounted for, lecture content seems to be the more important characteristic influencing student performance on cognitive tests.

The effects of teaching form upon cognitive test performance found in this study are consistent with the findings reported by Coats and Smidchens (1966) and Ware and Williams (1975). Cognitive test scores may be viewed as an indication of the knowledge gained by the students and their test performance. Both the "good form" and the "poor form" groups had an equal opportunity to learn the material, since they both had equal exposure to it. Differences in test outcome may be attributed to either a difference in motivation to perform well on the cognitive instrument or a difference in motivation to learn the subject matter, and, therefore, a difference in knowledge of the topic. The behaviors exhibited by the lecturer demonstrating good teaching form may have resulted in a higher level of motivation to either learn or perform well on the cognitive test.

Of the three characteristics affecting lecturer evaluations, teaching form was found to be the most important one. The effect, however, was not consistent across both levels of lecture content. That is, lecturer evaluations seem to be affected by the interaction between lecture content and teaching form. Students who viewed the lecturer demonstrating good teaching form evaluated the lecturer higher than students who viewed the lecturer exhibiting poor teaching form. The finding that students give higher evaluations for good teaching form across lecture content is in agreement with the findings reported by Naftulin, Ware and Donnelly (1973) and Ware and Williams (1975).

Differences in lecture content was most clearly indicated by the evaluations given by students in the high and low content groups who viewed the lecturer demonstrating poor teaching form. The differences were not, however, in the expected direction. Higher evaluations were given by students exposed to the low content lecturer than by those exposed to the high content lecturer.

Students who viewed the lecturer demonstrating good teaching form did not differentiate in their ratings of high content and low content lectures to the specified criterion of significance. Under the condition of good teachinf form, the mean evaluation scores of students who viewed the lecture which contained 26 specific facts were not statistically different from the mean evaluation scores of students who viewed the lecture which contained only four specific facts. Under the condition of good teaching form student evaluations did not reflect differences in lecture content although a trend in the expected direction was observed with the mean evaluation score for the high content group being slightly higher than the mean evaluation score for the low content lecture group.

The artificial nature of the experimental situation resulted in limitations on the conclusions which were drawn from the data. The use of videotaped lectures, while controlling many extraneous variables, eliminated the usual interaction which takes place between a teacher and his students.

Another limitation was the nature of the subject pool. The subjects were drawn from an introductory psychology course which is predominately freshmen. Because freshmen generally are less certain of their educational goals, their self-classification as majoring in one of the natural sciences or social sciences is more subject to change. This limitation may explain why academic major did not have a significant effect upon student-teacher evaluations.

Instructing teachers in techniques of good teaching form may have some value in terms of student achievement.

This instruction would include not only techniques to gain and maintain the attention of students and to make the learning experience more acceptable to the students, but also training to eliminate behaviors which are counter to good teaching form. The impact of teaching form upon evaluations will be the greatest when the lecturer is using poor form and the content of the lecture is high. That is, poor teaching form combined with high lecture content seems to have a detrimental effect upon evaluations of teacher effectiveness.

The belief that factors in addition to ability influence evaluations given by students is supported by the results of this study. Specifically, evaluations seem to be influenced by the teacher's form of delivery. Teachers who develop and master good teaching form are more likely to receive high student evaluations regardless of how much information is presented in their lectures. One implication is that teachers can obtain good ratings while students may be learning very little. To guard against the possibility of such a problem arising, any program which is designed to evaluate a faculty member's teaching ability should include a measure of student achievement.

Much of the published literature on faculty evaluation is not research but rather discussions of the parameters which are and those which are not appropriate for faculty evaluation research. The research that has been done has dealt primarily with the validity and reliability of student-teacher evaluations. Some of the empirical work has been directed toward the identification of the teacher, student and class characteristics which influence the evaluation of teacher effectiveness. Many of these studies have been post hoc using such descriptive variables as age, sex, faculty rank and college class. The relationships between some of the descriptive variables and the evaluations have ranged from significantly positive to significantly negative for the same variables. The obvious limitation of these studies is the inability to assign causality.

Based on the research which has been done, the influence and importance given to faculty evaluations in the decision making process is not warrented. Very few experiments have been done to determine the nonrelevant variables which influence faculty ratings. Variables such as motivation, perception, and class cohesion need to be studied in greater detail. If we are to understand the process of faculty evaluation, and make appropriate use of it, considerably more empirical work needs to be done to determine and understand the nature of the variables which influence faculty evaluations.

References

Anikeeff, A. M. Factors affecting student evaluation of college faculty members. Journal of Applied Psychology, 1953, 37, 458-460.

Bendig, A. W. The relation of level of course achievement to students' instructor and course rating in introductory psychology. Educational and Psychological Measurement, 1953, 13, 437-448.

Berns, R. The Berkeley uprisings: A study of the conditions of riotous behavior. American Journal of Psychiatry, 1967, 123, 1165-1169.

Bloom, B. S. (Ed.). Taxonomy of educational objectives, Handbook I: Cognitive domain. New York: McKay, 1956.

Blum, M. M. An investigation of the relation existing between students' grades and their ratings of an instructor's ability to teach. Journal of Educational Psychology, 1936, 27, 217-221.

Centra, J. A. Do student ratings of teachers improve instruction? Change, 1973, 5(3), 12-13.

Clark, K. E., & Keller, R. J. Student ratings of college teaching. In R. A. Eckert (Ed.), A university looks at its program. Minneapolis: University of Minnesota Press, 1954.

Coats, W. D. & Smidchens, U. Audience recall as a function of speaker dynamism. Journal of Educational Psychology, 1966, 57, 189-191.

Danielsen, A. L., & White, R. A. Some evidence on the variables associated with student evaluations of teachers. Journal of Economic Education, 1976, 7, 117-119.

Downie, N. W. Student evaluation of faculty. Journal of Higher Education, 1952, 23, 495-496.

Gage, N. L. The appraisal of college teaching: An analysis of end and means. Journal of Higher Education, 1961, 32, 17-22.

Goodhartz, A. S. Student attitudes and opinions relating to teaching at Brooklyn College. School and Society, 1948, 68, 345-349.

Guthrie, E. R. The evaluation of teaching. Educational Record, 1949, 30, 109-115.

Hays, W. L. _Statistics for the social sciences_ (2nd ed.). New York: Holt, 1973.

Heilman, J. D., & Armentrout, W. D. The rating of college teachers on ten traits by their students. _Journal of Educational Psychology_, 1936, _27_, 197-216.

Holmes, D. S. The relationship between expected grades and students' evaluations of their instructors. _Educational and Psychological Measurement_, 1971, _31_, 951-957.

Holmes, D. S. Effects of grades and disconfirmed grade expectancies of students' evaluations of their instructor. _Journal of Educational Psychology_, 1972, _63_, 130-133.

Kennedy, W. R. Grades expected and grades received--their relationship to students' evaluations of their faculty performance. _Journal of Educational Psychology_, 1975, _67_, 109-115.

Kirk, R. E. _Experimental design: Procedures for the behavioral sciences_. Belmont, Ca: Brooks/Cole, 1968.

Lasher, H., & Vogt, K. Student evaluation: Myth and realities. _Improving College and University Teaching_, 1974, _22_, 267-269.

Lovell, G. D., & Haner, C. F. Forced-choice applied to college faculty rating. *Educational and Psychological Measurement*, 1955, *15*, 291-304.

Mirus, R. Some implications of student evaluation of teachers. *Journal of Economic Education*, 1973, *5*, 35-37.

Naftulin, D. H., Ware, J. E., Jr., & Donnelly, F. A. The Dr. Fox lecture: A paradigm of educational seduction. *Journal of Medical Education*, 1973, *48*, 360-365.

Null, E. J., & Walters, J. E. Values of students and their ratings of a university professor. *College Student Journal*, 1972, *6*(4), 46-51.

Rader, N. F. College student ratings of instructors. *Journal of Experimental Education*, 1968, *37*, 76-81.

Ragan, T. J., Back, K. T., Stansell, V., Ausburn, L. J., Ausburn, F. B., Butler, P. A., Huckabay, K., & Burkett, J. R. *Cognitive style: A review of the literature*. Unpublished manuscript, Lowry AFB CO: Technical Training Division, Air Force Human Resources Laboratory, January, 1978.

Rao, C. R. *Linear statistical inference and its applications* (2nd ed.). New York: Wiley, 1973.

Remmers, H. H. The measurement of teaching personality and its relation to learning process. Education, 1930, 51, 27-35.

Remmers, H. H., & Elliot, D. N. The Indiana College and University Staff-Evaluation Program. School and Society, 1949, 70, 168-171.

Riley, J. W., Ryan, B. F., & Lifshitz, M. The student looks at his teacher. New Brunswick, N. J.: Rutgers University Press, 1950.

Rosenshine, B., Cohen, A., & Furst, N. Correlates of student reference ratings. Journal of College Personnel, 1973, 13, 269-272.

Schuk, A. J., & Crivelli, M. A. Animadversion error in student evaluations of faculty teaching effectiveness. Journal of Applied Psychology, 1973, 58, 259-260.

Snyder, C. R., & Clair, M. Effects of expected and obtained grades on teacher evaluation and attribution of performance. Journal of Educational Psychology, 1976, 68, 75-82.

Solomon, D. Teacher behavior dimensions, course characteristics, and student evaluations of teachers. American Educational Research Journal, 1966, 3, 35-47.

Tatsuoka, M. M. Discriminant analysis. Champaign, Ill: Institute for Personality and Ability Testing, 1970.

Walker, B. D. Anticipated grade and student rating of teaching. Improving College and University Teaching, 1974, 22, 220-221.

Ware, J. E., Jr. The Doctor Fox Effect: An experimental study of the effectiveness of lecture presentations and the validity of student ratings, Dissertation Abstracts International, 1975, Ser A 35(7), 4266A-4267A, (University Microfilms No. 75-146).

Ware, J. E., Jr., & Williams, R. G. The Dr. Fox Effect: A study of lecturer effectiveness and ratings of instruction. Journal of Medical Education, 1975, 50, 149-156.

Weaver, C. H. Instructor rating by college students. Journal of Educational Psychology, 1960, 51, 21-25.

Williams, R. G. & Ware, J. E., Jr. Validity of student ratings of instruction under different incentive conditions: A further study of the Dr. Fox Effect. _Journal of Educational Psychology_, 1976, _68_, 48-56.

Williams, R. G. & Ware, J. E., Jr. An extended visit with Dr. Fox: Validity of student satisfaction with instruction ratings after repeated exposures to a lecturer. _American Educational Research Journal_, 1977, _14_, 449-457.

Winer, B. J. _Statistical principles in experimental design_ (2nd ed.). New York: McGraw-Hill, 1971.

Witkin, H. A. _The role of cognitive style in academic performance and in teacher-student relations_. Research Bulletin 73-11. Princeton, N. J.: Educational Testing Service, 1973.

Witkin, H. A., Dyk, R. B., Faterson, H. F., Goodenough, D. R., & Karp, S. A. _Psychological differentiation_. Potomac, Md.: Erlbaum, 1962.

Witkin, H. A., Moore, C. A., Goodenough, D. R., & Cox, P. W. *Field-dependent and field-independent cognitive styles and their educational implications*. Research Bulletin 75-24. Princeton N. J.: Educational Testing Service, 1975.

Appendix A

PROSPECTUS

INTRODUCTION AND PROBLEM

The informal evaluation of teachers has been going on for hundreds of years. Students have always discussed among themselves the competencies or performances of their teachers. In more recent years there has been an increasing demand to formalize student evaluation of university faculty. This demand has come from legislators, the public and the student "consumer" of the teacher's wares. The interest of students in the evaluation of their teachers was compounded by the activism of the 1960's (see Berns, 1967). The students demanded more input into decisions concerning not only what was taught but who was to teach it. In an effort to deal with this problem, the American Association of University Professors made a policy statement to guide in the proper evaluation of teaching (AAUP, 1975) which included their position on the place of student evaluations in the overall evaluation of faculty.

There has been, however, considerable resistance to the formalization of teacher evaluation procedures by some faculty members (Bryan, 1968; Solbin, 1969). Many feel that students are not competent or qualified to judge their teaching performances (Rodin & Rodin, 1972). Some of these teachers believe that their students are either unaware or can not separate the teacher from the course content. In addition, some teachers feel students do not understand the role that "form" or style plays in teaching.

Some teachers also feel that there are a number of variables over which they have no control that also affect the evaluations they receive from their students. One of these variables is the student's academic major. Some of these teachers believe that a student whose major is one of the natural sciences places a higher value, in terms of his evaluations, on course content than on teaching form while a student whose major is one of the social sciences will assign greater value to form than to content.

A great deal of literature on the subject of faculty evaluation has been published. Most of this literature is not of an experimental nature, however. The authors of some of this nonexperimental literature are concerned with the appropriateness of "effective teacher" as a subject for investigation. Others discuss the problems of identifying the criteria of effective teacher; some of these same writers are also concerned with who should be involved in the identification of these criteria. Still others address the problems involved with the use of students as evaluators of teacher effectiveness. The experimental literature, on the other hand, deals primarily with the validity and reliability of student-teacher evaluations. Some of the experimental work has been directed toward the identification of the variables which affect teacher evaluations.

