Negative microfilm
received from university

66-6710

STAKENAS, Robert George, 1929-
EVALUATIVE STRESS, FEAR OF FAILURE AND ACADEMIC ACHIEVEMENT.

The University of Michigan, Ph.D., 1965
Education, psychology

University Microfilms, Inc., Ann Arbor, Michigan

EVALUATIVE STRESS, FEAR OF FAILURE AND ACADEMIC ACHIEVEMENT

by
Robert George Stakenas

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in the
University of Michigan
1965

Doctoral Committee:

Professor Wilbert J. McKeachie, Chairman
Professor Warren A. Ketcham
Professor John E. Milholland
Doctor Karl Lewis Zinn

ACKNOWLEDGMENTS

Research of this nature necessarily requires the cooperation and support of many persons. With this fact in mind, I would like to acknowledge my gratitude to my thesis committee: to Dr. McKeachie, the chairman, for his assistance and expert advice in dealing with this problem; to Dr. Milholland for his help in carrying out the statistical analysis; to Dr. Zinn for preparing computer programs for scoring the raw data; to Dr. Ketcham for his willingness to serve as a replacement for Dr. Morse; and to Dr. Morse for serving on the committee until his departure from Ann Arbor, and for his support and advice over the years.

One manifest outcome of this thesis is the training in research afforded me by the experience. A major contribution in this regard was made by the Research Project on the Characteristics of Effective College Teaching (Research Contract O.E. 4-10-001), which provided me with the opportunity and the material support to carry out the research presented in this thesis. Project staff members whose assistance is also gratefully acknowledged are: Dr. Yi-Guang Lin who served as a continuing source of information on the use of the computer for data analysis; Amy (Konopacki) Lewis who assisted in preparing and distributing the questionnaires; my two faithful and hard working assistants Morgen Griffith and Jane Golden who carried out the multitude of clerical tasks

set before them; and Audrey Warren and Ferne Galantai who assisted in typing the initial draft.

Thanks are also due Dr. Daniel Fusfeld, coordinator of Introductory Economics, Dr. Richard Mann, coordinator of Introductory Psychology, and to the teaching fellows who assisted in gathering data as well as serving as subjects.

The efficient work of Alice Sano in typing the final draft is also appreciated.

A special debt of gratitude is owed to my wife, Margaret, for her assistance in typing portions of the first draft, proofreading, and for being a source of support and encouragement during my tenure as a graduate student and in completing this thesis.

TABLE OF CONTENTS

		Page
ACKNOWLEDGMENTS		ii
LIST OF TABLES		vi
LIST OF APPENDICES		xi
INTRODUCTION		1
Chapter		
I	A REVIEW OF RELEVANT LITERATURE	3
	The Measurement of Anxiety	3
	Experimental Studies of Anxiety	5
	Correlational Studies	17
	The Experimental Approach in Real Life Situations	25
	Summary	30
II	THEORY AND HYPOTHESES	32
	The Problem	32
	The Concept of Anxiety	33
	Anxiety and Academic Achievement	35
	Achievement and Fear of Failure	36
	Hypotheses	38
	The Variables	41
	Fear of Failure	41
	Social Constraint	42
	Evaluative Stress	42
	Study Habits and Persistence	42
	Academic Performance	43
III	THE SURVEY STUDY	44
	Procedure	44
	Course Settings	44
	Student Populations	45
	Teacher Samples	46

Chapter Page

Measures of Student Characteristics 46
Measures of Teacher Behavior 47
Data Collection 48
Coding of Open-ended Items 49
Levels for the Treatments x Levels Design 51
Standardization of Grades 53

Results 53

Intercorrelations of Student Characteristics . . . 53
Teacher Ratings 53
The Analysis of Variance 59

Discussion 63
Summary 66

IV THE EXPERIMENTAL STUDY 68

Procedure 68

Student Subjects 68
Measures 69
Data Collection 69

Results 72
Discussion 76
Summary 83

V AN OVERVIEW 84

Implications for Future Research 86

APPENDICES 90

REFERENCES 179

LIST OF TABLES

Table		Page
1	Intercorrelations of Three Measures of Anxiety	18
2	Intercorrelation of Variables for 370 Men and 577 Women in Introductory Psychology (Fall, 1964)	54
3	Intercorrelation of Variables for 338 Men and 149 Women in Introductory Economics (Fall, 1964)	54
4	Mean Item Ratings for Item One	55
5	Mean Item Ratings for Item Two	55
6	Mean Item Ratings for Item Three	56
7	Mean Item Ratings for Item Five	56
8	Mean Item Ratings for Item Four	57
9	Frequency of Teacher Assigned Tests	58
10	Intercorrelation of Item Means for Teacher Produced Evaluative Stress	59
11	Mean Grade by Level of Ability for Students in Psychology and Economics (Fall, 1964)	60
12	Mean Grade by Level of Study Habits for Students in Psychology and Economics (Fall, 1964)	61
13	Mean Grade by Level of Debilitating Anxiety for Students in Psychology and Economics (Fall, 1964)	62
14	Zero-Order and Multiple Correlations for Ability, Study Habits, and Debilitating Anxiety in Predicting Grades in Psychology and Economics (Fall, 1964)	64
15	Intercorrelation of Variables for 101 Men and 108 Women in Introductory Psychology, Criteria Test Given Under Non-evaluative Conditions (Fall, 1964)	73

Table		Page
16	Intercorrelation of Variables for 96 Men and 99 Women in Introductory Psychology, Criteria Test Given Under Evaluative Conditions (Fall, 1964)	73
17	Mean Criteria Test Scores for Levels of Ability (Fall, 1964)	74
18	Mean Criteria Test Scores for Levels of Evaluative Stress (Fall, 1964)	75
19	Mean Criteria Test Scores for Levels of Debilitating Anxiety and Stress (Fall, 1964)	75
20	Mean Criteria Test Scores for the Interaction of Study Habits and Debilitating Anxiety (Fall, 1964)	76
21	Mean Scores for the Interaction of Debilitating Anxiety and Study Habits for Easy and Difficult Criteria Test Items, Men in Psychology (Fall, 1964)	81
22	Mean Scores for the Interaction of Debilitating Anxiety and Study Habits for Easy and Difficult Criteria Test Items, Women in Psychology (Fall, 1964)	81
23	Mean Scores for the Interaction of Debilitating Anxiety and Evaluative Stress for Difficult Criteria Test Items, Men and Women in Psychology (Fall, 1964)	82
24	Class Standing of Men and Women in Psychology and Economics, The Survey Study (Fall, 1964)	91
25	Means and Standard Deviations for Measures of Ability, Study Habits, Debilitating Anxiety and Grades for 370 Men and 577 Women in Introductory Psychology, The Survey Study (Fall, 1964)	125
26	Means and Standard Deviations for Measures of Ability, Study Habits, Debilitating Anxiety, and Grades for 338 Men and 149 Women in Introductory Economics, The Survey Study (Fall, 1964)	126
27	Means and Standard Deviations for Measures of Ability, Study Habits, Debilitating Anxiety, and Criteria Test Scores for Men, The Experimental Study (Fall, 1964) . .	126

Table		Page
28	Means and Standard Deviations for Measures of Ability, Study Habits, Debilitating Anxiety, and Criteria Test Scores for Women, The Experimental Study (Fall, 1964) .	127
29	Reliability Coefficients for the Student Measures	128
30	Scale Scores for Levels of Ability, Debilitating Anxiety, and Study Habits for Men in Psychology and Economics, The Survey Study (Fall, 1964)	129
31	Scale Scores for Levels of Ability, Debilitating Anxiety, and Study Habits for Women in Psychology, The Survey Study (Fall, 1964)	130
32	Scale Scores for Levels of Ability, Debilitating Anxiety, and Study Habits for Men in Psychology, The Experimental Study (Fall, 1964)	131
33	Scale Scores for Levels of Ability, Debilitating Anxiety, and Study Habits for Women in Psychology, The Experimental Study (Fall, 1964)	132
34	Analysis of Variance for Item Three, Men in Psychology (Fall, 1964)	133
35	Analysis of Variance for Item Four, Men in Psychology (Fall, 1964)	134
36	Analysis of Variance for Item Five, Men in Psychology (Fall, 1964)	135
37	Analysis of Variance for Teacher Assigned Tests, Men in Psychology (Fall, 1964)	136
38	Analysis of Variance for Item Three, Men in Economics (Fall, 1964)	137
39	Analysis of Variance for Item Four, Men in Economics (Fall, 1964)	138
40	Analysis of Variance for Item Five, Men in Economics (Fall, 1964)	139
41	Analysis of Variance for Frequency of Teacher Assigned Tests, Men in Economics (Fall, 1964)	140