The present experiment will investigate the three variables, form, content and academic major, in order to determine the effect these variables may have upon student-teacher evaluation and student achievement. An appropriate analysis of these latter two variables will allow for the assessment of the impact that form, content and academic major has on them.

Statement of Problem and Purpose

Research problem

Do form, content and academic major affect teacher evaluation and student performance?

The following questions were generated in an attempt to examine the three basic concepts of this study.

1. Are teacher evaluations and student achievement affected differently when a course contains a great amount of content than when it contains very little content?

2. Does the form of the lecturer, good or poor, affect the teacher's evaluation and the student's achievement?

3. Does the student's academic major affect his performance or his evaluation of his teacher?

Purpose

The purpose of this study is to investigate the effect that a teacher's form, the content of the course and the academic majors of the students may have upon students' evaluations of the teacher and students' achievement levels.

Hypotheses

The research problem and questions were the source for the following null hypotheses:

1. The three independent variables will not interact to affect either of the dependent variables.

2. The independent variables, taken pair-wise, will not interact to differentially affect the two dependent variables.

3. The dependent variables will not be affected differently by the levels of course content.

4. Different levels of teaching form will not affect the dependent measures differently.

5. Academic majors will not have a differential effect upon the student outcome measures.

Definition of Terms

For the purposes of this study definitions of the following terms are given:

Teacher: "A teacher is a person engaged in interactive behavior with one or more students for the purpose of effecting a change in those students." (McNeil & Popham, 1973, p. 219)

Form: the style or manner in which information, ideas and concepts are presented, and characterized by quality of voice, mannerisms and friendliness.

Good form: a style characterized by a well modulated voice, an absence of distracting mannerisms and a friendly manner.

Poor form: a style characterized by a monotonic voice, the presence of distracting mannerisms and an aloof manner.

Content: information, ideas and concepts relating to the biochemistry of memory.

Evaluation: the determination of the value or worth of a teacher's effort or effectiveness in communicating the lecture content to his listeners.

Natural science: The natural sciences shall include no other than the following disciplines: astronomy, biology, botany, chemistry, geology, mathematics, physics, and zoology.

Social science: The social sciences shall include the following disciplines to the exclusion of all others: anthropology, economics, history, philosophy, political science, psychology, and sociology.

Limitations

The study was conducted within the limits of the parameters listed below:

1. The evaluation is limited to the rating of teachers or faculty at the college and university level by college and university undergraduate students.

2. Faculty is limited to the teaching faculty and those individuals functioning in the role of teaching faculty.

REVIEW OF RELATED LITERATURE

Although the literature on faculty evaluation is broad, there is no great depth to it. The review of this body of literature will include the purposes for which evaluations are made, the criteria used in the evaluation of teachers, the methods by which evaluations are made and the variables most often involved in such evaluations.

Purpose of Evaluation.

There are several purposes of evaluation. The literature has shown that evaluations are most commonly used (a) to provide end-of-semester feedback to the instructor, (b) to assist administrators in making decisions concerning promotions, salary adjustments and tenure and (c) to provide the student body with information for selecting instructors. Because of their almost continuous contact with their teachers, students are in the best position to provide instructors with information concerning the effects of their behavior and teaching practices. Most people would agree that a primary purpose of a university is to teach its students, and, therefore, the evaluation of teacher effectiveness should be a significant input into administrative decisions concerning the teacher. Again, the student is in the most advantageous position to make such an evaluation. Based upon surveys (Astin & Lee, 1966; Gustad, 1961) of universities, administrators indicate that classroom teaching is a significant factor in the evaluation of faculty members, but the source of information on this point is usually someone or something other than the student. Astin and Lee (1966) found that of the 15 most frequently used sources of information, informal student opinion rated fifth and systematic student evaluation rated twelfth; in fact, almost 50 percent of the institutions surveyed reported that they did not use systematic student ratings to evaluate teaching effectiveness. Gustad (1961) also found that "with one or two exceptions the most frequently cited source of information can be described simply as hearsay" (p. 203). Remmers (1958), however, set the tone for the formal evaluation of teachers by their students when he noted,

> No teacher has any choice as to whether he wishes to be judged by his students. The only choice he has is whether he wishes to know how he is judged and thus possibly capitalize on this feedback (p. 20).

This statement may also be viewed as a justification for obtaining student ratings. Although the evidence offers some support for the assertion that student feedback has a positive effect on teaching performance, the evidence is less than conclusive. A number of studies have concluded that feedback in the form of student evaluations resulted in a change in teacher performance (Braunstein, Klein, & Pachla, 1973; Centra, 1973a; Marsh, Fleiner, & Thomas, 1975). Other investigators have not arrived at the same conclusions. Miller (1971), for example, found that end-of-semester evaluations for teachers in a feedback group did not differ significantly from end-of-semester ratings for teachers in a no-feedback group. He has suggested two factors that may explain, in part, the lack of obtained significance. Since the subjects were graduate teaching assistants, many of whom may not have committed themselves to teaching, their motivation to use the feedback to improve their teaching practices may have been less than adequate. In addition, the fact that the largest cell only had seven subjects may have resulted in reduced power.

No reports have been found on the improvement of instruction as a result of teachers being evaluated by their administrators. As a matter of fact, since administrators are the primary source of teacher evaluations (Astin & Lee, 1966; Luthans, 1967; Seldin & Wakin, 1974), a finding by Astin and Lee suggests that the contrary may be true; they state that

> citing 'classroom teaching' as a 'major factor' in personnel decisions does not encourage improved teaching as long as teaching ability is more likely to be evaluated on the basis of scholarly research and publication rather than information more directly related to effective performance in the classroom. (p. 372)

Buxton (1956) has suggested that if teachers are forced to participate in an evaluation program their reaction is more likely to be negative toward their ratings. These instructors, however, are "likely to accept student ratings as a source of personal evaluation and guidance" (p. 354) when given some choice in the matter. In a survey of 416 institutions offering a bachelor's degree, 48 percent of the respondents (149 institutions) indicated that the ratings are made known to the teacher only; 22 percent made the student ratings known to the teacher and one or more of the teacher's administrators (Bryan, 1968). In all but eight of the latter institutions participation was voluntary. All teachers were required to participate at five of these eight while three required participation by nontenured faculty only.

Administrators use a wide range of factors, both formal and informal, to assess faculty performance. As might be expected, most administrators indicate that classroom teaching is the most important factor in evaluating their faculty. The feeling among faculty members has been, however, that "teaching is paid only lip service, that other factors especially publication, really pay off in promotion" (Gustad, 1961, p. 202). A 1958 study (Caplow & McGee, 1958) found university faculty making similar comments. They felt they were hired to do one job, teach, and when they were evaluated for promotion, the judgment was perceived as being based primarily on research and publications.

> It is neither an overgeneralization nor an oversimplification to state that in the major universities in the United States, the evaluation of performance is based almost exclusively on publication of scholarly books or articles in professional journals. (p. 83)

The pessimism of the faculty is not without basis. While the trend has been to objectify the evaluation of the classroom performance of teachers, most of the information available to administrators is subjective. The six sources which seem to be used most often are as follows: informal and formal student opinion, classroom visitations, colleagues' opinions and the opinions of chairmen and deans. With the exception of formal ratings and, possibly, classroom visitations, these sources are unsystematic and subjective. Luthans (1967) analyzed the promotion process used by large universities. When asked how they made evaluations one-third of the deans and department heads reported that no consistent, objective evaluation procedure was used. Bibliographies were employed by half; publications were considered by slightly less than half; objective teaching ratings were not widely used.

The American Council on Education (ACE) initated a survey in 1966 dealing with faculty evaluation. The results of this survey were reported by Astin and Lee (1966). They indicated that systematic student evaluations were used by only 12.4% of the institutions which responded to the questionnaire and almost 50% of them didn't use them at all. In addition, the frequency of classroom visits as a source of information was almost as low (14.0%); this source was not used by 39.5% of the institutions. The opinions of chairmen, deans and colleagues and the informal opinions of students were reportedly used by 85.1%, 82.3%, 48.9% and 41.2% of the institutions, respectively, as a source of data. A follow-up study (Seldin & Wakin, 1974) reported an increase in the use of objective evaluations. The population, however, was more selective than that of the 1966 survey; they used only private liberal arts colleges, while Astin and Lee's report was based on responses from junior colleges, teachers colleges, liberal arts colleges and five colleges within the university. The follow-up study discovered a shift toward the use of systematic student ratings, a change from 11.2% to 29.3% of the responding schools; and a shift away from informal student ratings, from 47.2% to 17.8%. Only 5.1% of the schools reported using firsthand observation by colleagues, a decrease from the ACE study. While this downward trend is also seen in the use of colleagues' opinions, a decrement of 10.8 percentage points, the opinions of chairmen and deans have maintained their superior position (82 to 85% always use them).

Even with these gains in the use of data, faculty evaluation still has a long way to go. Sixty percent "of the deans admit that their personnel decisions are <u>not</u> based on rational, impersonal, and unprejudiced information" (Seldin & Wakin, 1974, p. 49). Although this situation may be understandable, it is far from acceptable. As reported in a statement prepared for the Council of the American Association of University Professors (1975) when an administrator uses his judgment as the "sole or determining factor" for the evaluation of teaching, he is "acting contrary to the policies set forth in the <u>Statement on College and University Government</u>" (American Association of College Professors, 1975, p. 200).

Supposedly the purpose of the publication of faculty evaluations is to provide a source of information concerning faculty and courses so that students might plan their academic programs better. Many faculty members, however, are dubious. They view such publications as an opportunity for students to publicize their anonymous "hatchet job". Because of the publication of student ratings, the poor teacher may feel so threatened that defensiveness and anxiety may inhibit any improvement that might otherwise be made. Even the best teachers may experience some anger over the "cheap shot" and thereby become resistant to the publication of student evaluations. A review of the 1965 and 1966 course critiques (Evans, 1969) at the University of Washington in Seattle revealed that when the vindictive remarks from student ratings were not included in the publication it was accepted "much less enthusiastically" (p. 103).

Although some teachers feel that the publication of their evaluations may bias their students when they are asked to rate them in the future offerings, such has not been the case. Students are unaffected by knowledge of their instructor's prior ratings when they are given the opportunity to evaluate them (McClelland, 1970). Crannell (1948) concurs; his data indicate that "any 'danger' of damage to an instructor's reputation through the uncontrolled anonymous attacks of a 'disgrunted minority' is certainly far from evident" (p. 11).

Evaluation Criteria.

The greatest single problem facing anyone who would do research in the area of teacher effectiveness is determining the appropriate criterion. A cursory review of the literature makes it obvious that there is no agreed upon criterion. Until the concept of teaching effectiveness is delimited, research in this area will lack a logical, coherent structure. This concept has no meaning apart from the criteria measures which are used to indicate some degree of success in teaching (Mitzel, 1960). Although "the" criterion has not been specified yet, criteria are used. As a result of Brownell's (1948) distinction between the process of learning and the product of learning, he has suggested two classifications of criteria: (a) process criteria and (b) product criteria. Mitzel (1960) has used the term process criteria to refer to "those aspects of teacher and student behavior which are believed to be worthwhile in their own right" (p. 1483). Research on teacher effectiveness has dealt primarily with teacher behaviors (Perry, 1969; Schulman & Trudell, 1972), as opposed to student behaviors. The investigator's interest may be focused on such teacher attributes as personality (Bendig, 1955; Heilman & Armentrout, 1936; Isaacson, McKeachie, & Milholland, 1963), knowledge (Elliott, 1950; Naflulin, Ware, & Donnelly, 1973) or classroom behavior (Withall, 1949). McNeil and Popham (1973) have voiced some skepticism of this approach to teacher effectiveness research when they said the following:

> By studying certain procedures employed by teachers and then assuming that these processes are related to pupil growth, the investigator gets at a readily accessible process criterion and hopes it reflects an outcome criterion. (p. 220)

Mitzel (1958) concurs; process criteria are sometimes looked for in the classroom "because of their assumed mediating effects on product criteria" (p. 1483). He and McNeil and Popham are criticizing the use of a process criterion because of the <u>assumed</u> relationship between it and a product criterion. Saadeh (1970) has argued that the teaching process alone is insufficient for investigation of teaching competence and that such a criterion must have as its basis student outcomes. There is strong support for Saadeh's position; a number of researchers, as well as professional organizations, have stated that in the final analysis the criterion for teacher competence must be student learning (e. g., American Association of University Professors, 1975, American Educational Research Association, 1952; Astin & Lee, 1969; Wilson, 1976).

Product criteria are a function of a set of specific and specified objectives toward which the teacher's efforts are directed. (This review will not deal with the philosophical questions of the value or worthiness of these goals, or, given that they are of value, that they are appropriate as educational objectives.) The degree to which these efforts are successful is usually determined by the degree to which the students have achieved. This achievement of the student being an indication or a measure of student growth is, by definition, the intent of teachers. However, even with the teacher's intentions and the student achievement it does not necessarily follow that the latter is due entirely to the former. And herein lies a very significant problem. How can that portion of the student's gain which is a direct result of the teacher's effectiveness be determined? Neeley (1968) has stated that "student gain would be an excellent measure if we could tell which student gained from which teacher" (p. 207). It is very difficult in most cases, if not impossible, to "designate and partial out the contribution to a particular product made by a specified aspect of the producing situation, such as a teacher" (Ryans, 1960, p. 1487). This difficulty is reflected in the limited use of achievement as a measure of teaching effectiveness. The research by Gessner (1973), McKeachie, Lin and Mann (1971), Lewis and Orvis (1973) and White (1976) has attempted to deal with this problem, and has demonstrated, to some extent, the efficacy of performance tests as measures of teacher performance. Gessner (1973), for example, used two performance criteria; the first was the cumulative raw score on three departmental examinations and the second was the obtained score on the National Medical Board Examination. He concluded:

> The high correlation between student ratings and class performance on national normative examinations suggests that such examinations themselves could be used as a measure of teacher effectiveness. (p. 569)

Sources of Evaluation.

By what method should teacher effectiveness be evaluated? The procedure of assessment is, to a great extent, a function of the purpose for which the evaluation is being conducted. To some extent, however, student ratings have a place in teacher evaluation regardless of the purpose. Evaluation by administrators and peers, as well as the review of teacher and course documentation, are the sources primarily used for institutional purposes. In terms of instructional improvement, data obtained from students and alumni may prove to be useful; and, to a lesser degree, self-evaluation could facilitate increasing teacher effectiveness. Rarely, if ever, is any source of data except student ratings used to inform students as to the worth of an instructor or a course taught by a particular instructor. Whatever the source of the evaluation may be, each offers different strengths and weaknesses.

Many investigators feel that students, because of their continual exposure to the teacher's effort, are in the best position to give information concerning the teacher's effectiveness. This data is derived either informally, via conversations with deans or department chairmen, or formally. The formal sources of student information are (a) student evaluations and (b) student performance. Although informal student opinion has consistently been identified as one of the most frequently used sources of teacher evaluation (see Astin & Lee, 1966; Gustad, 1961, 1967; Seldin & Wakin, 1974), research in this area is rare. Student ratings of teacher effectiveness, on the other hand, have received more consideration (Felman, 1972; Leonard & Beatty, 1975; Greenwood & Renner 1975; Kolevzon & Wiltse, 1973). The attention given to the relationship between student evaluations of teaching and achievement of educational goals and objectives has been well documented (Costin, Greenough, & Menges, 1971). There has been, however, some disagreement among researchers as to what that relationship is.

The results of a study by Rodin and Rodin (1972) indicate that the relationship between the mean scores of the students' teacher evaluations in 12 recitation sections and the mean student ratings of the teaching assistants for their respective sections was a partial correlation of -.746. This study has received considerable criticism, however, because of methodological problems (Marsh, et al., 1975; Rodin, Frey, & Gessner, 1975). Frey (1974) and others (Cohen & Berger, 1970; Gessner, 1973; Morsh, Burgess, & Smith, 1956; Remmers, Martin, & Elliot), found the relationship between student ratings and student achievement to be positive. Frey designed his study to replicate the Rodin and Rodin (1972) research, with some modifications. First, he used faculty members as the unit of study. These teachers taught two different courses; eight instructors taught Introductory Calculus and five taught Multidimensional Calculus. Each class, which met with the professor three hours per week, also met one hour a week with a teaching assistant for a quiz section. Second, the instructors in each course used a common syllabus, a common text and a common final examination.

After the course was completed, those students whose Scholastic Aptitude Test (SAT) scores were available to the investigator were sent an instructor evaluation form. The return rate was about 75%; of the 474 forms sent to the students, 354 forms were returned. The form consisted of six instructional factors. Frey found that the correlations between the average final examination score for an instructor's students and the average factor ratings for the instructor were all positive. The mean correlation for the two courses on each of the six factors ranged from +.31 to +.87.

Earlier in this review, administrators as sources of information for the evaluation of teacher effectiveness was discussed. Research cited in that discussion indicated that deans and department chairmen possess very little direct information concerning teacher classroom behavior or teacher effectiveness. Most of the information they obtain concerning the faculty's teaching is indirect and comes from reviews of research and publications and anecdotal reports (Astin & Lee, 1966). Other secondary sources which are available to administrators include colleague evaluations and course documentation (e. g., course syllabi and examinations). As a result of using informal student ratings and peer evaluations as sources of information, the administrator's judgment of a teacher's effectiveness will be no less biased than the judgment of either of these sources. There also tends to be a degree of self-selection in the use of these assessments; Mayhew (1967) said it this way:

> Every academic administrative officer hears a number of student remarks about teachers. The tendency is to make mental notes of these. But memory is selective and faulty. One remembers the good

> things one hears about people one likes and the bad about others. (p. 9)

The essence of a statement by Braunstein and Benston (1973) runs through the professional literature relating to administrators as teacher evaluators:

> University administrators typically have no organized way of obtaining information about the quality of the classroom performance of their faculty. (p. 244)

When an administrator uses research and publications to evaluate an instructor's teaching effectiveness, he is assuming that there is a high correlation between research and teacher effectiveness. Guthrie (1949) found that there is no significant association between the two. Another researcher (Hayes, 1971) has concluded that, based upon significant data, "in the department head's opinion, good teaching tends to be associated with high research ability" (p. 228). He also reported that the relationship between student ratings and research ability is weak and nonsignificant. Braunstein and Benston (1973) obtained instructor and course evaluations for 713 individual courses taught over a period of four semesters (summer of 1968 through the summer of 1969). They also collected data from departmental chairmen concerning the teaching and research standing of 347 professors. "In light of all this data," they concluded, "the popular use of research and publication as a method of evaluating university teaching is highly suspect" (p. 248). Gustad's (1961) report makes the point even stronger:

> The responses to the question dealing with research evaluation suggest quite clearly that research is not, in fact, actually evaluated; it is <u>counted</u> (p. 206).

A survey reported by Astin and Lee (1966) indicated that almost 50% of the 1,110 responding colleges and universities used colleagues' opinions as a source of information in teacher effectiveness evaluation. In another survey reported five years earlier (Gustad, 1961), of the 546 colleges and universities that returned the survey form, colleagues' opinions were the fifth most cited source of evaluation data. A third survey (Seldin & Wakin, 1974) indicated that almost 40% of the deans responding used peers' opinions in judging teacher effectiveness. Clearly, the opinions of colleagues have been a significant source of information for the rating of teaching performance. There is some degree of relationship between student ratings of a teacher and colleague ratings of the same teacher (Guthrie, 1949); Maslow & Zimmerman, 1956). In one of the first studies on the relationship of student and peer evaluations, Guthrie (1949) reported a high correlation among student evaluations (r=.89) and among faculty juries of seven colleagues (ranging from r=.64 to r=.76), but a weak correlation between the evaluations of the two groups. He suggests that one source of the difference in the faculty evaluations of teaching and in the student ratings is that

> students, when called upon to judge a teacher, have sat through from ten to fifty hours of his course, at least one-half of its total. The faculty are dependent on student hearsay, on the observation of the presumed effects of other men's instruction on their own students, and on inferences from their personal acquaintance with men and their academic record. (p. 113)

Maslow and Zimmerman (1956) reported a more substantial relationship. Data were collected on 86 Brooklyn College faculty members. Students and department colleagues rated each of the faculty members as a "teacher". They concluded that student judgments

> correlate so well (r=.69) with judgments of these same teachers that a faculty cannot take student judgment lightly without casting aspersions on its own competence to judge. (p. 189)

Starrack (1934) also reported a close agreement between student and faculty evaluations in more than 75% of the cases analyzed.

The use of documents to evaluate teacher effectiveness involves the review of publications, which was discussed in the section Sources of Evaluation, or course syllabi, outlines, materials and examinations. The surveys by Gustad (1961, 1967) indicate that course syllabi and examinations rank about tenth as a source of data in the evaluation of teaching. Astin and Lee (1966) reported that they rank seventh; 26.4% of the institutions polled used this source of data and 28.0% never used it. The use of this source to evaluate teachers should not inhibit them, discourage innovation or place any limits on the kind of teaching styles utilized by the teachers. This survey also placed alumni opinion last in the list of the 15 most frequently used sources of teacher evaluation data; less than 10% of the reporting institutions used it. Researchers have reported correlations between student and alumni ratings from r=.40 to r=.85. Drucker and Remmers (1951), discussing student evaluations of instructors and alumni ratings of the same instructors 10 years after graduation (correlations were between .40 and .68), indicate that "to a large extent the same relative ratings are given to instructors by current students ... and by alumni" (p. 139). Another investigation (Braunstein & Bentson, 1973) found comparable correlations. A correlation of .43 was reported for a natural science department, and for two other departments, social science and humanities, the r values were .85 and .84. Teachers who are rated as effective are generally remembered as being effective for years after graduation.

It should be noted that the three surveys which have been cited (Astin & Lee, 1966; Gustad, 1961, 1967) are at least ten years old and the information reported in them may not reflect the extent to which the different sources of teacher evaluation are used today by colleges and universities.

Evaluation Variables.

Good student evaluations are the result of more than just good teaching. There are basically three factors which may influence a student's rating of a teacher's effectiveness: (a) teacher characteristics, (b) student characteristics and (c) class characteristics. Many teachers argue that since they have little or no control over personal and environmental variables, such as age of faculty, academic rank, student's academic major and class size, and since these factors do influence student ratings, then student evaluations do not adequately reflect the faculty's contribution to the learning process. The evaluations will either be inflated or deflated depending upon the pattern of the various factors. Before this argument can stand as a viable criticism, it must first be determined whether or not the factors do, in fact, influence student evaluations.

A number of teacher characteristics including age, sex, academic rank, academic degree attained, experience and teaching style have been investigated. Heilman and Armentrout (1936) found no significant differences between student ratings and teaching experience, age, or sex. This study, conducted at the Colorado State College of Education, involved 46 teachers and 50 classes. The lack of a relationship of student rating to sex of teacher has also been reported by others (Remmers, 1930; Remmers & Elliot, 1949). Downie (1952) reported no difference between the ratings of 237 faculty members under age forty and 169 aged forty and over. Lasher and Vogt (1974) collected data at Bowling Green State University over six quarters from 120 faculty members and 1,072 courses. They reported significant differences between student evaluations of teaching effectiveness and age, faculty rank, academic degree and experience. At Brooklyn College, Riley, Ryan and Lifshitz (1950) found, for the most part, inverse relationships between student ratings and age, academic rank and academic degree. The elimination of age as an influence altered the relationships pertaining to rank and degree only slightly. Rayder (1968) reported "older,more educated and more experienced instructors were rated lower than younger and less experienced instructors" (p. 80). Other researchers have also found significant relationships existing between student ratings and several of the teacher characteristics, but they have identified the relationships as positive. Downie (1952) reported full professors tended to receive higher student evaluations than did other ranks. Clark and Keller (1954) and Gage (1961) indicated that associate professors and full professors were rated higher than were instructors or assistant professors. And Centra (1973b) demonstrated that "students rated women more favorably than men teachers" (p. 13).

In an interesting study of lecture style, Naftulin, Ware and Donnelly (1973) hired an actor to portray an authority, Dr. Fox, on the application of mathematics of human behavior. The actor, who knew nothing of the topic, was instructed to teach "charismatically and nonsubstantively" and to respond to questions with "double talk, neologisms, non sequiturs and contradictory statements." The subjects in this study, who were mental health educators (psychiatrists, psychologists and social workers), made significantly more favorable than unfavorable responses to a satisfaction questionnaire. These authors have demonstrated that students, in this case professional educators, can be "seduced" into believing that they have learned, and consequently, giving a teacher a high satisfaction rating. A weakness of this study is the lack of a measure of achievement. Researchers at Southern Illinois University, Carbondale, have dealt with this weakness in their studies of the effect of teaching style on student evaluations. These investigators found that seductive behavior (friendliness, expressiveness, enthusiasm, humor, etc.) in a lecture presentation had a greater influence on student ratings and student achievement than did lecture content (Ware & Williams, 1975). In another study, they reported that even after the students experienced a second lecture of the same style and content, ratings of instruction still did not accurately reflect lecture content and student performance (Williams & Ware, 1977). In a third study, Williams and Ware (1976) manipulated expressiveness, content and incentive. Their results from the manipulation of expressiveness were not in total agreement with their earlier findings; student achievement was not affected by expressive manipulation. Their data were, however, congruent with their previous research relative to student ratings; that is, student ratings were not sensitive to differences in lecture content under the high-expressiveness conditions, but under the low-expressiveness conditions they indicated variations in lecture content.

Most student characteristics make little difference in the evaluations that teachers receive from them. A number of studies found no significant differences in the rating of a teacher by male and female students (Goodhartz, 1948; Rayder, 1968; Remmers, 1930; Remmers & Elliot, 1949). Rayder (1968) reported no substantial relationship of student's age, college class or GPA. Goodhartz (1948) reported similar results; he concluded "there is no conclusive evidence for believing that the ratings given to an instructor are affected ... by such factors as the student's sex or college class" (p. 348). The research on the effect of the student's college class on the ratings of teachers is not conclusive, however. Downie (1952) and Rosenshine, Cohen and Furst (1973) found that upper division students tend to rate teachers more favorably than do lower division students. Also, graduate students are more inclined to give higher ratings than are undergraduates (Remmers & Elliot, 1949).