Table		Page
42	Analysis of Variance for Item Three, Women in Psychology (Fall, 1964)	141
43	Analysis of Variance for Item Four, Women in Psychology (Fall, 1964)	142
44	Analysis of Variance for Item Five, Women in Psychology (Fall, 1964)	143
45	Analysis of Variance for Frequency of Teacher Assigned Tests, Women in Psychology (Fall, 1964)	144
46	Mean Grade by Level of Ability for Students in Psychology and Economics (Fall, 1964)	145
47	Mean Grade by Level of Study Habits for Students in Psychology and Economics (Fall, 1964)	146
48	Mean Grade by Level of Debilitating Anxiety for Students in Psychology and Economics (Fall, 1964)	147
49	Analysis of Variance for Experimental Conditions, Men in Psychology (Fall, 1964)	148
50	Analysis of Variance for Experimental Conditions, Women in Psychology (Fall, 1964)	149
51	Per Cent Passing for the Ten Easiest and Ten Most Difficult Items on the Introductory Psychology Criteria Test, Form X (Spring, 1962; N = 64-67)	151
52	Mean Scores by Level of Ability for Ten Easy and Ten Difficult Criteria Test Items, Men and Women in Psychology (Fall, 1964)	152
53	Mean Scores by Conditions for Ten Easy and Ten Difficult Criteria Test Items, Men and Women in Psychology (Fall, 1964)	152
54	Mean Criteria Test Scores for the Interaction of Ability and Conditions, Men and Women in Psychology (Fall, 1964)	153
55	Analysis of Variance for Ten Easy Criteria Test Items, Men in Psychology (Fall, 1964)	157

Table | Page

56 Analysis of Variance for Ten Difficult Criteria Test Items, Men in Psychology (Fall, 1964) 158

57 Analysis of Variance for Ten Easy Criteria Test Items, Women in Psychology (Fall, 1964) 159

58 Analysis of Variance for Ten Difficult Criteria Test Items, Women in Psychology (Fall, 1964) 160

LIST OF APPENDICES

Appendix		Page
A	CHARACTERISTICS OF THE STUDY SAMPLES BY COURSE, YEAR, AND SEX	91
B	INSTRUCTIONS TO ECONOMICS AND PSYCHOLOGY INSTRUCTORS, AND THE LETTER OF INTRODUCTION	92
C	MEASURES OF STUDENT CHARACTERISTICS	96
D	EVALUATIVE STRESS ITEMS AND THE SUBJECTIVE REPORT QUESTIONNAIRE	108
E	FREQUENCY OF EVALUATION QUESTIONNAIRE	114
F	CONVERSION TABLE BASED ON NORMS FOR UNIVERSITY OF MICHIGAN FRESHMAN (1955)	117
G	CODING CATEGORIES FOR OPEN-ENDED RESPONSES ON THE SUBJECTIVE REPORT QUESTIONNAIRE	122
H	DISTRIBUTION CHARACTERISTICS OF VARIABLES FOR THE SURVEY AND EXPERIMENTAL STUDIES	125
I	RELIABILITY COEFFICIENTS FOR THE STUDENT MEASURES	128
J	SCALE SCORES FOR LEVELS OF THE TREATMENTS BY LEVELS DESIGN	129
K	THE COMPLETE ANALYSIS OF VARIANCE FOR THE SURVEY STUDY	133
L	THE COMPLETE ANALYSIS OF VARIANCE FOR THE EXPERIMENTAL STUDY	148
M	ANALYSIS OF CRITERIA TEST DATA WITH ITEM DIFFICULTY CONTROLLED	150
N	RAW DATA	161

INTRODUCTION

Educational psychologists have devoted a great deal of effort to identifying the factors which are associated with academic performance. The advent of mental testing through the work of Binet, Terman, Otis, and others, seemed to provide a real breakthrough in terms of defining a quantifiable dimension related to academic performance. But even as mental tests were refined, it became increasingly clear that other factors must be operating also. Validity coefficients between mental test scores and academic achievement continued to range between .30 and .80 with the median coefficient being about .50 (Tyler, 1956). Needless to say, the magnitude of these coefficients leaves a good deal of the variance unaccounted for.

Methods of grouping and instruction have also been studied extensively. The results of research in these areas have also failed to account for additional variance in any convincing way. (See Wallen and Travers, 1963; McKeachie, 1963.)

In order to account for the lack of conclusive findings in studies on teaching methods, McKeachie (1961) argues that differences in individual achievement are being masked by the use of group comparisons. While certain students might benefit from a particular method, others could be detrimentally affected. Therefore, when group comparisons are made, the gains may be offset by "unexpected" losses so that there

appears to be little, if any, overall gain in achievement attributable to method.

In order to test his assumptions, McKeachie (1961) examined relationships between student personality characteristics and differing teaching styles. He was able to find significant interactions between variables such as students' need Affiliation and teacher warmth in determining grades in college courses.

In view of these results, it would be advantageous to identify other individual and situational characteristics in the classroom which are subject to similar interactions. Since the effect of anxiety on learning and performance has been demonstrated in the laboratory (Child, 1954; Atkinson, 1960; I. G. Sarason, 1960), this personality variable should be a relevant factor in academic performance. Therefore, the present study examines relationships between student anxiety, teacher behavior, and academic achievement.

CHAPTER I

A REVIEW OF RELEVANT LITERATURE

The Measurement of Anxiety

The theoretical construct of anxiety could hardly be considered new. For example, Freud's (1936) detailed discussion of anxiety was translated and published in America as early as 1936. Although his influence can be seen in current theories of anxiety (e.g., Dollard and Miller, 1950; Mandler and Sarason, 1952), experimental study of this concept with human subjects did not flourish until the appearance of the Manifest Anxiety Scale (MAS) and the Test Anxiety Questionnaire (TAQ) in the early 1950's.

Taylor (1951, 1953) constructed the Manifest Anxiety Scale for use in a study investigating the relationship between eyelid conditioning and anxiety as a measure of Hull's D (drive). Items were selected from the Minnesota Multiphasic Personality Inventory and submitted to a panel of clinicians who made judgments as to whether or not the items satisfied an a priori definition of manifest anxiety. Items on which there was 80 per cent agreement or higher were included in the final scale.

In contrast to the Manifest Anxiety Scale, which is considered to be a measure of chronic anxiety level, Mandler and Sarason (1952) developed a questionnaire which was concerned with attitudes and

feelings in testing situations. The items attempt to tap manifest anxiety as inferred from physiological and emotional reactions such as feelings of uneasiness, accelerated heartbeat, perspiration, and worry before and during intelligence tests and course examinations. Thus, the TAQ is a measure of anxiety specific to testing situations.

Dissatisfied with existing scales for assessing anxiety, Haber and Alpert (1960) attempted to derive an even purer measure of specific anxiety, the Achievement Anxiety Test (AAT). An innovation in their instrument was the inclusion of two sub-scales, one to measure the facilitating effects of anxiety (AAT+), the other to measure the debilitating effects of anxiety (AAT-) on achievement test performance. Final examination scores and course grades were used as external criteria for item selection. Items which were highly correlated with the criteria but uncorrelated with each other were included in the final instrument. Item content varies somewhat from the MAS and TAQ in that physiological symptoms of manifest anxiety are not included. Instead, questions ask the test-taker to assess how much "nervousness," "fear of a bad grade," "blocking on questions," "time pressure," etc., affect his academic test performance.

Inspection of the item content of these scales suggests face validity for the construct of anxiety but the questions of construct validity and predictive validity remain to be answered. The extent to which these scales are measuring the same concept will be taken up shortly in the section entitled "Correlational Studies." Construct validity for these measures of anxiety will have to be gleaned from the accumulated experimental and correlational evidence which follows.

Experimental Studies of Anxiety

A review of the literature reveals that experimental studies of anxiety incorporate varying methodologies and dependent variables. For example, one frequently used design classifies subjects on some measure of anxiety and then examines relationships between level of anxiety and learning on tasks of varying difficulty. It was initially used by Taylor (1951) and Taylor and Spence (1952) in order to test basic assumptions underlying Hullian learning theory.

In the Hullian model, it is assumed that each organism has a hierarchy of responses of varying habit strength. Emission of responses is a function of the drive level of the organism, with manifest anxiety being considered as one source of drive. An increase in drive level would bring already existing responses above emission threshold. Therefore, when habit strength of a goal response is initially low, responses of higher habit strength (though they may be irrelevant) are thought to compete for emission. High levels of anxiety drive would hinder acquisition of a new response because of competing response tendencies already in the habit hierarchy. High anxiety drive, therefore, would have a facilitating effect on performance of already acquired responses but a debilitating effect on learning or acquisition of new, complex responses. Since anxiety drive is assumed to be a stable characteristic of an individual, no attempt is made to arouse the drive. Instead, drive level is manipulated by selection of subjects rather than by experimental arousal procedures.

Typical of this experimental design is a study by Taylor and Spence (1952) who tested the Hullian drive model on a verbal maze

learning task. MAS scores were used to classify subjects on level of anxiety. A T-maze was mounted on a memory drum to facilitate standardized administration procedures. Subjects negotiated the maze by responding with the words "right" or "left" at each choice point. As predicted, high-anxious subjects made significantly more errors and required significantly more trials to reach the criterion level of performance. It was also found that the largest difference in errors between high- and low-anxious subjects took place at the most difficult choice points in the maze.