The effect of the student's academic major upon his teacher evaluation has received only minimal attention. An incidental finding in a study by Centra (1973b) was that natural science courses as opposed to courses in education, humanities and social sciences were viewed by students "as having a faster pace, as being more difficult and as being less likely to stimulate student interest" (p. 13). A more substantive study was done by Null and Walter (1972) at Purdue University. Major field of study was one of ten variables they were interested in studying. They asked 192 students enrolled in an introductory biology course to evaluate ten behavioral characteristics of their teacher. They found that academic major did not have a significant impact on any of the teacher characteristics evaluated; patterns were revealed, however, by an examination of mean scores. Biology majors (89 of the 192 students) preceived nine of the ten dimensions of instructor behavior somewhat better than non-biology majors. In another study involving academic major as an independent variable, Rayder (1968) reported that "student ratings of instructors were not substantially related to student's ... major area ... " (p. 78). In view of the cited research, very little can be said, with any degree of comfort, concerning the role that a student's academic major plays in how he will evaluate his instructor's teaching effectiveness.

The relationship between grades, expected or received, and teacher evaluations is not clear, even though it has been extensively researched. Many studies have found no relationship between these variables (Bendig, 1953; Blum, 1936; Clark & Keller, 1954; Homes, 1972; Rayder, 1968; Remmers, 1930). On the other hand, studies by Anikeeff (1953), Kennedy (1975), Schuk and Crivelli (1973), Snyder and Clair (1976), Weaver (1960) and Walker (1974) have found that students' grades are significantly related to their evaluations of instructors. These relationships do not necessarily hold for all aspects of the evaluation. For example, in Homes' (1971) study, "relationships existed between the grades students expected and the degree to which they reported they were stimulated by the instructor and the degree to which they felt the grading system was fair" (p. 956), but expected grades were not related to items assessing the instructor's presentation.

Many faculty members who teach large classes suggest that they receive lower evaluations because students feel isolated from the teacher, that they are viewed as simply a number or an unknown face in the class. To compound the influence of class size on the evaluations, most large classes are also required, as opposed to elective courses. And this is viewed by the faculty as another variable that they have no control over and one which negatively influences their ratings. Research on these two class characteristics offers very little support for their beliefs. While Lovell and

Haner (1955) reported that instructors of classes over 30 received lower ratings, Mirus (1973) and others (Danielsen & White, 1976; Heilman & Armentrout, 1936) have reported that class size are teacher evaluations are positively related. Gage (1961), on the other hand, reported a curvilinear relationship; "teachers in courses with 30 to 39 students consistently received lower ratings than did those in courses with more or fewer students" (p. 17). Other researchers (Goodhartz, 1948; Guthrie, 1954; Solomon, 1966) have obtained results which indicate that there is no relationship between class size and teacher ratings.

The research on the influence of required vs elective courses on teacher evaluations is as inconclusive as most of the other research on the teacher, student and class characteristics. Some investigators (Gage, 1961; Lovell & Haner, 1955) have found that student ratings of teachers of required courses are significantly lower than those of teachers of elective courses. In contrast, Goodhartz (1948), Heilman and Armentrout (1936) and Mirus (1973) have reported no significant difference between the required and elective courses in terms of teacher evaluations.

METHOD

Design

In order to answer the questions posed in Chapter I, an experiment was designed in which two levels of each of the three independent variables will be studied simultaneously. Two of the independent variables, content and form, will be manipulated through the use of four videotapes; each tape will contain one level of each of these two variables. By blocking students into two classifications, natural science or social science, the third dependent variable, academic major, will be manipulated. This 2(high content, low content) x 2(good form, poor form) x 2(natural science, social science) factorial design is shown in Figure 2.

	High Content		Low Content	
	Good Form	Poor Form	Good Form	Poor Form
Natural Science				
Social Science				

Figure 2: Three-way (content x form x academic major) factorial design.

Subjects

Undergraduate students enrolled in the introductory psychology course at the University of Oklahoma will participate in this study to particially fulfill a course requirement. There will be 10 subjects in each of eight groups (.99 power against a one standard deviation difference at the .05 level of significance for a main effect with two levels). Students identifying themselves as majoring in either a natural science or a social science will voluntarily sign-up for this experiment. Half of the subjects in each of these groups will participate during one time period and the remaining subjects during a subsequent time period. During the first time period, the natural science subjects will be randomly assigned to either the high content-good form condition or the low content-good form condition. The social science subjects will also be randomly assigned to the same two conditions. During the second time period the subjects will be similarly assigned to either the high content-poor form condition or the low content-poor form condition.

Materials

Videotape equipment..

The following list of videotape equipment will be utilized in the production of the four videotape recordings.

Two (2) Panasonic color video cameras

Two (2) Sure low impedance microphones

One (1) Panasonic 5000 special effects generator

One (1) Panasonic color monitor

One (1) Sure 4-channel microphone mixer

One (1) Panasonic color video cassette recorder

Two (2) 60-minute Memorex color videotape cassettes

When the stimulus tapes are presented to the subjects, they will be played on a Panasonic color videotape player and viewed on a Panasonic color monitor.

Stimulus tapes..

Four color videotape recordings of a "teacher" lecturing a class will be used in this experiment. The topic of the lecture is the biochemistry of memory. The high content videotape recordings will contain 26 facts concerning the subject matter, while the low content videotapes will relate only four such facts. The recordings in which good form is demonstrated will present a teacher who is friendly, lacks distracting mannerisms and uses a well-modulated voice. The tapes which will present poor form will have a teacher who appears aloof, speaks in a monotone and exhibits several distracting mannerisms.

Tape 1. This stimulus videotape presents the condition of high subject content and good teaching style.

Tape 2. The condition presented in this recording is high subject content and poor teaching style.

Tape 3. The low subject content and good teaching style condition is presented in this videotape recording.

Tape 4. Presented in this recording is the condition of low subject content and poor teaching style.

In order to control as many extraneous variables as possible across the four videotapes, an actor (from the University of Oklahoma Drama Department) and verbatim scripts (see Appendix C) were used in the production of the videotapes. The scripts were developed by Ware and Williams at Southern Illinois University, Carbondale, and used by Ware in his doctoral dissertation (1974). The scripts, as well as the evaluation questionnaire and the test instrument, are used with his written consent (see Appendix B). Teacher exposure time to the subjects was an important variable to be considered. The goal for each videotape was 20 minutes. In general, the goal of equal exposure was accomplished. The average length of the videotape presentations is 19 minutes and 9 seconds. The presentations ranged from 17 minutes, 35 seconds (high content-poor form) to 20 minutes, 57 seconds (low content-good form). The duration of each videotape is presented in Table 7.

Before these tapes are used as stimuli for evaluation by the subjects, they will be reviewed and judged by four (4) judges (doctoral students in Educational Psychology or Counseling Psychology). The judges will rate each videotape recording using a six-item questionnaire (see Appendix D). The items are concerned with the lecturer's mannerisms, tone of voice, enthusiasm, sense of humor and interest in his students. Each item will be rated on a five-point scale.

Table 8

Duration of Videotaped Lectures in Minutes and Seconds

		Content: High	Content: Low
Form	Good	19:26	20:57
	Poor	17:35	18:39

<u>Evaluation questionnaire</u>..

The teacher will be evaluated using a modified form of the Instructor Improvement Questionnaire (IIQ) (Ware, 1974). Ware modified this instrument to make it appropriate for the evaluation of a lecturer after a single presentation. He selected eight of the original 40 items and added 11 more. The new items were added to increase the comprehensiveness of the evaluations in terms of the lecture content. The point-biserial correlations (item to total score) ranged from .60 to .82, with a median correlation of .75. An evaluation score is obtained by summing the responses to the 18 items (item 19 was omitted from the total score as unreliable). Ware also reported that the internal consistency (KR-20) for the 18-item modified IIQ instrument was .96.

The evaluation form which was used by the students to evaluate the "teacher" in this study is shown in Appendix E.

Test instrument..

The multiple choice test (see Appendix F) over the lecture topic (biochemistry of memory) was developed by Ware (1974). The test contains 26 four-choice items, one item for each of the facts presented in the high content lecture. The students were instructed to choose the best answer to each item and to guess when in doubt.

The point-biserial correlations (item to total score) ranged from .04 to .41 for the 26 test items; the median correlation was .19. A range of 19 to 66 was observed for item difficulties. The median difficulty level was 33 for the 26 items. Using the total number of correct items, a total test score was computed for each student. The internal consistency reliability (KR-20) for the 26-item test was reported by Ware to be .61.

Procedure

The academic major variable will be ignored at the time of data gathering. This will allow subjects who have natural science majors and those who have social science majors to view the tapes at the same time for each of the four conditions: high content-good form, high content-poor form, low content-good form and low content-poor form. All four groups will receive the following verbal instructions from the experimenter prior to their viewing the videotapes:

Many colleges within the University have been considering producing some materials to be used as supplements to classroom lectures and textbooks. Most of you are aware that supplemental readings in terms of books and articles are put on reserve in the library, and that materials are handed out in class. Another supplemental source of materials that is being investigated now is the use of videotapes. Many professors have considered putting selected materials on videotape to expand or supplement class lectures. These would be available to students on the same basis as selected readings in the library.

Some professors in association with the Instructional Services Center in the College of Education have produced videotapes directed toward this end. Before these tapes are made available to students, however, it is necessary that they be evaluated in terms of their usefulness. Because of your course work in psychology, you have been asked to participate in the evaluation of one of these videotapes. In keeping with their objective of only being supplemental material, they have been kept short. The videotape is approximately 20 minutes long.

After you have seen the tape, you will be asked to evaluate it. You will also be asked to take a 26-item test covering the lecture topic.

Please take your time and consider each item on the evaluation and test forms before you respond to it. You have as much time as you need to complete the forms. Do not put your name on the forms; your answers are confidential and will be known by no one but yourself.

Thank you for your time and your cooperation in participating in this evaluation.

Each group will then be instructed to complete Part I of the evaluation questionnaire. Part I includes demographic data (age, sex, highest grade completed, GPA, and academic major) and questions concerning the individuals feelings about the lecture topic (knowledge of the topic, expected grade in a course covering the topic and interest in the topic). After Part I of the questionnaire is completed by everyone, the videotaped lecture will be presented. At the completion of the lecture, the students will be instructed to complete Part II of the questionnaire. This part is concerned with the students satisfaction with the lecturer. After everyone has completed the evaluation, the group will be instructed to complete the test over the lecture topic.

After all the data have been collected, each group will be debriefed. The purpose of the study and the hypotheses will be explained to the students, and all questions pertaining to the study will be answered. To ensure that all subjects have an opportunity to be exposed equally to the lecture subject, a copy of the high content (poor form) script will be made available to each student in the two groups that viewed the low content videotape recordings.

Analysis

The first dependent measure for each subject will be obtained by summing the individual responses to each of the 18 evaluation items (evaluation score), and the second dependent measure by summing the subject's responses to the 26 items on the subject matter test (content score).

A multivariate analysis of variance technique will be used to test the hypotheses. The program OUMANOVA will be utilized in the analysis.

The amount of variance accounted for by the interaction effects and each of the main effects will be determined for each of the dependent variables.

References

American Association of University Professors, Committee C on College and University Teaching, Research, and Publication. Statement on teaching evaluation. _American Association of University Professors Bulletin_, 1975, _61_, 200-202.

American Educational Research Association, Committee on the Criteria of Teacher Effectiveness. Report of the committee on the criteria of teacher effectiveness. _Review of Educational Research_, 1952, _22_, 238-263.

Anikeeff, A. M. Factors affecting student evaluation of college faculty members. _Journal of Applied Psychology_, 1953, _37_, 458-460

Astin, A. W. & Lee, C. B. T. Current practices in the evaluation and training of college teachers. _Educational Record_, 1966, _47_, 361-375.

Bendig, A. W. The relation of level of course achievement to students' instructor and course rating in introductory psychology. Educational and Psychological Measurement, 1953, 13, 437-448.

Bendig, A. W. Ability and personality characteristics of introductory psychology instructors rated competent and empathetic by the students. Journal of Educational Research, 1955, 48, 705-709.

Berns, R. The Berkeley uprisings: A study of the conditions of riotous behavior. American Journal of Psychiatry, 1967, 123, 1165-1169.

Braunstein, D. N., & Benston, G. J. Student and department chairman views of the performance of university professors. Journal of Applied Psychology, 1973, 58, 244-249.

Braunstein, D. N., Klien, G. A., & Pachla, M. Feedback expectancy and shifts in student ratings of college faculty. Journal of Applied Psychology, 1973, 58, 254-258.

Brownell, W. A. Criteria of learning in educational research. In Improving educational research. American Educational Research Association, 1948.

Bryan, R. C. Student rating of teachers. Improving College and University Teaching, 1968, 16, 200-202.

Buxton, C. E. College teaching: A psychologist's view. New York: Harcourt Brace, 1956.