In an experiment designed to bring out differences in emission of responses of varying habit strength, Farber and Spence (1953) tested subjects on two tasks. Negotiating a stylus maze was considered to be a complex response initially of low habit strength while blinking of the eyelid was an initially dominant response tendency. Again subjects were selected on the basis of extreme MAS scores. Learning the stylus maze was complicated by the fact that subjects were shielded from seeing it. Correct responses were verbally reinforced by the experimenter. The criterion of learning was two successive errorless trials. At a later date, these subjects were also subjected to an eyelid conditioning experiment. The maze learning of the high-anxious subjects was significantly poorer than that of the low-anxious subjects, with the more difficult choice points providing the largest difference in errors between the two groups. Anxious subjects proved to be superior in eyelid conditioning.

Other tests of the Hullian model were carried out with essentially the same results. Montague (1953) manipulated the degree of response competition by varying association value and intralist similarity of

nonsense syllables. Ramond (1953) used words of differing association value to induce response competition. As predicted, low-anxious subjects responded correctly significantly more often when the weaker associations were correct. On lists with a minimum number of competing associations, high-anxious subjects tended to excel. Using water jar problems and an anagrams task, Maltzman, Fox, and Morrisett (1953) found that low-anxious subjects tended to be more flexible in deriving water jar solutions, whereas high-anxious subjects tended to persist in the method of solution induced during early trials. However, on the anagrams task, high-anxious subjects made fewer errors when the dominant response tendency was the correct one.

The competing responses hypothesis derived from the Hullian drive model is well supported. Studies utilizing a wide variety of responses as the dependent variable give highly consistent results. Tentative generalizations from the findings of these studies would include (a) high anxiety drive facilitates performance of dominant response tendencies, (b) high anxiety drive impedes the acquisition of complex, initially nondominant response tendencies, and (c) anxiety drive is a relatively stable characteristic and hence will not vary a great deal across situations or tasks. However these generalizations have been derived from only one methodological approach. An unanswered question remains. How generalizable are these results to other situations and tasks? Results of other approaches are therefore necessary in order to assess the validity of these findings.

A second common methodological approach also involves relationships between anxiety, learning, and motivational conditions. Control

on level of anxiety is not always used, however, based on the assumption that anxiety can be aroused by experimental manipulation, contrary to assumptions held by the Hullian group. Anxiety producing stress is induced through either symbolic or physical threat, e.g., threatened or actual electric shock, verbal communications which evaluate the subject's performance as being inadequate, etc. An important aspect of this design is the neutral or control condition which is intended to minimize anxiety arousal.

Principal exponents of this arousal or "reactivity" model are Mandler and Sarason (1952). In the "reactivity" model they assume that anxiety is a drive learned as a response to stimuli present in specific situations. Drive is assumed to require reduction. Therefore, anxiety drive would elicit responses that tend to reduce that drive. These responses are of two general types: (a) responses to anxiety which are not specifically connected with learning or performance of a task and (b) responses which reduce anxiety through completion of the required task. Included in the first category would be self-oriented responses aroused by the symbolic threat of a testing situation. Examples of self-oriented responses would include experiencing feelings of inadequacy or helplessness to cope with the situation, anticipation of punishment for poor performance, and implicit attempts to leave the psychological field through fantasy, etc. These reactions are believed to be generalized from previous experiences in similar situations. Anxiety responses in the second category are more realistic attempts to deal directly with the source of threat in that anxiety reduction takes place through acts instrumental toward completion of the task. Thus, in an

aroused state, persons with a predisposition to emit self-oriented responses would experience much internal conflict. The self-oriented responses, whose aim is to reduce anxiety drive, would compete with responses necessary to learn or to complete the required task. When response competition is strong, the level of learning and performance is reduced. On the other hand, persons who react more realistically to the situation would find their performance facilitated by an increase in level of arousal and absence of interfering self-oriented response tendencies.

Mandler and Sarason (1952) carried out an experiment to test their assumptions on the role of internal drive states and resultant responses on learning. The variables of interest were level of test anxiety, success versus failure in task completion, and learning to perform a task. Subjects were chosen from the extremes of a distribution of TAQ scores. To enhance initial ego-involvement, subjects were told that they were being given an intelligence test the results of which would be placed in their college records. There were practice trials in which the subject was allowed to work on a Kohs Block Design to completion, and the Digit Symbol Test which could not be completed in the time allowed. Thereupon, differential instructions were communicated to the subjects telling them that they had done well (success condition), done poorly (failure condition), or asked to continue (neutral condition). Then more trials with a new Kohs Block Design and a modified form of the Digit Symbol were administered.

The rationale underlying the design of this study was that failure or lack of successful completion would induce self-oriented responses in subjects who are high in test anxiety. Therefore, learning and

consequently performance should be lower for them under the failure stress condition. Results of the study supported their assumptions in large part as low-anxious subjects performed better on the first five trials of the Block Design although as learning proceeded, the high-anxiety group tended to improve rapidly. Surprisingly, both success and failure reports improved performance for the low-anxious subjects but depressed scores for the highs. The prediction that lack of task completion would hinder the high-anxious subjects was not supported as differences in Digit Symbol performance were not significant.

In another test of the "reactivity" model, Sarason, Mandler, and Craighill (1952) again used differential motivating instructions. Subjects in Group I were told that they should be able to finish the task in the allotted time. In Group II they were told that they were not expected to finish. The assumption being tested was that noncompletion would evoke irrational responses in high-anxious subjects which would interfere with their performance. There were no significant differences in performance of high-anxious subjects under either condition. Low-anxious subjects in the "expected to finish" condition did significantly better than the low-anxious "not expected to finish" group or either of the high-anxious groups. In a second study, these same experimenters found that ego-involving instructions tended to reduce the performance level of high-anxious subjects on a stylus maze task. They concluded that stress-producing instructions can have opposite effects on learning as a function of level of anxiety specific to testing situations.

Another type of research design used to investigate the effects of stress on performance simply assigns subjects to a high- and

low-stress condition on a random basis without controlling for individual differences in anxiety level. The assumption is made that anxiety is aroused by the stressful stimuli. This research design was used by Lazarus and Eriksen (1952). All subjects were given ego-involving instructions initially. After a period of practice, subjects were divided into an experimental and a control group. The control group was told that it had done well and was asked to perform a Digit Symbol task again. The experimental group was informed that it had done poorly but that it would be given an opportunity to try to improve. Results revealed little difference between the two groups. Further analysis revealed that subjects with high grade point averages showed a tendency to improve under stress while the reverse was true for subjects with low academic standing. Although there was no control on level of test anxiety, these findings seem to be parallel to those reported by Sarason et al. (1952) in that individuals vary with regard to their reactions to stress.

A puzzling finding is reported by Deese, Lazarus, and Keenan (1953). These experimenters assigned subjects scoring at the extremes of a distribution of scores derived from Winne's Neuroticism scale, to one of three conditions: control, avoidance-learning, and nonavoidance-learning. The task was to learn a list of nonsense syllables. Under the avoidance-learning condition subjects were shocked for wrong responses as well as hesitant ones. Subjects were shocked at random whether they were right or wrong under the nonavoidance condition. The high-anxious subjects were superior under all three conditions although the

difference was significant only for the avoidance-learning group. However, these differences seemed to be due to learning decrements of the low-anxious group as their performance curve was negatively accelerated whereas the curve for the high-anxious group maintained a constant slope throughout the trials. Therefore it was concluded that stress did not increase the performance of the high-anxious subjects, while it decreased the performance level of low-anxious subjects. Both results were in contradiction to findings anticipated by the Iowa studies. In view of these results, Deese et al. felt that more needed to be known about task variables as well as subject populations.

Before proceeding further, the implicit circularity involved in the terms learning and performance should be considered. Learning is defined as the acquisition of new responses. Performance refers to the emission of responses. The differentiation between the two terms becomes blurred because increments in learning must be inferred from increases in observed performance. This distinction is an important one especially when the strength of a response is being considered as in the Hullian framework. It should be noted that the previous studies emphasized the effect of anxiety drive, and its arousal, upon response acquisition or learning. Experiments investigating the effect of anxiety on already established habits (performance per se) are few in number.

In order to study the effect of anxiety on performance, Westrope (1953) selected subjects from the extremes of an MAS distribution and allowed them twelve practice and three control trials before placing them under stress. The stressful situation was created in an observation room by means of a mechanical apparatus. An electrode was fastened

to the subject's hand, also. Subsequently, the subject was told that he might receive a shock if his performance were below standard. A red light would flash to warn him when he was performing inadequately. However, the red light was turned on for most of the performance interval and each subject received a shock just before the end of each of three experimental trials. The results showed only a main effect for conditions. Stress significantly impaired the Digit Symbol performance of subjects irrespective of their level of anxiety or sex although there was a slight tendency for low-anxious men's and women's performance to be less detrimentally affected. This unexpected outcome might be explained at least in part by the nature of the experimental condition. The fact that performance was detrimentally affected for both high- and low-anxious subjects suggests that the threat inducing stimuli were too highly distracting and stressful for both the high- and low-anxious subjects during the brief period of time allowed for them to work.

Irwin Sarason (1957) also studied the effect of anxiety arousal on verbal learning and performance. His subjects were designated as high-, medium-, and low-anxious on the basis of MAS scores. There were three conditions: (a) neutral, (b) ego-involving or subject oriented, (c) experimenter-oriented. The motivating conditions were used only on Day I. An analysis of variance for the learning phase revealed significant main effects for level of anxiety, motivating instructions, and an instructions x anxiety interaction. High-anxious subjects excelled all others under the neutral condition while low-anxious subjects were superior under the subject-oriented condition. The middle-anxiety group was poorest under all conditions. Twenty-four hours later all subjects

were re-tested on the same list of nonsense syllables under neutral conditions. Only anxiety had a significant main effect, with the high- and low-anxious subjects again excelling the middle group. Although the difference was not significant, high-anxious subjects were superior to the lows. Except for the deviant performance of mid-anxiety subjects, the results from Day II are in line with the notion that level of anxiety has much less effect upon acquired (i.e., dominant) response tendencies.

A third methodological approach studies anxiety level, task difficulty, and stress simultaneously. This design was used by Farber and Spence (1956) in an attempt to determine whether anxiety drive is "chronic" or "reactive." They used a discrimination task of varying difficulty with reaction time as the dependent variable. Stress was induced by threat of an electric shock for inadequate performance. All subjects were given practice trials before beginning the stress condition. Although differences were found between male and female performance and complexity of the task, no significant main effects or interactions were found for level of MAS or for experimentally induced stress. Dissatisfied with their initial effort, Farber and Spence carried out a second experiment in which shock was actually used in the stress condition. Subjects scoring at five different levels of MAS were also included in the design. Results from this second experiment also failed to substantiate predictions one would make from the earlier Iowa studies. Differences in learning and performance could be accounted for by the variables of sex, task complexity, and practice, with the differences between sexes becoming relatively smaller at the end. There was no main

effect or interaction for anxiety. There was a neutral condition in the first experiment but not in the second. A variety of explanations can be offered for their unexpected results. The fact that subjects were given practice trials prior to stress may have been sufficient to develop response tendencies strong enough to withstand the stress used. As such, their data would be more representative of performance phenomena than response acquisition phenomena, a distinct difference from the earlier Iowa studies.

A series of experiments by Sarason and Palola (1960) was designed to examine relationships between these same variables. Moreover, subjects were categorized on both the TAQ and the MAS. Under the neutral condition subjects were informed that they were performing a series of experimental tests. Subjects in the stress condition were given ego-involving instructions with emphasis on the fact that their performance would be compared with other students. A triple interaction of TAQ, instructions, and task difficulty was found as high-TAQ subjects were superior on an easy form of the Digit Symbol Test under both stress and neutral conditions. On a more difficult form of the Digit Symbol Test, low-TAQ's did better under stress, but high-TAQ's were superior under the neutral condition.

By changing the dependent variable to learning and performance on an arithmetic task of varying difficulty, it was found that performance for both high- and low-TAQ groups was almost identical on the easy form. On the more difficult problems, low-TAQ subjects were superior to high-TAQ subjects under both conditions. Analysis of performance

scores on these same subjects with control on MAS revealed no main effect for that variable nor any interactions with other variables. It was concluded that the specific anxiety scale is a better predictor of performance in a testing situation than a more general measure of anxiety such as the MAS.

This outcome can be interpreted as support for the assumptions made by Mandler and Sarason that anxiety responses are learned in connection with specific situations. Their measure of anxiety specific to testing situations helped to differentiate groups on the dependent variable with greater certainty than the more general measure of anxiety. Furthermore, the interaction of Test Anxiety, motivating conditions, and task difficulty is in line with predictions one would make from the "reactivity" model. These results also suggest that the initial habit strength of the high-anxious subjects must have been sufficient for the easier tasks as stress did not appear to hamper their performance.

The differential effects of anxiety arousal on learning and performance was brought out in an experiment by Moldowsky and Moldowsky (1952). Their aim was to study the effect of motivating instructions on tasks varying in degree of previous learning. Although they did not control on level of anxiety, subjects were carefully matched on Digit Span and Verbal scores derived from an earlier testing on the Wechsler-Bellevue Scale. Stress instructions implied that the subject had done poorly on the initial administration of the Wechsler scale and was encouraged to improve his score. Control subjects were tested under regular clinical testing conditions. The control group showed an

increase in Digit Span performance over their initial testing while the scores for the experimental group dropped. Both groups improved in vocabulary scores, however. It was concluded that anxiety arousal impedes learning of a relatively complex nature, e.g., immediate memory on the Digit Span Test, but does not affect already existing or over-learned responses. In support of their results, these experimenters cite a study by Montague and Schneider (1949) who found that once high-anxiety subjects had reached a certain criterion of learning, their retention of nonsense syllables was as good as that of low-anxious subjects. Whether stressful conditions were used in their study is not clear.

Correlational Studies

The experimental studies have raised two issues which deserve further consideration. They are: (a) the nature of the relationship between different measures of anxiety, and (b) the nature of the relationship between anxiety and measures of intellectual ability. The latter point is especially relevant because the dependent variable in most of the previous studies has been learning or response acquisition. Since a measure of intellectual ability is an index of one's ability to learn, can the results observed be accounted for in terms of differences in aptitude for learning? Moreover, many of the task variables are sub-scales used in assessing intellectual aptitude, e.g., Digit Symbol Test, Kohs Block Designs, etc. Following are several studies which should be helpful in clarifying these issues.

The results of studies using more than one measure of anxiety should help to clarify the nature of the relationship between different instruments purported to be measures of that variable. For example, the experimental study by Sarason and Palola (1960) used both MAS and TAQ scores as independent variables. No main effect or interaction was found for MAS while several other significant effects were observed when Test Anxiety was used. This outcome suggests that they are indeed tapping different dimensions.

The results of studies in which three measures of anxiety were intercorrelated may be noted in Table 1.

TABLE 1

INTERCORRELATIONS OF THREE MEASURES OF ANXIETY

Experimenter(s)		MAS[a]·TAQ[b]	MAS·AAT-[c]	TAQ·AAT-
Sinick (1956)	Men	.45		
	Women	.35		
I. Sarason (1959)	Men	.41		
	Women	.49		
Haber and Alpert (1960)	Men	.32	.38	.64

(All coefficients are significant, p = .05 or less.)

[a]Manifest Anxiety Scale.

[b]Test Anxiety Questionnaire.

[c]Debilitating Anxiety sub-scale of the Achievement Anxiety Test.

As can be seen from the table, the relationships are all positive but their magnitude suggests that the amount of common variance between the specific scales and the general scale is small. Although the

correlation of AAT- and TAQ is larger (.64), suggesting that they are more nearly measuring the same variable, the common variance is still less than 50 per cent.

In another study, Gordon and Sarason (1955) report results concerning the TAQ and a scale to measure generalized anxiety (GA). Item content in the GA scale dealt with manifestations of anxiety in a variety of situations. A correlation of .468 (N=389) was found between GA and TAQ scores. Although it was apparent that anxiety in a testing situation was related to generalized anxiety, they also examined extremes of the TAQ distribution for intraindividual differences with respect to GA. It was found that both high- and relatively high-scorers on the TAQ were a great deal alike in their endorsement of items on the GA scale. The low- and relatively low-scorers on the TAQ were found to differ in their endorsement of items varying in phobic content. On the basis of this result it was suggested that low scorers on the TAQ attempt to deal with their anxiety in testing situations in a different manner than do subjects who score high on the TAQ.

I. G. Sarason (1957) intercorrelated two measures of ability with scores on the TAQ and the GA scale (Gordon and Sarason, 1955). TAQ scores correlated negatively with scores on the verbal part of the Scholastic Aptitude Test (SATV) as well as the mathematics part (SATM). GA was independent of these ability measures but did correlate <u>positively</u> with grade point average over the course of four years in college. On the other hand, TAQ scores tended to correlate negatively with college grade point average during the first two years and then reduce to zero

over the last two. Sarason concluded that relationships between anxiety and achievement variables depend to a large extent on the nature of the instrument employed to measure anxiety.

I. G. Sarason (1961) found that Test Anxiety is a less detrimental factor for course grades than for aptitude tests. For both males and females the relationship of TAQ to performance was negative and significant for 11 out of the 13 achievement and aptitude variables studied. There were indications that the negative correlations were stronger between TAQ and aptitude test scores than for grade point average. The observed differences in degree of relationship were attributed to the nature of the testing situation. Since the content on aptitude tests is usually unknown, there is no opportunity to prepare oneself ahead of time in order to compensate for anxiety by overlearning. In academic performance, content requirements are usually known and can be overlearned, hence the lower degree of relationship between TAQ and grade point average. Using high school students for subjects, I. G. Sarason (1963) administered the TAQ and the School and College Ability Test (SCAT). Test Anxiety was found to be negatively correlated with SCAT scores. The negative correlations obtained tended to be larger for female than for male subjects. Two possible alternatives were given for the recurring relationship between anxiety and intellectual ability. They were (a) high TAQ subjects are realistically more anxious about taking tests because of an actually lower intellectual level; (b) subjects differing in TAQ scores do not differ in intellectual level but do differ in the extent to which anxiety interferes with their performance on ability tests.