Caplow, T., & McGee, R. J. The academic marketplace. New York: Basic Books, 1956.

Centra, J. A. Effectiveness of student feedback in modifying college instruction. Journal of Educational Psychology, 1973, 65, 395-401 (a).

Centra, J. A. Do student ratings of teachers improve instruction? Change, 1973, 5(3), 12-13 (b).

Clark, K. E., & Keller, R. J. Student ratings of college teaching. In R. A. Eckert (Ed.), A university looks at its program. Minneapolis: University of Minnesota Press, 1954.

Cohen, S. A., & Berger, W. G. Dimensions of students' ratings of college instructors underlying sebsequent achievement of course examinations. Proceedings of the 78th Annual Convention of the American Psychological Association, 1970, 5, 605-606.

Costin, F., Greenough, W. T., & Menges, R. J. Student ratings of college teaching: Reliability, validity, and usefulness. Review of Educational Research, 1971, 41, 511-535.

Crannell, C. W. An experiment in the rating of instructors by their students. College and University, 1948, 24, 5-11.

Danielsen, A. L., & White, R. A. Some evidence on the variables associated with student evaluations of teachers. Journal of Economic Education, 1976, 7, 117-119.

Downie, N. W. Student evaluation of faculty. Journal of Higher Education, 1952, 23, 495-496, 503.

Drucker, A. J., & Remmers, H. H. Do alumni and students differ in their attitudes toward instructors? Journal of Educational Psychology, 1951, 42, 129-143.

Elliott, D. H. Characteristics and relationships of various criteria of college and university teaching. Purdue University Studies in Higher Education, 1950, 70, 5-61.

Evans, E. D. Student activism and teaching effectiveness: Survival of the fittest? Journal of College Student Personnel, 1969, 10, 102-108.

Feldman, R. A. Towards the evaluation of teaching competence in social work. Journal of Education for Social Work, 1972, 8, 5-15.

Frey, P. W., Lenard, D. W., & Beatty, W. W. Student ratings of instruction: Validation research. American Educational Research Journal, 1975, 12, 435-447.

Gage, N. L. The appraisal of college teaching: An analysis of end and means. Journal of Higher Education, 1961, 32, 17-22.

Gessner, P. Evaluation of instruction. Science, 1973, 180, 566-569.

Goodhartz, A. S. Student attitudes and opinions relating to teaching at Brooklyn College. School and Society, 1948, 68, 345-349.

Greenwood, G. E., & Renner, R. R. Student ratings of college teaching: Their validity and use in administrative decision-making. Science Education, 1975, 59, 493-498.

Guthrie, E. R. The evaluation of teaching. Educational Record, 1949, 30, 109-115.

Gustad, J. W. Policies and practices in faculty evaluation. _Educational Record_, 1961, _42_, 194-211.

Gustad, J. W. Evaluation of teaching performance: Issues and possibilities. In C. B. T. Lee (ed) _Improving college teaching_. Washington, D. C.: American Council on Education, 1967.

Hayes, J. R. Research, teaching and faculty fate. _Science_, 1971, _172_, 227-230.

Heilman, J. D., & Armentrout, W. D. The rating of college teachers on ten traits by their students. _Journal of Educational Psychology_, 1936, _27_, 197-216.

Holmes, D. S. The relationship between expected grades and students' evaluations of their instructors. _Educational and Psychological Measurement_, 1971, _31_, 951-957.

Holmes, D. S. Effects of grades and disconfirmed grade expectancies of students' evaluations of their instructor. _Journal of Educational Psychology_, 1972, _63_, 130-133.

Isaacson, R. L., McKeachie, W. J., & Milholland, J. E. Correlation of teacher personality variables and student ratings. _Journal of Educational Psychology_, 1963, _54_, 110-117.

Kennedy, W. R. Grades expected and grades received--their relationship to students' evaluations of their faculty performance. Journal of Educational Psychology, 1975, 67, 109-115.

Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, Ca: Brooks/Cole, 1968.

Kolevzon, M. S., & Wiltse, K. T. Student ratings and teacher effectiveness: A reappraisal. Journal of Education for Social Work, 1973, 9(2), 24-30.

Lasher, H., & Vogt, K. Student evaluation: Myth and realities. Improving College and University Teaching, 1974, 22, 267-269.

Lewis, D. R., & Orvis, C. C. A training system for graduate student instructors of Introductory Economics at the University of Minnesota. Journal of Ecomonic Education, 1973, 5, 27-34.

Lovell, G. D., & Haner, C. F. Forced-choice applied to college faculty rating. Educational and Psychological Measurement, 1955, 15, 291-304.

Luthans, F. The faculty promotion process: An empirical analysis of the administration of large universities. Iowa City: University of Iowa, 1967.

McClelland, J. N. The effect of student evaluations of college instruction upon subsequent evaluations. California Journal of Educational Research, 1970, 21, 88-95.

McKeachie, W. J. Student ratings of faculty. American Association of University Professors Bulletin, 1969, 55, 439-444.

McKeachie, W. J., Lin, Yi-Guang, & Mann, W. Student ratings of teacher effectiveness: Validity studies. American Educational Research Journal, 1971, 8, 435-445.

McNeil, J. D., & Popham, W. J. The assessment of teacher competence. In R. M. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand McNally, 1973.

Marsh, H. W., Fleiner, H., & Thomas, C. S. Validity and usefulness of student evaluations of instructional quality. Journal of Educational Psychology, 1975, 67, 833-839.

Maslow, A. H., & Zimmerman, W. College teaching ability, scholarly activity and personality. Journal of Educational Psychology, 1956, 47, 185-189.

Miller, M. T. Instructor attitudes toward, and their use of, student ratings of teachers. Journal of Educational Psychology, 1971, 62, 235-239.

Mirus, R. Some implications of student evaluation of teachers. Journal of Economic Education, 1973, 5, 35-37.

Mitzel, H. E. Teacher effectiveness. In C. W. Harris (Ed.), Encyclopedia of educational research (3rd ed.). New York: Macmillan, 1960.

Morsh, J. E., Burgess, G. G., & Smith, P. N. Student achievement as a measure of instructor effectiveness. Journal of Educational Psychology, 1956, 47, 79-88.

Naftulin, D. H., Ware, J. E., & Donnelly, F. A. The Dr. Fox lecture: A paradigm of educational seduction. Journal of Medical Education, 1973, 48, 360-365.

Neeley, M. A. Teacher's view of teacher evaluation. Improving College and University Teaching, 1968, 16, 207-209.

Null, E. J., & Walters, J. E. Values of students and their ratings of a university professor. College Student Journal, 1972, 6(4), 46-51.

Perry, R. R. Evaluation of teaching behavior seeks to measure effectiveness. College and University Business, 1969, 47, 18, 22.

Pohlmann, J. T. Evaluating instructional effectiveness with the instructor improvement questionnaire. Technical Report, Southern Illinois University, 1973.

Rader, N. F. College student ratings of instructors. Journal of Experimental Education, 1968, 37, 76-81.

Remmers, H. H. The measurement of teaching personality and its relation to learning process. Education, 1930, 51, 27-35.

Remmers, H. H. On students' perception of teachers' effectiveness. In W. J. McKeachie (Ed.), The apprasial of teaching in large universities. Ann Arbor: University of Michigan, 1958.

Remmers, H. H., & Elliot, D. N. The Indiana College and University Staff-Evaluation Program. School and Society, 1949, 70, 168-171.

Remmers, H. H., Martin, F. D., & Elliot, D. N. Are students' ratings of instructors related to grades? _Purdue University Studies in Higher Education_, 1949, _66_, 17-26.

Riley, J. W., Ryan, B. F., & Lifshitz, M. _The student looks at his teacher_. New Brunswick, N. J.: Rutgers University Press, 1950.

Rodin, M., Frey, P. W., & Gessner, P. K. Student evaluation. _Science_, 1975, _187_, 555-559.

Rodin, M., & Rodin, B. Student evaluations of teachers. _Science_, 1972, _177_, 1164-1166.

Rosenshine, B., Cohen, A., & Furst, N. Correlates of student reference ratings. _Journal of College Personnel_, 1973, _13_, 269-272.

Ryans, D. Prediction of teacher effectiveness. In C. W. Harris (Ed.), _Encyclopedia of educational research_ (3rd ed.). New York: Macmillan, 1960.

Saadeh, I. Q. Teacher effectiveness: A new direction in the evaluation of teaching. _Journal of Teacher Education_, 1970, _21_, 73-91.

Schulman, B. R., & Trudell, J. W. California's guidelines for teacher evaluation. Community and Junior College Journal, 1972, 43, 32-34.

Seldin, P., & Wakin, E. How deans evaluate teachers. Change, 1974, 6, 48-49.

Schuk, A. J., & Crivelli, M. A. Animadversion error in student evaluations of faculty teaching effectiveness. Journal of Applied Psychology, 1973, 58, 259-260.

Snyder, C. R., & Clair, M. Effects of expected and obtained grades on teacher evaluation and attribution of performance. Journal of Educational Psychology, 1976, 68, 75-82.

Slobin, D. Y., & Nichols, D. G. Student rating of teaching. Improving College and University Teaching, 1969, 17, 244-248.

Solomon, D. Teacher behavior dimensions, course characteristics, and student evaluations of teachers. American Educational Research Journal, 1966, 3, 35-47.

Starrack, J. A. Student rating of instruction. Journal of Higher Education, 1934, 5, 288-290.

Walker, B. D. Anticipated grade and student rating of teaching. Improving College and University Teaching, 1974, 22, 220-221.

Ware, J. E. The Doctor Fox Effect: An experimental study of the effectiveness of lecture presentations and the validity of student ratings, (Doctoral dissertation, Southern Illinois University, 1973). Dissertation Abstracts International, 1975, Ser A 35(7), 4266A-4267A, (University Microfilms No. 75-146).

Ware, J. E. & Williams, R. G. The Dr. Fox Effect: A study of lecturer effectiveness and ratings of instruction. Journal of Medical Education, 1975, 50, 149-156.

Weaver, C. H. Instructor rating by college students. Journal of Educational Psychology, 1960, 51, 21-25.

White, R. A. Some added support justifying administrative use of student evaluations of teachers. Journal of Economic Education, 1976, 7, 120-124.

Williams, R. G. & Ware, J. E. An extended visit with Dr. Fox: Validity of student satisfaction with instruction ratings after repeated exposures to a lecturer. American Educational Research Journal, 1977, 14, 449-457.

Williams, R. G. & Ware, J. E. Validity of student ratings of instruction under different incentive conditions: A further study of the Dr. Fox Effect. Journal of Educational Psychology, 1976, 68, 48-56.

Wilson, B. G. Faculty/course evaluation by participation. Business Education Forum, 1976, 30, 10, 12-13.

Withall, J. G. Development of a technique for the measurement of social-emotional climate in classrooms. Journal of Experimental Education, 1948, 17, 347-361.

Appendix B

Ware's Consent Letter

Rand
SANTA MONICA, CA. 90406

June 19, 1978

Mr. Jim Spence
2017 Elmhurst
Norman, Oklahoma 73071

Dear Mr. Spence:

As per our recent telephone conversation I am enclosing four reprints of articles pertaining to our "Doctor Fox" research. In addition you may be interested in the replication performed by Robert S. Meier; see his doctoral dissertation at Purdue University (1977).

You certainly have our permission to use the scripts and instruments from our research as you wish for purposes of your dissertation research. If I can be of any assistance in this regard let me know.

In case you are looking for ideas, I think it would be interesting to study whether student preferences regarding lecturer types (measured before viewing a lecture) interact with ratings and or learning (after the lecture). It may be that the so-called "Doctor Fox Effect" interacts with such preferences.

Good luck with your study.

Sincerely,

John E. Ware, Jr. (dj)

John E. Ware, Jr.

JEW/df

Enclosures

Appendix C

Videotape Scripts

Footnotes Regarding Changes Made in Converting From Good To Poor Form Scripts

Various changes were required in order to convert good form lecture scripts into poor form scripts. At points where changes were required, the changed material is enclosed in parentheses and the section is numbered. The nature of the change is explained by the following footnotes:

I. Footnotes Regarding Conversion from High Content-Poor Form to High Content-Good Form

1. Entry in parentheses omitted for Poor Form lecture.
2. Entry in parentheses omitted for Poor Form lecture.
3. Entry in parentheses omitted for Poor Form lecture.
4. Entry in parentheses omitted for Poor form lecture.
5. Entry in parentheses omitted for Poor Form lecture.
6. Entry in parentheses omitted for Poor Form lecture.
7. Substitute "eliminating him from the experiment" for the Poor Form lecture.
8. Entry in parentheses omitted for Poor Form lecture.
9. Entry in parentheses omitted for Poor Form lecture.

II. Footnotes Regarding Conversion from Low Content-Poor Form to Low Content-Good Form

1. Entry in parentheses omitted for Poor Form lecture.
2. Entry in parentheses omitted for Poor Form lecture.
3. Entry in parentheses omitted for Poor Form lecture.
4. Entry in parentheses omitted for Poor Form lecture.
5. Entry in parentheses omitted for Poor Form lecture.