The MAS also has been studied in relation to measures of ability and academic achievement. For example, Klugh and Bendig (1955) reported no relationship between the Manifest Anxiety Scale and scores on the American Council on Education Psychological Test (ACE). MAS scores were independent of grade point average, contrary to some earlier findings. Spielberger (1958) made a rather thorough analysis of these variables, also. By analyzing the distribution characteristics of both measures in several samples of male subjects, he found that significant relationships seemed to arise only when subjects with extremely low ACE scores were included in the samples. He concluded that the inverse correlation between MAS and ACE scores would be obscured by selection procedures which restrict the range of intelligence.

Spielberger and Katzenmeyer (1959) reassessed the problem by correlating MAS and college grade point average with control on level of ability. The ACE distribution was divided into an upper 20 per cent, lower 20 per cent, and a middle 60 per cent. Tests for linear and curvilinear regression indicated that grades were unrelated to MAS scores for the high and low ACE groups. Grades varied inversely with anxiety level for the average ACE subjects, however. On the basis of these results they concluded that college work is too difficult for low ACE subjects whose low grades were unrelated to MAS. High ACE subjects tended to obtain high grades regardless of their anxiety level. Anxiety seemed to have its largest effect for students of middle ability.

A comprehensive study of the relationship between measures of anxiety and academic achievement was carried out by Haber and Alpert (1960). Scores on the MAS, TAQ, their own AAT+ and AAT- sub-scales,

and grade point average were intercorrelated on several samples. It was found that the specific scales (TAQ, AAT+, and AAT-) correlated much higher with each other and with college grades than with the MAS. It was also found that the specific anxiety measures were all related to verbal aptitude while the general measure was not. On the basis of their correlational analysis, Haber and Alpert concluded that specific anxiety scales measure something quite different from whatever is measured by general measures of anxiety.

In spite of the relationship between AAT scores and ability, Alpert and Haber also found that their anxiety subscales made a significant addition to ability measures in the prediction of grades by a multiple regression equation. A somewhat contradictory finding is reported by Lin (1964). Using multiple regression techniques on several large samples of male and female subjects, he found that ACE or SAT measures occurred in each of four regression equations used in predicting grades in an introductory psychology course, while AAT+ or AAT- failed to appear in any of the equations. It should be pointed out, however, that Lin was predicting grades for only one course while Haber and Alpert were predicting grade averages for several courses and hence would have a much more reliable measure of academic performance.

A broader interpretation was given to level of Test Anxiety by Atkinson and Litwin (1960). They used the TAQ as an index of a motive to avoid failure. Using only male subjects, they found that high TAQ males spent significantly less time working on a final examination and earned significantly lower grades as well as showing a low preference for intermediate risks in a ring toss game. The strength of the

relationship was enhanced by classifying subjects simultaneously on need for Achievement and level of TAQ. Atkinson felt that his interpretation of Test Anxiety as a measure of failure avoidance was not incompatible with the interfering response theory on which the instrument is based.

It seems evident from these results that not all anxiety scales are measuring the same dimension. The specific anxiety scales are better predictors of academic achievement than the MAS although they correlate more highly with measures of ability which are potent predictors of academic achievement also. At first glance, the inverse relationship between measures of achievement and test anxiety seems puzzling. What is the basis for their sharing of common variance? A result obtained by Grice (1955) is disconcerting in this respect. Using high- and low-MAS subjects, he found that the low scoring subjects were superior in performance of a complex discrimination-reaction-time task. When these results were controlled for level of ability, differences between the two anxiety groups disappeared. Spielberger and Katzenmeyer's (1959) study controlled on intelligence and found that level of anxiety was associated with performance at the middle levels of ability only. As indicated previously, Haber and Alpert (1960) found that their measures of anxiety were able to account for additional variance above and beyond that accounted for by ability even though their measures were correlated with ability, also. It could well be that this inverse relationship reflects a component of personality which cannot be separated from intellectual performance. As suggested by I. G. Sarason (1963), the recurring relationship between anxiety and intellectual ability

could be due to the fact that persons differing in Test Anxiety do not differ in intellectual level but do differ in the extent to which anxiety interferes with their performance on ability tests.

In light of this accumulated evidence, it seems clear that specific anxiety scales are indeed measuring a different behavior-related dimension than the more general assessment devices. The relationship between specific anxiety measures and intellectual aptitude is a more difficult one to unravel. It seems likely that this relationship reflects a nonintellectual component of intellectual performance. By studying these two dimensions simultaneously, more recent studies have found that specific measures of anxiety account for additional variance, even with ability controlled, in the prediction of academic achievement. Because of the ambiguity of this relationship, it would seem advisable to take both variables into account when designing studies in which the dependent variable contains a large component of learning ability. Precise designs of this nature should help to clarify this issue more satisfactorily.

The large body of research reported here lends a great deal of support to the notion that highly anxious persons suffer decrements in acquiring new responses. But level of motivation is a key factor in performance. A question might be raised as to the validity of the studies on the grounds that ego-involvement of subjects in experiments is marginal. Results of studies carried out under "real-life" stress could add appreciably to the construct validity of findings generated in the laboratory. Following are studies attempted in such "real-life" settings.

The Experimental Approach in Real Life Situations

Carrier (1956) studied the effect of stress on performance of course examinations in an introductory psychology course. Under the stress condition, subjects were assigned to a strange room and a strange proctor for the exam. Instructions also included the statement that there would be a penalty for guessing but that no question was to be left unanswered. There were two very difficult trial questions which were intended to induce failure. Under relaxed conditions, the examination was administered by the regular instructor who tried to maintain a friendly, relaxed atmosphere. There were no difficult practice questions or guessing penalty. In addition, a special answer sheet with space for writing optional comments was used. This format had been shown to be beneficial in reducing effects of anxiety in previous studies (McKeachie, Pollie and Speisman, 1955). Statistical controls on level of ability and past academic performance were used by means of a multiple regression equation with the final analysis being done on residuals (predicted achievement minus actual achievement). Instead of the standard measures of test anxiety, the Social Approachability-Stability personality questionnaire was used (Smith, Wood, Downer, and Raygor, 1956). The SA-dimension is purported to tap susceptibility to influence by internal and external stimuli, the S-dimension a stability-anxiety dimension reflecting nervous tension and desurgency.

It was found that women were more detrimentally affected by stress than men. Subjects scoring high on the SA (permeability) dimension were detrimentally affected by stress also. Although it had been

predicted that low S (stability-anxiety) subjects would experience performance decrements under stress, this hypothesis was not statistically confirmed although observed effects were in the expected direction. If a low S (stability-anxiety) score can be assumed to be indicative of anxiety, then the effect of stress on the performance of high anxious subjects is small. It should be pointed out, however, that the SA-S measures general rather than specific anxiety.

French (1961) studied the effect of emotional states on the performance of college entrance examinations. Each subject was used as his own control. The stress condition was the regular administration of the College Board Scholastic Aptitude Test (SAT). The relaxed condition took place either a few days before or after the regular administration of the SAT so that order and practice effects could be counterbalanced. Performance measures under the two conditions were parallel form subtests taken from an outdated SAT battery.

Subjects were told they were taking experimental tests for research purposes although the scores would be reported to the student's school. The measure of anxiety was a forced-choice version of the MAS. Results showed the effects of anxiety to be small as there was no significant anxiety effect for boys, while anxiety seemed to enhance performance on the mathematics test for girls. Moreover, the concurrent validity of the tests under relaxed conditions in predicting grades was not significantly different from that obtained using test results under stress conditions. It is unfortunate that a measure of general anxiety was used since several of the studies cited previously have shown that aptitude test performance is more closely related to test anxiety than

to general anxiety. This fact raises some question as to the conclusiveness of French's findings. On the other hand, since subjects did not know what to expect on the tests, they could not study for them. Thus, since previous learning is being tapped rather than rate of new response acquisition, anxiety arousal would tend to have a smaller effect on test performance as found by Moldowsky and Moldowsky (1952).

Using several different paper and pencil tests, Sperber (1961) studied the effect of anxiety level and defense preference on performance under stress. His subjects were recruits undergoing aptitude testing for placement in the Air Force. Since each subject's future in the Air Force depended upon the outcome of this assessment, it was assumed that the subjects would be highly engaged in task performance as well as experiencing a realistic threat of evaluation under the stress condition. In addition, the first task under stress was insoluble. The remaining tests were long enough so that they could not be completed in the time allowed even though subjects were told they should be able to finish them. In the neutral condition, subjects were told they were taking experimental tests and were allowed an initial success experience. The TAQ was used as the measure of anxiety with subjects being selected from the extremes of the distribution. Control on level of intelligence and years of education was also used in the analysis. Sperber's findings were surprising in that high-TAQ recruits under stress outperformed similar high-TAQ subjects in the neutral condition on two of three number matching tests. Furthermore, high-TAQ subjects outperformed low-TAQ subjects under the stress condition, also. Under low stress the low-TAQ

subjects seemed to do better than the highs on letter substitution and number matching. Sperber explained his findings on the basis of complex interactions between variables such as anxiety drive, motives, and characteristics of the task. Since many of the tasks were of a simple yet boring and repetitive nature, it was reasoned that high anxiety drive assisted high-TAQ subjects to maintain a higher level of effort in view of the payoff for good performance under the "real" condition. Since payoff was much less evident in the "experimental-low stress" situation, high-TAQ subjects apparently failed to channel their anxiety drive into effort since it was safe to be unconcerned about personal performance.

Eriksen and Paul (1964) investigated the effect of test anxiety on course examinations in an introductory psychology course. Following a regular examination female students were seen in small informal groups. At this time they were given an experimental examination over the same content that had been covered by the regular test that morning. Moreover, the subjects were told that scores on this test would not count toward their final grade. Intercorrelation of variables showed that scores on the TAQ were negatively related to aptitude scores on the SCAT and the regular exam grade. TAQ scores were not related to scores on the experimental exam. There were no significant differences between levels of anxiety and conditions for the total sample. However, by excluding extreme levels of ability as Spielberger (1958) had done, a significant interaction was obtained between TAQ and conditions with low-TAQ middle ability subjects doing better under stress of evaluation, and high-TAQ middle ability subjects doing better under relaxed conditions. Examination of test scores on two later examinations in the

course did not reveal any significant differences between the high- or low-anxious middle ability groups. Level of anxiety apparently did not enhance nor hinder performance under "real life" stress with level of ability held constant.

Stakenas (1964) investigated interactions between level of debilitating anxiety (AAT-) and teacher behavior. It had been hypothesized that the learning and performance of highly anxious students would be hampered by teachers who emit anxiety cues, e.g., emphasizing grades. Contrary to expectation, students high in debilitating anxiety tended to earn higher grades in an introductory psychology course in classrooms rated as high in anxiety cues.

A word of caution should be invoked before concluding that stress and anxiety do not interact in "real life." There is a major methodological difference between the laboratory and "real life" studies. In general, the laboratory studies required the subject to work on an unfamiliar task so that the dependent variable is heavily loaded with a learning or response acquisition component. The "real life" studies, on the other hand, present the subjects with tasks with which they are already familiar, e.g., examinations based on course content previously assigned. Hence, the dependent variable is heavily loaded on performance or response emission. Therefore the seemingly contradictory outcomes may stem more from differences in methodology than from inconsistencies in theory.

Summary

A review of research literature on anxiety reveals that a variety of methodologies and dependent variables have been used to study the effect of this variable. A multimethod appraisal of the effect of anxiety on learning and performance suggests the following conclusions.

1. Not all scales purporting to measure anxiety are measuring the same thing, as intercorrelations across scales are relatively low and effectiveness of prediction depends upon the task and the scale used. Scales measuring anxiety specific to the dependent variable are better predictors than more general measures.
2. The inverse relationship between specific anxiety scales and measures of intellectual ability is not well understood. Few anxiety studies have been done using control on ability also. Studies which have used this technique seem to indicate that measures of specific anxiety can account for additional variance beyond that accounted for by ability alone. Test anxiety, therefore, appears to be a noncognitive component of intellectual performance.

The terms learning and performance are often used as synonyms and as such can lead to invalid conclusions. Although learning must necessarily be inferred from increments in performance, it is important to distinguish between these concepts. Learning is most useful in describing the phenomenon of response <u>acquisition</u>. Performance denotes rate of response <u>emission</u>. With this distinction in mind, it is concluded that:

3. Anxiety arousal is most detrimental to acquisition of unfamiliar, complex responses.
4. Anxiety arousal is least detrimental to performance of responses already acquired and may even facilitate emission of highly developed habits.
5. Existing studies tend to examine the "chronic" and "reactive" properties of anxiety. An area still largely unexplored is anxiety as an indicator of motive.

CHAPTER II

THEORY AND HYPOTHESES

The Problem

The formulation of this problem is an outgrowth of the work of McKeachie (1961) who has proposed and demonstrated that academic performance in college is influenced by interactions between student characteristics and situational cues stemming from teacher behavior. In his social-psychological model he assumes that the behavior of individuals is affected by the arousal of motives which have been acquired from past experience. Therefore, if motives instrumental to attaining course goals can be aroused in students, then their academic achievement should be enhanced. Since the teacher determines the goals of the course, evaluates student performance, accepts or rejects student friendship, etc., each of these behaviors can serve as a cue for motive arousal contingent upon the presence of relevant motives in the student.

Examples of arousable motives well known in the psychological literature are need Achievement (McClelland, 1953), need Affiliation (Atkinson, 1958), and need Power (Atkinson, 1958). The classroom motive arousal model was tested by McKeachie using the above mentioned variables. The most promising results were obtained using the measure of affiliation motivation, as students' need Affiliation was found to interact with teacher warmth in determining course grades. Students high in

need Affiliation tended to receive higher grades in classes taught by teachers rated warm and friendly than in classes rated low on this dimension (McKeachie, 1961).

In view of the evidence presented in Chapter I, the construct of anxiety, and its arousal, appears to be a variable worthy of consideration within the context of the social-psychological classroom model proposed by McKeachie. Therefore, the problem under investigation in this study is to determine the relationship between student anxiety level, teacher produced stress cues, and academic achievement.

The Concept of Anxiety

There is a distinct similarity between Freud's discussion of anxiety and the various interpretations proposed by experimental psychologists such as Taylor and Spence, or Mandler and Sarason. The most distinguishable points of overlap deal with physiological manifestations of anxiety in the case of the former experimentalists, and the task-relevant and task-irrelevant responses assumed to exist by the latter pair. This should become more apparent in light of the following discussion.

For Freud (1936), anxiety is an unpleasant affective state, but, he hastens to point out, not all unpleasant affects are anxiety, e.g., sadness or sorrow. In his words:

> Anxiety, therefore, is a specific state of unpleasure accompanied by motor discharge along definite pathways. . . . An increase of excitation underlies anxiety, an increase which on one hand is responsible for its unpleasant character and on the other is relieved through the discharge referred to [Freud, 1936, p. 70].

He goes on to distinguish between <u>true</u> and <u>neurotic</u> anxiety. True anxiety is aroused by the presence of a real danger which is readily

apparent. Neurotic anxiety is fear of what Freud termed instinctual danger, or danger which is not readily known. He posits that even in the presence of real danger there can be two reactions, (a) an affective response or outbreak of anxiety, and (b) an instrumental response leading to protection from the danger. When there is purposeful cooperation between these reactions, the anxiety becomes a signal for the initiation of the instrumental response leading to safety. However, there can also be conflicting reactions whereby the anxiety immobilizes the instrumental response: "paralysis through fear" he calls it. Thus from Freud's point of view, anxiety arousal can facilitate instrumental behavior under certain circumstances, but immobilize it under other more unfavorable circumstances.

Traumatic experience is the basis for the development of neurotic anxiety. In order to protect onself from similar traumas in the future, one begins to foresee and anticipate the presence of danger. It is immaterial whether or not real danger is present. Situational cues associated with the earlier trauma elicit the danger signal, anxiety. Freud relates this psychodynamic account:

> I anticipate that a situation of helplessness will come about, or the present situation reminds me of one of the traumatic experiences which I have previously undergone. Hence I will anticipate this trauma; I will act as if it were already present as long as there is still time to avert it. Anxiety, therefore, is the expectation of the trauma on the one hand, and on the other, an attenuated repetition of it. . . . The danger situation is the recognized, remembered, and anticipated situation of helplessness. Anxiety is the original reaction to helplessness in the traumatic situation [Freud, 1936, p. 114].

Thus arousal of anxiety in conjunction with feelings of helplessness to remove the perceived threat, whatever it might be, would produce less

appropriate behaviors than in the case where one felt confident in his ability to deal with a potential trauma.

It is but a short step from Freud's conception of the psychogenesis of anxiety to Mandler and Sarason's concept of "test anxiety." Because of felt inadequacies and helplessness in the face of achievement situations, individuals come to fear and to experience heightened levels of anxiety in the face of them. Feelings of helplessness and anticipation of punishment for poor performance would be covert responses which interfere with the acquisition and emission of desired instrumental behavior.

Anxiety and Academic Achievement

In the previous chapter it was concluded that anxiety arousal has a detrimental effect on the acquisition of new, complex responses. In terms of academic achievement situations, a translation from theory would state that when habit strength of cognitive responses (e.g., content knowledge and problem solving skills relevant to course goals) is high, anxiety arousal should be least detrimental and could even facilitate the emission of these responses. On the other hand, unfamiliarity with content and problem solving skills (low habit strength) at the time of testing should interact with anxiety arousal to impede recall of content poorly assimilated as well as blocking spontaneous learning on test items requiring a high degree of unpracticed problem solving skill or new insight.

If the assumption that anxiety interferes with response acquisition is valid, then it seems likely that what the student does in the

way of studying and verbal rehearsal prior to test performance is a more potent determiner of achievement level than the level of arousal during the examination itself. This is not to say that anxiety arousal is irrelevant to academic performance. It may be that manifestation of anxiety takes its heaviest toll when students confront reading assignments and other achievement related tasks during initial contacts when unfamiliarity with achievement goals is maximal.

But what determines whether or not a student engages in the instrumental behaviors essential for academic achievement? The most plausible explanation would seem to lie in the area of motivation and motive arousal.