6. Entry in parentheses omitted for Poor Form lecture.

7. Entry in parentheses omitted for Poor Form lecture.

8. Entry in parentheses omitted for Poor Form lecture.

9. Substitute "thus eliminating him from the experiment" for Poor Form lecture.

10. Entry in parentheses omitted for Poor Form lecture.

11. Entry in parentheses omitted for Poor Form lecture.

12. Entry in parentheses omitted for Poor Form lecture.

Lecture on the Biochemistry of Memory

These days, educators in America talk a great deal about the innovative hardware of education about computer-assisted instruction, 8mm cartridge-loading projectors, microtransparencies, and other devices. The emphasis is on devices, but tomorrow? Well, in the not too distant future they may well be talking about enzyme-assisted instruction, protein memory consolidators, antibiotic memory repellers, and the chemistry of the brain. That may be the phrase of the future--the chemistry of the brain. Although the psychologists' learning theories derived from the study of maze-running rats or target-pecking pigeons have failed to provide insights into the education of children, there is nothing wrong with learning theory--it's just inadequate. It is unlikely, very unlikely, that what is now being discovered by the psychologist, chemist, and neurophysiologist about rat-brain chemistry can deviate widely from what we will eventually discover about the chemistry of the human brain.

Consider this example, most adults, (even college professors who are not senile)[1] can repeat a series of seven numbers--8,4,8,8,3,9,9--immediately after the series is read. If, however, they are asked to repeat these numbers thirty minutes later, most will fail (including even the college professors--most will fail)[2]. Now in the first instance, we are dealing with the immediate memory span; and in the second, with long-term memory. These basic behavioral observations lie behind what is called the two-stage process theory.

According to a common variant of these notions, immediately after each learning trial--indeed, after every experience--a short-lived electrochemical process is established in the brain. This process, so goes the assumption, is the physiological mechanism which carries the short-term memory. Within a few seconds or minutes, however, this process decays and disappears; but before doing so, if all systems are go, the short-term electrochemical process triggers a second series of events in the brain. This second process is chemical in nature and involves, primarily, the production of new proteins and the induction of higher enzymatic activity levels in the brain cells. This process is more enduring and serves as the physiological substrate of our long-term memory.

It would follow that one approach to testing our theory would be to provide a subject with some experience or other, then interrupt the short-term electrochemical process immediately--before it has had an opportunity to establish the long-term process. If this were done, our subject should never develop a long-term memory for that experience.

Now at the Albert Einstein Medical school in New York, Dr. Murray Jarvik has devised a "step-down" procedure based on the fact that when a rat is placed on a small platform a few inches above the floor, the rat will step down onto the floor within a few seconds. Now the rat will do this consistently, day after day. Suppose that one day the floor is electrified, and stepping onto it produces a painful shock. When the rat is afterward put back on the platform--even 24 hours later--it will not budge from the platform; it will remain there until the experimenter gets tired and calls the experiment quits. That rat has thus demonstrated that he has a long-term memory for that painful experience.

Now, if we take another rat, but this time interfere with his short-term memory process immediately after he has stepped onto the electrified floor, the rat should show no evidence of having experienced a shock when tested the next day, since we have not given his short-term electrochemical memory process an opportunity to initiate the long-term protein-enzymatic process. To interrupt the short-term process, Jarvik passes a mild electric current across the brain of the animal. Now, the current is not strong enough to cause irreparable harm to the brain cells, but it does result in a very high level of activation of the neurons in the brain, thus disrupting the short-term electrochemical memory process. If this treatment follows closely enough after the animal's first experience with the foot shock, and we test the rat a day later, the rat acts as if there were no memory for yesterday's event; the rat jauntily and promptly steps down from the platform with no apparent expectation of shock.

When a long-term interval is interposed between the first foot shock and the electric current (through the brain) treatment, the rat does remember that foot shock, and it remains on the platform when tested the next day. This, again, is what we should have expected from our theory. The short-term electrochemical process has now had time to set up the long-term chemical memory process before it was disrupted.

Some well-known effects of accidental human injury seem to parallel these findings. Injuries which produce a temporary loss of consciousness but no apparent damage to brain tissue can cause the patient to experience a "gap" in his memory for the events just preceding the accident. This retrograde amnesia can be understood on the assumption that the events immediately prior to the accident were still being carried by the short-term memory processes at the time of the injury, and their disruption by the injury was sufficient to prevent the induction of the long-term memory processes. The patient asks "Where am I?" not only because he does not recognize the hospital, but also he cannot remember how he became injured.

Work, conducted by Dr. Bernard Agranoff at the University of Michigan Medical School, supports the hypothesis that the synthesis of new brain proteins is crucial for the establishment of the long-term memory process. He argues that if we could prevent the formation of new proteins in the brain, then--although the short-term electrochemical memory process is not interfered with-- the long-term memory process could never become established.

Much of Agranoff's work has been done with goldfish. Now he does this in this manner. The fish is placed in one end of a small rectangular tank, which is divided into two halves by a barrier which extends from the bottom to just below the surface of the water. When a light is turned on, the fish must swim across the barrier into the other side of the tank within twenty seconds--otherwise, he receives an electric shock. This training is continued for several trials until the animal learns to swim quickly to the other side when the light is turned on. Most goldfish learn this shock-avoidance task quite easily and they remember it for many days. Immediately before--and in some experiments, immediately after--training, Agranoff injects the antibiotic puromycin into the goldfish's skull. Now puromycin, as some of you may know, is a protein inhibitor and prevents the formation of new proteins in the brain's neurons. (If by any chance you did't know that before, you do know that now.)[3] After injection, Agranoff finds that the goldfish are not impaired in their acquisition of the shock-avoidance task, but, when tested a day or so later, they show almost no retention for the task.

These results seem to mean that the short-term memory process, which helps the animal remember from one trial to the next and thus permits him to learn in the first place, is not dependent upon the formation of new proteins; but that the long-term process which helps the animal remember from one day to the next and thus permits him to retain what he had learned, is dependent upon the production of new proteins. Again, as in the instance of Jarvik's rats, if the puromycin injection comes more than one hour after learning, it has no effect on later memory--the long-term memory process presumably has already been established and the inhibition of protein synthesis can no longer affect memory. In this antibiotic, therefore, we have found our first chemical memory erasure--or, perhaps more accurately, a chemical long-term memory preventative. Almost identical findings, by the way, have been reported by other workers in other laboratories working with such animals as mice and rats, which are far removed from goldfish.

Dr. Holger Hyden in Sweden also postulates that the creation of protein is essential to the formation of long-term memory. Dr. Hyden conducted experiments and established that when a normally left-handed rat is forced to learn to use his right hand to get his food, cells in the most highly developed part of the brain, the cortex, produced a special kind of substance called RNA as well as proteins. He discovered that stimulation of human neurons, nerve cells, causes some of the millions of ribonucleic acid, RNA, molecules inside them to give orders to glia cells to manufacture new proteins. The nature and pattern of these proteins contain an imprint of something that has been perceived and may become part of memory.

In fact, some of my colleagues have provided preliminary evidence suggesting that a learned response can be transferred from one animal to another by giving the untrained animal brain extract from the animal who has learned the response. (Now the possibilities of this brings to mind a story regarding one possible outcome of this work. Imagine, if you will, an intellectual supermarket. Now an individual walks up to the meat counter of this market and requests a pound of brains. The grocer inquires as to the kind desired. He indicates that he has monkey brains on special for $1.69 a pound, he has some human brains at $3.49 a pound, and professors' brains at $117.95 a pound. Now the customer asks him, "How come this discrepancy in price between human brains and professor brains?" Well, the grocer looks at the man and says, "My dear fellow, DO YOU REALIZE HOW MANY PROFESSORS IT TAKES TO MAKE A POUND OF BRAINS!" There may be a bit more truth in that particular story than many of us would like to think.)*

To go on with Dr. Hyden in Sweden, he has also determined that the injection of a reasonably pure extract of brain DNA into some animals produces a doubling in the rate of protein synthesis within an hour. This may have great implications for man since we know that output of RNA and this protein synthesis decreases with age! Now if it is determined that the type of protein produced by injection of DNA is a functionally valuable type and is a lasting effect, we may be on the way to procedures for decreasing senility!

It is important, of course, that correlates of some of these findings have been reported for humans. I am somewhat concerned when all of the evidence is from rat studies. Rats are (cute little devils but they are)[5] just rats! I am particularly concerned when findings from rat studies don't agree. I am even more concerned when the results with rats seem to differ according to the theoretical position of the researcher. A number of writers have talked about this expectance or experimenter bias effect. Well, maybe it is just that our rats on the East Coast are different from rats in California. However, I am reminded of the practice in some laboratories of eliminating from an experiment those rats that don't perform properly. (Do you know what I mean? (Reference to Audience) Well, I see some of you smiling so some of you do know what I mean.)[6] Well, what do you do with one rat out of 30 in an experiment who insists on always turning left in a maze? He keeps going left when all of his reinforcers say go right! Go right, young man, go right. Well, some scientists pick him up by the tail and they whack him on the side of the test stand (sending him to that "great maze" in the sky, where those left hand turns are permissible.)[7] So much for those kind of individual differences. Maybe that is why we have Gestalt rats in some places and Skinnerian rats in other places. Hey, have you every considered, we may be selectively breeding rats who learn and perform according to the prevalent theoretical position at a given university. (Now that's a thought!)[8]

To continue, thus far, ladies and gentlemen, I have been talking about disrupting or preventing the formation of memory as well as natural stimulation of the brain. Dr. James L. McGaugh of the University of California at Riverside has argued that injection of central nervous system stimulants such as strychnine, picrotoxin, or metrazol should enhance, fortify, or extend the activity of the short-term electrochemical memory processes and thus increase the probability that they will be successful in initiating long-term memory processes. From this it follows that the injection of CNS stimulants immediately before or after training should improve learning performance. That is precisely what McGaugh found--together with several additional results which have important implications for our concerns today.

In one of McGaugh's most revealing experiments, eight groups of mice from two different hereditary backgrounds were given the problem of learning a simple maze. Immediately after completing their learning trials, four groups from each strain were injected with a different dosage of metrazol--from none to five, 10 and 20 milligrams per kilogram of weight. First, it was apparent that there are hereidtary differences in learning ability--a relatively bright strain and a relatively dull one. Secondly, by properly dosing the animals with metrazol, the learning performance increased appreciably. Under the optimal dosage, the metrazol animals showed about 40 per cent improvement in learning ability over their untreated brothers. The improvement under metrazol, in fact, was so great, that the dull animals, when treated with 10 milligrams, did slightly better than their untreated but hereditarily superior colleagues.

In metrazol we not only have a chemical facilitator of learning, but one which acts as the "Great Equalizer" among hereditarily different groups. As the dosage was increased for the dull mice from none to five to 10 milligrams, their performance improved. Now beyond the 10 milligram point for the dull mice, however, and beyond the five-milligram point for the bright mice, increased strength of the metrazol solution resulted in a deterioration in learning. From this, we can draw two morals. First, the optimal dosage of chemical learning facilitators will vary greatly with the individual taking the drug. There is, in other words, an interaction between heredity and drugs. Second, there is a limit to the intellectual power of (even a hopped-up Southern California super)[9] mouse!

Now we already have available a fairly extensive class of drugs which can facilitate learning and memory in animals. A closer examination of McGaugh's results and the work of thers, however, also suggests that these drugs do not work in a monolithic manner on something called "learning" or "memory". In some instances, the drugs see to act on "attentiveness"; in some , on the ability to vary one's attacks on a problem; in some, on persistence; in some, on immediate memory; in some, on long-term memory. Now, think, different drugs work differentially for different strains, for different individuals, for different intellectual tasks, and for different learning components.

Do all these results mean that we will soon be able to substitute a pharmacopoeia of drugs for our various school-enrichment and innovative education programs, and that most educators will soon be technologically unemployed--or will have to retool and turn in their schoolmaster's gown for a pharmacist's jacket. No. No. The answer is no--as the Berkeley experiments on the influence of education and training on brain anatomy and chemistry suggests the opposite. Now, I will have much more to say about this in the lectures to follow in this Advancing Science Series on the Science of Memory.

But, we have made a start here. We have done some pretty impressive things in our animal laboratories. Given this start we need to go on with an intensive further program of research--with animals and with children--which seeks to spell out the further complexities between chemicals and memory. An understanding of these interrelationships, I feel, can pay off handsomely. Certainly we need not wait for the future. The future is now. The future is here. WE KNOW ENOUGH RIGHT NOW TO GET STARTED!!

The biochemist and the teacher of the future will combine their technologies for the educational and intellectual development of the child. Tommy needs a bit more immediate memory so we give him a stimulator; Jack could do with a chemical attention-span stretcher; Rachel needs an anticholinesterase to slow down her mental processes; Joan, some puromycin--she remembers too many details, and she gets lost.

Sure, our data has come from goldfish, and rats, and rodents, and worms. But is anyone so certain that the chemistry of the brain of, say, a rat which is after all a fairly complex mammal, is so different from that of the brain of a human being that he dare neglect this challenge---or even gamble--when the stakes are so high? We have so much to gain.

Think about it. (Pause)

Think about it.

Thank you.

Lecture on the Biochemistry of Memory

These days, educators in America are talking a great deal about the innovative hardware of education, computer assisted instruction, 8mm cartridge-loading projectors, microtransparencies, and other devices. In the not too distant future, they may well be talking about enzyme-assisted instruction, protein memory consolidators, antibiotic memory repellers, and the chemistry of the brain. The chemistry of the brain--perhaps the phrase of the future. Although the psychologists' learning theories derived from the study of maze-running rats or target-pecking pigeons have failed to provide insights into the education of children, there is nothing worong with learning theory. It is just inadequate. It is unlikely that what is now being discovered by the psychologist, the chemist, the neurophysiologist about rat-brain chemistry can deviate widely from what we will eventually discover about the chemistry of the human brain.

The data are not, however, restricted to findings from studies of sub-human animals. We have some very convincing direct and indirect evidence from studies of humans as well as clinical reports which, in some cases parallel the findings from animal laboratories very closely. I will have more to say about this later. I especially want to take a moment later to stress what I believe is a very important point, namely, the necessity of confirmation of experimental results. That is, we must proceed very cautiously until controversial findings from a given investigatror's laboratory are also reported by at least another scientist using the same methodology. I feel we must have some indication or strong trend suggesting that there is comparability between processes in animals and the implications of the process before we start yelling from the rooftops, so to speak. Therefore, I have chosen to focus on those teaching points in this lecture which relate to the work of a number of investigators. When there is a question, I will call that to your attention. When there are parallel data from direct and indirect evidence gathered regarding humans, I will call that to your attention. Although summeries are uniquely difficult in such a complicated area as this with results which are difficult to digest, I will try to put a handle on it all for you.

Now, as much as one might be inclined to assume that it is obvious what we mean when we say "memory", let me be simplistic and go over an example. Most adults, except those who are senile, (including even college professors,)[1] can repeat a series of seven numbers 8,4,8,8,3,9,9 immediately after the series is read. If, however, they are asked to repeat these numbers later, most (,including the college professors,)[2] will fail. As simple and as obvious as this example may seem it says something very important about the process of memory and the kinds of functions it serves. Evidence of the importance placed on the ability to retain digits and the differences which exist in people's ability to make such retention is found by analyzing current intelligence tests. Now, it is beyond the scope of this first segment of the lecture to go into the very interesting theoretical and empirical work on the relationship between memory and intelligence but, suffice it to say, for the purposes of my point, that memory span may be an important aspect of I. Q.

So what have I said? I have used a simple example to make the point that memory is a process. However, one should not be ready to accept the notion that the word process in any way gives us some magic insight into the intricate complexities of what that word so easily encompasses. As you will see from the studies I will cite, one must carefully work with the subtle aspects of the processes in order to tease out an understanding of the relationships which are involved.

Still another problem is the terminology employed in the study of memory which, like many other areas of science, is unique and holds little in common with the ways in which the vocabulary generally applies. For example, so-called "stages" of memory are often discussed in the literature of the biological sciences as well as in the psychology literature. These stages have been referred to by a variety of names--however, I want you to know that regardless of the terminology employed, the concepts are basically the same.

The fact that memory is a process--that is the point I want to make now--I don't want to belabor that. Let me conclude this part of my discussion by saying that the important facts to follow are those which have to do with the nature of this process. Various aspects of this nature will be pointed out in the context of findings from specific studies done in laboratories on the East coast as well as in the laboratories of my colleagues.

As I have suggested, important aspects of the process of memory are "teased" out under the highly controlled conditions of the experimental animal laboratory. Let me give you an example of such a strategy. One approach would be to interrupt the memory process as it begins. Now, if this is done experimentally, the subject probably will not remember what was underway at the time of the interruption. You have probably experienced this many times yourselves.

At the Albert Einstein School of Medicine which is located in the state of New York, a scientist by the name of Dr. Murray Jarvik has devised a procedure for conducting just such an investigation. His approach is essentially a "step down" procedure based on the fact that when a rat is placed on a small platform a few inches above the floor, the rat will step down onto the floor within a few seconds. Dr. Jarvik does not teach his rats this response but rather is using a response which occurs naturally in rats. The memory or learning which he is investigating is rather a modification of that response on the basis of what the rat experiences during a series of trials in which he steps down onto the floor. When the rat is placed on the platform during each trial he will step down a few seconds. The rat will do this consistently, day after day. Suppose that on one day the floor is electrified, and stepping onto it produces a painful shock. When the rat is afterward put back on the platform--even 24 hours later--it probably will not budge from the platform, but will remain there until the experimenter gets tired and calls the experiment quits. The rat has thus demonstrated retention, i.e., memory for that painful experience. To carry the study design a step further, Dr. Jarvik took rats that had learned <u>not</u> to step down on the electrified floor and he passed a mild electrical current through their brains. The current used is not strong enough to cause irreparable harm to the brain cells, but it does result in a very high level of activation of the neurons. Of course, there are many variations used in this type of research. For example, the interval of time between learning trials and the passage of the disruptive current across the brain is an important consideration in the results observed.

Some conclusions can be drawn from this type of study. Additional variations on this design are needed in order to address some of the pressing theoretical questions. Dr. Jarvik has done many of these experiments and his findings have been a real contribution.

A different approach is evidenced in the work of Dr. Agranoff. He has focused on still another aspect of the process of memory. His subjects are different--he uses goldfish--and his research hypothesis concerns the chemicals which are involved in retention. In a typical experiment, a fish is placed in one end of a small rectangular tank which is divided into two halves by a barrier which extends from the bottom of the tank to just below the surface of the water. When a light is turned on, the fish must swim across the barrier into the other side of the tank--within 20 seconds. Otherwise, he receives an electric shock. Now, this training is continued for several trials until the animal usually learns to swim quickly to the other side when the light is turned on. Goldfish learn the shock-avodiance task quite easily and they remember it for many days. Immediately before, and in some experiments, immediately after the training, Agranoff injects the antibiotic puromycin into the goldfish's brain.

Now, there have been other important studies. Dr. Holger Hyden in Sweden has conducted experiments in which normally left-handed rats have been forced to use their right hand to get food. Dr. Hyden has studied the effects of these experimental manipulations on the chemicals of the brain. He has concluded that the nature and pattern of these changes contain an <u>imprint</u> of something that has been perceived and may become a part of memory.

In fact, some of my colleagues have provided preliminary evidence suggesting that a learned response can be transferred from one animal to another by giving the untrained animal brain extract from the animal that has learned the response. (Now this particular experiment and conclusion brings to mind a story regarding one possible outcome of this work. Now, imagine, if you will, an intellectual supermarket. An individual walks up to the meat counter and requests a pound of brains. The grocer inquires as to the kind of brains he desires. Well, the customer indicates that he just wants a pound of brains, what kind have you got available. So the grocer says he has monkey brains on special for $1.69 a pound, he has some human brains at $3.49 a pound, and professors' brains at $117.95 a pound. "Well", the customer says, "why this huge discrepancy in price between human brains and professor brains?" And the grocer looked at him and said, "My dear fellow, DO YOU REALIZE HOW MANY PROFESSORS IT TAKES TO MAKE A POUND OF BRAINS!" Now, if I may return, perhaps more directly to the subject.)[3]

Dr. James McGaugh, who is at the University of California at Riverside, has done some very interesting work in this field. In one of his most revealing experiments, Dr. McGaugh studied a number of groups of mice from two different hereditary backrounds who were given the problem of learning a simple maze. Immediately after the learning trials, some groups were injected with a different dosage of a drug--from none to five, 10 and 20 milligrams per kilogram of weight. There were some fascinating findings related to the numerous different variables which were manipulated in the research. However, on the other hand, we know that we often plunge into this type of scientific pursuit and we learn some harsh lessons. Where the findings at first seem like a real gold mine, they turn out to be ashes, a can of worms. In the case of the latter study, it seems that the treatment solution that was at first beneficial, may actually become harmful to learning. Old Mother Nature, she's shown us once again that too much of a good thing is just that--too much. In other words, (contrary to a popular cartoon,)[4] there was even a limit to the intellectual power of a (hopped-up Southern California Super)[5] Mouse.

Today we already have several drugs to help learning and memory in animals. However, if we look at the work of some scientists we see that these drugs do not just work one way on mearning or memory. They work on various aspects and types of memory. The historical background of many of these investigations, with an evolving hypothetical farmework, further complicates the task of trying to interpret the results. You see, we are not dealing with a simple process here. Rather, the process is extremely complex. What I am trying to say is that unless some methodological flaw is discovered in the previous research---particularly in those studies in which the researcher reported negative findings---it seems that further investigation with regard to polymacromolecular aggregations should be directed to micro aspects of the cell and its membrane. I will have more to say about this, by the way, later in the presentation. I feel that the dozen or so years that we have under our belt can serve as the basis for some very definitive statements about the directions we should take in the future and the priorities which should be set. This, let us face it, is not a simple task.

But, let me share with you for a moment some of the frustrations of a colleague of mine who spent four decades studying memory processes through the experimental induction of lesions in the brains of laboratory animals. Primarily, these were rats. In an attempt to locate the sites of memory storage, he selectively cut brain tissue in live animals every way possible. He destroyed areas around suspected memory sites as well as destroyed memory sites while leaving surrounding areas intact. The results were not very conclusive; in fact, they were downright confusing. One might even conclude from this that memory is really not possible. Seriously, however, we know it is. In fact, the irony of it all is that we are using this illusive process at the very moment we are thinking about it. Perhaps, as some ancient philosophers may have thought, our minds don't want to reveal their inner-most secrets.

From a purely mechanical point of view, memory can be thought of as a physical change. Although the question may, at first, sound absurd, let me ask whether or not windows have memories. Does a window remember the times during the year when decorations are placed on it? Well, there are some residual tape markings on a window after you tape a paper pumpkin on it. However, these are removed with time, with a razorblade, or the wind or rain if the markings are on the outside. Maybe that's memory--it just isn't permanent memory. Does a window remember it when a rock is thrown through it--or a baseball? How about the time when Johnny threw a baseball through the kitchen window. Now there is a memory--a memory for all of us including maybe even the window. Certainly a permanent change resulted in the window as a result of that experience. Maybe this kind of structural change is analogous to memory. I really don't know--it's just an idea--a thought I wanted to share with you. I thought that maybe says something about how memory works. What do you think?

It is important that correlates of some of these findings be reported from some of these studies of other animals and also from studies of humans. I have a tendency to become somewhat concerned when all of the evidence is from rat studies. Rats are (cute little devils but they are)[6] just rats. I am particularly concerned when findings from rat studies don't agree. I am even more concerned when results with rats seem to differ according to the theoretical position of the researcher. A number of writers have talked about this expectance or experimenter bias effect. But, anyway, maybe it's just that our rats on the East Coast are different from rats in California. However, I am reminded of the practice in some laboratories of eliminating from an experiment those rats that don't perform properly. (You know what I mean? Well, I see some of you are smiling. Some of you do.)[7] Well, what do you think? What do you do with one rat of 30 in an experiment who insists on always turning left in a maze? He keeps going left when all the reinforcers say go right. (Go right, young man, go right.)[8] Well, sometimes you pick him up by the tail and whack him on the side of the test stand (sending him to that great maze in the sky, where those left hand turns are permissible.)[9]

So much for those kinds of individual differences. You know that may be why we have Gestalt rats some places and Skinnerian rats in other places. Have you considered that we may be selectively breeding rats who learn according to the prevalent theoretical position of a given university? (Now, there's a thought.)[10]

There has been additional work, much additional work. Work conducted by Dr. Bernard Agranoff at the University of Michigan Medical School supports the hypothesis that the synthesis of new brain protein is crucial for the establishment of a long-term memory process. Now he argues that if we could prevent the formation of new proteins in the brain, then although the short-term electro-chemical memory process is not interfered with, the long-term memory process could never become established.

What conclusions can we draw? Do all these results mean that we will soon be able to substitute a pharmacopoeia of drugs for our various schools' enrichment and innovative educational programs, that most educators will soon be technologically unemployed or they will have to turn in their school master's gown for a pharmacist's jacket? (No, no, do not live in fear,)[11] The answer is no. As the Berkeley experiments on the influence of education and training on brain anatomy and chemistry suggest, just the opposite. Now, with regard to this, I will have much more to say during subsequent lectures produced in this advancing series by the Advancing Science Series on Memory.

But--for a moment--let me make it clear that we have made a start here. We have done some pretty impressive things in our animal laboratories. Given this start, we need to go on with an intensive further program of research--with animals and with children--which seeks to spell out the further complexity between chemicals and memory. Of course, we need some more government money. However, it is a good investment. An understanding of these interrelationships, I feel, can pay off handsomely. Certainly, we need not wait for the future. The future is now. We know enough right now to get started. The bio-chemist and the teacher of the future will combine their technologies for the education and intellectual development of the child. Tommy needs a bit more immediate memory so we give him a stimulator; Jack could do with a chemical attention-span stretcher; Rachel needs a anticholinesterase to slow down here mental process; Joan needs some puromycin--she remembers too many details and she gets lost.