Before proceeding further, definitions and assumptions with regard to the term motive are in order. Motive, as used here, refers to behavioral dispositions that strive for general goal-states which are gratifying to the individual. It is assumed that these behavioral tendencies are formed during childhood experience and are relatively stable over time. Motives are thought to be latent until situational cues arouse them. A motive is aroused when the person perceives or anticipates that he will gain satisfaction through the performance of an instrumental act. Specific experiences in the person's life history determine the content of motives in terms of goal-states and the preferred instrumental acts which lead to motive satisfaction.

Achievement and Fear of Failure

The Theory

The motive to avoid failure is assumed to be a disposition to

become anxious about the possibility of failing whenever the person perceives that his instrumental behavior will be evaluated. Therefore, a person motivated in this way finds achievement tasks unattractive and will avoid them unless constrained by external social forces, e.g., regulations specifying levels of performance necessary for maintaining one's status as a college student (Atkinson and Litwin, 1960).

Achievement tasks are unattractive because they arouse anxiety and feelings of helplessness in dealing with them. The avoidance motive functions to reduce anxiety and the anticipated pain of mortal embarassment associated with inadequate performance. Since achievement tasks are unattractive it would seem reasonable to assume that situations associated with achievement, e.g., engaging in rehearsal of academic assignments, would also be unattractive and therefore avoided. When penalties for low performance are present and inescapable, the fear of Failure person would be in a state of conflict because avoidance of achievement tasks would result in the undesirable consequence of failure. Consequently, approach tendencies must overcome avoidance tendencies if failure is to be avoided at all. Thus, we have a picture of the fear of Failure person as the last minute "crammer." How effective his cramming will be depends upon his ability and the magnitude and complexity of the achievement task. Cramming strategies could result in low academic performance when the body of content to be assimilated is large or especially complex.

Since it is assumed that the fear of Failure person must be constrained by social forces before he will engage in achievement behavior maximally, this assumption also suggests that he is principally

motivated by factors extrinsic to himself. Extrinsic factors would include the frequency of public evaluation, the magnitude and complexity of achievement tasks, and the severity of penalties for poor performance Since these variables are controlled by the teacher, the manner in which he structures course content and achievement tasks could have an effect on the ultimate level of performance of students high in fear of Failure.

Hypotheses

In order to assess the validity of assumptions underlying the theoretical model presented above, several hypotheses are proposed for empirical testing.

> Hypothesis I: Students high in fear of Failure will tend to achieve higher grades under conditions of high teacher produced evaluative stress than under conditions of low evaluative stress.

The rationale for this hypothesis stems from characteristics of the achievement situation. Although the college classroom clearly satisfies the condition of constraint, the instrumental behavior prior to achievement testing is assumed to be the key determiner of level of performance. High evaluative stress in the form of frequent demands for achievement should increase effort expended toward avoidance of failure. Low evaluative stress would minimize the demand for achievement and hence reduce the perceived necessity for adequate performance. In view of the magnitude and complexity of the content represented in course examinations, cramming should be less effective in increasing levels of performance for large bodies of content arising from infrequent evaluation, than in situations where the magnitude of the task is small. This prediction is not at variance with Atkinson's interpretation of

constraint in that both high and low fear of Failure students would be highly motivated during the achievement test. The major difference lies in the quantity of achievement related instrumental behavior prior to the testing situation when constraint is less evident.

> Hypothesis II: Evaluative stress should have little, if any, systematic effect on the academic performance of students low in fear of Failure.

This hypothesis is based on the assumption that achievement tasks and achievement related behaviors are not seen as being unattractive by low fear of Failure students, or as arousing affects which would be disruptive to attaining achievement goals.

Although it has often been assumed that stress disrupts the performance of anxious individuals, studies in "real life" situations have challenged this assumption. A careful review of the experimental literature suggests that this assumption should be modified to emphasize the fact that disruptive effects seem to be maximal during early stages of learning. On the other hand, if we assume adequate levels of habit strength, would test-anxious subjects perform equally well under high and low levels of stress? Before answering this question, Atkinson or Carney (1961) would determine the degree of constraint surrounding the achievement situation. Which is most relevant, then, anxiety level or degree of constraint? It could well be that both are operative thereby creating difficulties in designing a suitable experiment to separate these effects. Since stress is usually induced by a threat of public evaluation, how can a satisfactory testing condition be created where evaluative stress is minimized within acceptable limits of social

constraint? Evaluative stress implies that there is a payoff for good performance and penalties for inadequate performance. The latter point is a necessary condition underlying constraint. Therefore, constraint does not appear to be independent of evaluative stress but to be positively related to it. What is a relaxed condition? Is it chiefly an unconstrained one?

In order to clarify this issue, the additional hypotheses are proposed. On the assumption that stress inhibits performance in testing situations, it is predicted that.

> Hypothesis III: Students high in achievement anxiety (fear of Failure) will exhibit lower levels of performance in testing situations perceived as being high in evaluative stress than in testing situations perceived as being low in evaluative stress.

However, in order to be consistent with the rationale underlying the avoidance motive and its relationship to social constraint, the supplementary hypothesis should also be considered.

> Hypothesis IV: Students high in achievement anxiety (fear of Failure) will exhibit higher levels of performance in testing situations that are perceived as being high in constraint than in testing situations low in constraint.

Conditions specified in both hypotheses are similar. The essential difference stems from assumptions concerning causality. Hypothesis III assumes the causal variable to be anxiety arousal and consequent disruption of performance. Hypothesis IV predicts the opposite outcome based on the interaction of the avoidance motive with constraint, i.e., penalty for nonachievement. It seems clear that anxiety will be aroused under the stress condition regardless of the hypothesis specified. The

observed outcome should help to distinguish whether anxiety arousal per se is a principal factor in affecting performance, or whether motive arousal interacts with the dimension of constraint to affect performance also.

What is the effect of stress and constraint on the test performance of low fear of Failure subjects? The theoretical assumptions lead us to predict that:

> Hypothesis V: Students low in achievement anxiety (fear of Failure) will be less affected by differences in degree of constraint than subjects high in achievement anxiety.
>
> Hypothesis VI: Students low in achievement anxiety (fear of Failure) will perform better than students high in achievement anxiety to the extent that the achievement test is relatively brief and unfamiliar.

Performance on tasks which are simple or overlearned are relatively ineffective in differentiating groups varying in anxiety level. Therefore, the factors of time and unfamiliarity are specified in order to enhance the opportunity for performance decrements to occur.

In view of results from previous studies on achievement and avoidance motivation (McClelland et al. 1953; Atkinson, 1958), the predicted outcomes are anticipated with more confidence for male than for female subjects.

The Variables

Fear of Failure

Strength of the motive to avoid failure will be inferred from level of achievement anxiety as measured by the debilitating anxiety (DA) scale of the Achievement Anxiety Test (Haber and Alpert, 1960).

A high score on the debilitating anxiety scale indicates manifestations of nervousness, etc., which lower one's level of test performance. The presence of these manifestations are assumed to be indicative of the presence of an irrational fear of failing.

Social Constraint

A social-psychological dimension which implies that the person will suffer loss of status and self-esteem by withdrawing from required achievement tasks. Carney (1961) presents evidence to show that this dimension exists and is relevant to academic achievement situations. For the purposes of this study, constraint is determined by the University regulation requiring satisfactory academic performance as a condition for maintaining one's status as a student. Achievement situations not under the jurisdiction of this regulation are assumed to be of low constraint.

Evaluative Stress

Evaluative stress is aroused by situational cues which imply that achievement is expected and that performance is being evaluated. Whether or not cues are stressful depends on the legitimacy of penalties invoked for inadequate performance, legitimacy referring to the constraint dimension. Level of evaluative stress will be inferred from self-reports of perceived stress and the relative frequency of teacher behaviors assumed to be indicative of an emphasis on evaluation.

Study Habits and Persistence

Study habits will be inferred from the subject's score on a

scale of study habits and attitudes (Brown and Holtzman, 1955). It is assumed that achievement test performance depends on the quality and quantity of cognitive rehearsal prior to testing. Other things being equal, persons who have developed systematic approaches in dealing with subject matter should attain higher levels of achievement, in the long run, than persons who have poor instrumental study habits.

Academic Performance

During the course of a semester, students have the opportunity to adapt to anxiety arousing stimuli emitted by the instructor, to become familiarized with new content areas, etc. Course grade has been chosen as the dependent variable to examine the effect of anxiety arousal over long periods of time. In order to study the short term effects of anxiety arousal, students will be asked to complete the Introductory Psychology Criteria Test (Milholland, 1964). This test was selected because it is based on the content of introductory psychology and the fact that the test items require complex problem solving skills. Since problems on this test are relatively unique, a component of learning, i.e., unpracticed problem solving, is assumed necessary for successful completion thereby enhancing the opportunity for disruptive anxiety effects to occur.

A more detailed description of the instruments used to assess the above variables can be found in Appendix C. The research design used to test Hypotheses I and II is presented in Chapter III; the remaining hypotheses are treated in Chapter IV.

CHAPTER III

THE SURVEY STUDY

Results presented in this chapter were generated by the research design which was formulated in order to test Hypotheses I and II. Subject populations, assessment devices, and procedures incorporated into the design are described below.