Sure, our data have come from goldfish and rats and rodents and even worms. But, is anyone so certain that the chemistry of the brain of say a rat, which after all is a fairly complex mammal, is so different from that of a brain of a human being, that he dare neglect this challenge or even gamble when the stakes are so high. We have so much to gain.

Think about it. (Do it. Think about it.)[12] (Pause)

Thank you.

Appendix D

<u>Videotape Evaluation Form</u>

Videotape Evaluation Form

The effectiveness of the classroom teacher in communicating his course content is important to both himself and his students. This effectiveness is a function of a number of variables, and among these is "teaching style". Videotape simulations of teaching styles have been produced to investigate several components of teaching style: tone of voice, mannerisms and friendliness. Before these videotapes can be used for research purposes, however, the researcher needs to know how the teaching styles are perceived by the viewer. You will be shown four recordings. After each recording, you will be asked to rate the teacher using a six-item questionnaire. Each item will be rated on a five-point scale.

Please rate the first videotape recording using the evaluation form on Sheet I; rate the second recording on Sheet II; the third on Sheet III; the fourth on Sheet IV.

Please take your time and consider each item on the evaluation forms before you respond to it.

Thank you for your time and your consideration in the evaluation of these videotape recordings.

Sheet I

<u>LECTURER</u> <u>EVALUATION</u>: Below you will find a series of statements about the lecturer in the videotape you have just seen. Please read each statement carefully, and respond to each using the following scale:

SA - I <u>strongly</u> <u>agree</u> with the statement.
A - I <u>agree</u> with the statement.
N - <u>Neutral</u>, I nether agree nor disagree .
D - I <u>disagree</u> with the statement.
SD - I <u>strongly</u> <u>disagree</u> with the statement.

Indicate your response to each statement by circling the corresponding code to the left of the statement number.

THE LECTURER

SA A N D SD 1. Exhibited distracting mannerisms.

SA A N D SD 2. Spoke in a monotone.

SA A N D SD 3. Was enthusiastic about the subject.

SA A N D SD 4. Has a sense of humor.

SA A N D SD 5. Was interested in his viewers.

SA A N D SD 6. Made learning enjoyable.

Sheet II

LECTURER EVALUATION: Below you will find a series of statements about the lecturer in the videotape you have just seen. Please read each statement carefully, and respond to each using the following scale:

SA - I strongly agree with the statement.
A - I agree with the statement.
N - Neutral, I nether agree nor disagree .
D - I disagree with the statement.
SD - I strongly disagree with the statement.

Indicate your response to each statement by circling the corresponding code to the left of the statement number.

THE LECTURER

SA A N D SD 1. Exhibited distracting mannerisms.

SA A N D SD 2. Spoke in a monotone.

SA A N D SD 3. Was enthusiastic about the subject.

SA A N D SD 4. Has a sense of humor.

SA A N D SD 5. Was interested in his viewers.

SA A N D SD 6. Made learning enjoyable.

Sheet III

LECTURER EVALUATION: Below you will find a series of statements about the lecturer in the videotape you have just seen. Please read each statement carefully, and respond to each using the following scale:

SA - I strongly agree with the statement.
A - I agree with the statement.
N - Neutral, I nether agree nor disagree .
D - I disagree with the statement.
SD - I strongly disagree with the statement.

Indicate your response to each statement by circling the corresponding code to the left of the statement number.

THE LECTURER

SA A N D SD 1. Exhibited distracting mannerisms.

SA A N D SD 2. Spoke in a monotone.

SA A N D SD 3. Was enthusiastic about the subject.

SA A N D SD 4. Has a sense of humor.

SA A N D SD 5. Was interested in his viewers.

SA A N D SD 6. Made learning enjoyable.

Sheet IV

LECTURER EVALUATION: Below you will find a series of statements about the lecturer in the videotape you have just seen. Please read each statement carefully, and respond to each using the following scale:

SA - I strongly agree with the statement.
A - I agree with the statement.
N - Neutral, I nether agree nor disagree .
D - I disagree with the statement.
SD - I strongly disagree with the statement.

Indicate your response to each statement by circling the corresponding code to the left of the statement number.

THE LECTURER

SA A N D SD 1. Exhibited distracting mannerisms.

SA A N D SD 2. Spoke in a monotone.

SA A N D SD 3. Was enthusiastic about the subject.

SA A N D SD 4. Has a sense of humor.

SA A N D SD 5. Was interested in his viewers.

SA A N D SD 6. Made learning enjoyable.

Appendix E

Evaluation Questionnaire

PART I. DIRECTIONS: Please do not write your name anywhere on this questionnaire. We want you to feel free to respond honestly and frankly. However, for statistical purposes we need to know a few things about you. These data will be analyzed on a group basis.

1. What is your age (in years)? _________

2. What is your sex? _____Female _____Male

3. Circle the highest number of years you have completed during your schooling.

9 10 11 12	13 14 15 16	17 18 19 20
High school	College	Post graduate

4. Mark the one category which best represents your overall undergraduate GPA (4.0 = A).

_____3.6 to 4.0 _____2.6 to 3.0

_____3.1 to 3.5 _____1.0 to 2.5

5. Write your academic major in the space provided.

____

Lecture topic: The Biochemistry of Memory.

We are interested in what you think about the lecture topic. Please answer the questions below. Place an X above the one answer to each question which best represents how you feel.

6. How much do you think you know about this topic?

:________:________:________:________:________:

A great deal	Quite a bit	Some	Little	Nothing

7. If you took an introductory course on the topic, what grade would you expect?

:________:________:________:________:________:

A	B	C	D	F

8. How interested are you in this topic?

:________:________:________:________:________:

Extremely Interested	Very Interested	Somewhat Interested	No Interest At All	Annoyed By It

STOP! PLEASE DO NOT GO TO PAGE 2 UNTIL TOLD TO DO SO.

PART II. LECTURER EVALUATION, DIRECTIONS: Below you will find a series of statements about the lecturer in the videotape you just saw. Please read each statement carefully, and respond to each using the following scale:

E - Exceptional performance
V - Very good performance
G - Good performance
W - Weak performance
I - Improvement definitely needed

Indicate your response to each statement by circling the corresponding code to the left of the statement number.

THE LECTURER:

E V G W I 1. Spoke understandably.

E V G W I 2. Knew if students understood him.

E V G W I 3. Showed an interest in students.

E V G W I 4. Increased your appreciation for the subject.

E V G W I 5. In general, taught effectively.

E V G W I 6. Gave several examples to explain complex ideas.

E V G W I 7. Knew his subject matter.

PART II. LECTURER EVALUATION (CONTINUED): Below you will find another series of statements about the lecturer in the videotape you just saw. Please read each statement carefully, and respond to each using the following scale:

SA - I strongly agree with the statement.
A - I agree with the statement.
N - Neutral, I nether agree nor disagree .
D - I disagree with the statement.
SD - I strongly disagree with the statement.

Indicate your response to each statement by circling the corresponding code to the left of the statement number.

THE LECTURER

SA A N D SD 8. Stressed important material.

SA A N D SD 9. Was as effective lecturer.

SA A N D SD 10. Has a good sense of humor.

SA A N D SD 11. Organized and presented subject matter well.

SA A N D SD 12. Inspired confidence in his knowledge of the subject.

SA A N D SD 13. Broadened my interest in the subject.

SA A N D SD 14. Explained the subject clearly.

SA A N D SD 15. Increased my knowledge of the subject.

SA A N D SD 16. Stimulated my thinking.

SA A N D SD 17. Was enthusiastic about the subject.

SA A N D SD 18. Made learning enjoyable.

SA A N D SD 19. Often left me confused.

STOP! PLEASE DO NOT GO TO PAGE 4 UNTIL TOLD TO DO SO.

Appendix F

Test Instrument

THE BIOCHEMISTRY OF MEMORY

<u>DIRECTIONS</u>: This is a quiz over the lecture on the Biochemistry of Memory. Please answer the questions to the best of your ability. No one is expected to get all answers correct. If you are in doubt about the answer to a question, then guess.

Please place your answers on the computer-scored answer sheet which has been provided. Please do not mark on the test form. The numbers running scross the page on your answer sheet corresponds to the items on your test form. Indicate your choice for each item in the following manner. If your choice to Item 30 is the first of the five available choices, mark "A" as shown.

30. X B C D E

All responses on the answer sheet must be made using a pencil. Please choose the one best answer for each item.

1. According to the theory presented in the lecture, immeditely after every learning trial or experience
   A. a permanent change in the brain tissue occurs.
   B. a temporary electrochemical process is established in the brain.
   C. new brain cells are generated.
   D. no observable change in the brain is expected.

2. The two stages in the two-stage process of theory of learning are
   A. input and output.
   B. short term memory and long term memory.
   C. encoding and decoding.
   D. processing and storage.

3. The two-stage process theory states that
   A. the production of new proteins and the induction of higher ensymatic activity levels in the brain helps explain short term memory.
   B. no physiological or chemical changes occur in the brain until long after every learning trial or experience.
   C. an electrochemical process which triggers a series of events in the brain helps explain short term memory.
   D. short and long term memory are not accounted for by the two-stage theory at all.

4. Humans can retain knowledge of an event in short term memory for periods:
   - A. up to a few minutes.
   - B. up to a half hour.
   - C. varying from 15 minutes to a half hour depending on age.
   - D. varying from 5 minutes to 15 minutes depending on sex.

5. According to the two-stage process theory of memory
   - A. long term memory is independent of short term memory.
   - B. disruption of short term memory mildly inhibits the establishment of long term memory.
   - C. disruption of short term memory prevents the establishment of long term memory.
   - D. disruption of long term memory also disrupts short term memory.

6. Dr. Jarvik passed a mild electric current across the brain of a rat immediately following a learning event with the result that
   - A. the long term memory processes of the rat were destroyed.
   - B. the rat's memory was in no way affected.
   - C. the short term memory process of the rat was interrupted.
   - D. the short term memory ability of the rat was facilitated.

7. When Dr. Jarvik passed a mild electric current across the brain of the rat following a long interval since the learning event:
   - A. the rat's memory for the event was facilitated.
   - B. the rat's memory for the event was inhibited.
   - C. the rat's memory for the event was unaffected.
   - D. the rat exhibited a high level of emotional behaviors.

8. The term "retrograde amnesia" means
   - A. inability to recall events even though they are still stored in memory.
   - B. permanent loss of memory for early events in one's life.
   - C. loss of memory for specific traumatic events from one's past.
   - D. hypnosis induced memory loss.

9. Retrograde amnesia appears to be the result of
   A. disruption of short term memory.
   B. disruption of long term memory.
   C. disruption of glial cell production.
   D. deterioration of synaptic nodes.

10. Which of the following statements best summarizes observations following accidental human injury as they pertain to the two-stage process theory?
   A. gaps in memory following temporary loss of consciousness in humans cannot be explained by the two-stage process theory.
   B. well-known effects of accidental human injury parallel animal findings.
   C. there is no relationship because loss of consciousness does not occur in humans without permanent damage to brain tissue.
   D. short term memory loss in humans can be explained but not permanent loss.

11. Long term memory appears to be a result of
   A. the production of new proteins and the induction of higher enzymatic activity.
   B. the production of neuronal DNA and the ensuing induction of higher enzymatic activity.
   C. the storage of electrical representation of events in a form similar to morse code.
   D. a holographic representation of events which is stored and recalled electrically.

12. Dr. Agranoff, working with goldfish, held the hypothesis that
   A. the production of new glia cells is required for the establishment of long term memory.
   B. preventing the formation of new proteins in the brain has no measurable effect.
   C. the systhesis of new brain proteins is crucial for the establishment of long term memory.
   D. the systhesis of new brain proteins is crucial for the establishment of both short and long term memory.

13. Dr. Agranoff found that the antibiotic puromycin
   A. improved short term memory.
   B. improved long term memory.
   C. interfered with the development of long term memory.
   D. interfered with the development of short term memory.
   E. none of the above are accurate statements.

14. Agranoff's position was based on findings in which injection of puromycin into the goldfish's brain:
    A. allowed learning of a shock avoidance task but resulted in forgetting a day or so later.
    B. prevented immediate memory of a previously learned task but resulted in recall of that task a day or so later.
    C. prevented immediate and long term memory for a task.
    D. facilitated short term memory when small doses (less than 10 milligrams) were injected but inhibited short term memory in larger doses.

15. Agranoff found that injection of puromycin one hour or more after learning:
    A. produced a decrease in long term memory.
    B. produced an increase in long term memory.
    C. had no effect on long term memory.
    D. produced a small increase in short term memory.

16. The research findings of Agranoff have indicated that
    A. long term memory will follow immediate memory and researchers cannot interfere with this natural process once it has been initiated.
    B. immediate and long term memory connot be distinguished from each other.
    C. all forgetting is a function of a failure in the long term memory phase of intellectual processing.
    D. the formation of long term memory can be artificially stopped.

17. The results of experiments in which goldfish were injected in the brain with a chemical that inhibits formation of new protein, show that
    A. protein is not involved in the creation of memory.
    B. protein is involved in the creation of memory.
    C. protein is involved in the formation of long term memory.
    D. protein is involved in the creation of short term memory.