Procedure

Course Settings

The subject population included students enrolled in the Introductory Psychology Course and the Introductory Economics Course at the University of Michigan, fall semester of 1964. Both were four credit hour courses requiring a large group lecture and three small discussion group meetings per week. Lectures were given by senior staff members with the discussion groups being conducted by teaching fellows. The two courses differed greatly in evaluation procedures, however.

In Psychology, the teaching fellows were given full responsibility for test construction, scoring criteria, and assignment of final grades. In Economics, department wide examinations were given and scoring of tests was supervised to insure uniform grading criteria. Moreover, Economics grades were assigned on the basis of an absolute scale of test score points, whereas teaching fellows in Psychology

tended to assign grades from the distribution of scores relative to their own sections.

Teaching fellows in both departments were supervised by a senior staff member who also was the lecturer. Weekly meetings were held for the purpose of discussing course-related problems and coordinating instruction in general. Both departments administered a student rating form for the purpose of evaluating the course and effectiveness of instruction.

Student Populations

The sampling goal was to survey the entire Introductory Psychology Course and the endeavor was successful in large part. Of the 1,131 students who received a final grade in the course, complete data were obtained from 370 men and 577 women. Twenty-five subjects were rejected from the initial survey because they lacked ability test scores. One teaching fellow failed to administer the pre-test instruments as planned, which accounted for 29 additional subjects. The remaining 130 students were absent from class on the day pre-testing was done; no follow-up was attempted on this group. Therefore, the study sample included subjects from 39 of the 40 sections of Introductory Psychology or 84 per cent of the total enrollment.

In order to replicate the study on an independent sample, the Introductory Economics Course was also surveyed. Useable data were obtained on 338 men and 149 women. Thirteen subjects were rejected because they lacked ability test scores and 98 were absent from class

on the day pre-test data were collected. Subjects in two sections of the course had to be rejected because another student rating form had been used to assess the teacher's behavior. Therefore, the sample from the Economics Course included 26 of the 28 sections, and 488 of the 613 students who received a final grade in the course. For a more complete description of the study samples see Appendix A.

The Introductory Economics Course included four sections for honors students while there were none designated as such in Introductory Psychology. Because grades were determined on an absolute scale in Economics and the fact that level of ability was to be controlled in the statistical analysis, honors sections were kept in the study sample.

Teacher Samples

The 39 sections included in the Psychology sample were taught by eight female and 15 male instructors, four of whom were Ph.D's. The remaining group consisted of second and third year graduate students with varying degrees of previous teaching experience. The 26 sections comprising the Economics sample were taught by two female and 13 male instructors. This group also consisted of graduate students at various levels of advanced study and years of previous teaching experience.

Measures of Student Characteristics

Student characteristics of interest in this study were study habits and skills, debilitating anxiety, and intellectual ability. The study habits scale consisted of 20 items which had been selected from the Brown-Holtzman Survey of Study Habits and Skills[1] (Brown and

---

[1]Permission has been obtained from the Psychological Corporation

Holtzman, 1955). The content of these items deals with use of study time, organization of course material, preparing for examinations, integration of subject matter, etc. Fear of Failure was inferred from the debilitating anxiety sub-scale of the Achievement Anxiety Test (Haber and Alpert, 1960). A complete description of the study habits and anxiety scales may be seen in Appendix C. The linguistic score on the American Council on Education Psychological Test (ACE) and verbal score on the Scholastic Aptitude Test (SAT) were used as the measures of ability.

A 10-item Subjective Report Questionnaire was administered to students in Economics. These same 10 items were incorporated into the 21-item rating form used in Psychology (Appendix D). This modification was employed in order to avoid duplication of items already included in the Economics Course rating form. Items in the Subjective Report Questionnaire included 5-interval Likert-type scales pertaining to effort expended in studying, difficulty of the course, perceived pressure and anxiety, as well as open ended questions asking students to explain their particular endorsement of the Likert-type items. The purpose of this instrument was to provide self report data which could be used to validate our measure of evaluative stress.

Measures of Teacher Behavior

Teacher evaluation forms are used in both departments as a matter of standard procedure. Therefore, a five-item instrument designed to assess teacher behaviors thought to be indicative of evaluative

---

to use this measure for research purposes.

stress was included in the rating form ordinarily used by each department. These items required the student to judge the relative frequency with which the teacher emphasized grades, threatened to give unannounced quizzes, judged performance in class, etc. One other item asked students to estimate the amount of competition for grades in order to determine the possible influence of peer group phenomena on academic performance. Relationships between the teacher's behavior and competition for grades could also be determined from this item (Appendix D). Teaching fellows were asked to fill out a questionnaire describing the frequency and kinds of assignments which they used to evaluate their students (Appendix E). The teacher's formal demands for achievement in the form of tests, term papers, etc., were assumed to be concrete forms of evaluative stress.

Data Collection

The questionnaire containing the study habits and anxiety scales was administered to each class by the regular teaching fellow. In order to standardize testing procedures, teaching fellows were asked to read aloud a set of instructions provided by the experimenter. An introductory letter familiarizing students with the study preceded the testing. (The instructions and the introductory letter can be found in Appendix B.) So that pre-testing could be incorporated into functioning of classes as smoothly as possible, each teacher selected the date on which the tests would be given. Data collection of student characteristics was begun on the sixth week and completed by the eleventh week of the semester.

Ability test scores were obtained from the University's Evaluations and Examinations Division. The ACE and SAT had been administered prior to election of the course. In order to use all subjects for whom ability test scores were available, ACE percentile ranks were converted to comparable scaled scores in the SAT distribution. This was done by means of a conversion table based on norms for University of Michigan Freshman compiled in 1955 (Appendix F).

The student rating forms were administered at the close of the semester. A student in each class was placed in charge of distributing and collecting the forms. In keeping with the usual policy, instructors left the classroom while the rating forms were being filled out. Completed forms were not returned to the instructor for his use until after all final grades had been recorded.

Students were not asked to identify themselves on the student rating forms. However, they were asked to indicate their names on the face sheet of the Subjective Report Questionnaire. These face sheets were removed and identification recoded before these forms were made available to teaching fellows. In this way the identity of respondents remained unknown to the teaching staff.

Coding of Open-ended Items

The rationale for using open-ended questions in the Subjective Report Questionnaire was based on the need for validating the evaluative stress scale. It seemed reasonable to assume that if the teacher's behavior has a significant impact on motivation to achieve, then students would be able to report this in response to questions such as

"What factors do you think account for the amount of studying that you did in this course?" On the basis of a priori considerations, it seemed reasonable to expect that significant motivational effects could stem from (a) the teacher's demands for achievement, (b) the student's attraction for the content, and (c) the level of difficulty of the subject matter itself. Thus, the goal was to separate motivational effects that could be due to any one of these three basic factors.

In order to derive a coding scheme, a large sample of questionnaires was examined and responses were categorized into the areas mentioned above as well as new areas suggested by the responses themselves. On the basis of this sampling procedure, several nominal categories were defined for items 7, 9, and 11 of the Subjective Report Questionnaire. The coding schemata for these items can be found in Appendix G.

Actual scoring was begun only after the two coders had worked on practice materials. Per cent agreement for four independent blocks of approximately 80 practice trials was 80 per cent, 89 per cent, 89 per cent, and 91 per cent. It was observed that discrepancies during the practice trials seemed to be due to human errors rather than to ambiguities inherent in the coding categories. Because of the relatively weak nominal scale employed, it seemed important to keep scoring errors at an absolute minimum. Therefore, both coders scored all of the questionnaires independently. By continuously checking for interscorer agreement, errors were detected. Whenever disagreements occurred, they were resolved by discussion so that the overall agreement was very close to 100 per cent.

Levels for the Treatments x Levels Design

In view of the reported intercorrelations between specific anxiety scales, intellectual ability, and academic achievement, statistical techniques were used to control for the effect of these dimensions on the dependent variable, course grade. Therefore, a treatments x levels (fixed effects) analysis of variance design was used to estimate the level of significance of the results. However, the existence of correlations among the independent variables created problems in maintaining equal cell frequencies for the statistical design. When test score distributions were trichotomized independently, it was found that certain cells were nearly vacant while other cells were disproportionately full. This was due chiefly to the inverse relationship between debilitating anxiety and ability (see Tables 2 and 3). For example, very few subjects were found in the high ability-high anxiety or low ability-low anxiety cells. In order to meet the condition of equal cell frequencies and still preserve matching of subjects across treatments, the following procedure was used.

1. Trichotomize the ability distribution first since this variable is most highly correlated with the dependent variable.
2. Trichotomize anxiety scores within each level of ability.
3. Dichotomize study habits scores within each level of anxiety.

Although this procedure resulted in different breaking points on study habits and anxiety across levels of ability, subjects were well-matched within each level of ability and across conditions. But even with the use of this modified classification procedure, there were still

unequal cell frequencies for any given four-dimensional array. Therefore, it was decided to strive for a minimum of five subjects per cell. After being classified on the four dimensions, subjects within a cell were numbered-off and then five were selected by means of a random numbers table.

Due to chance factors in sampling and the difficulty encountered in classifying subjects simultaneously on four dimensions, it was necessary to supply missing data in order to maintain five replications per cell. Scores for missing subjects were determined by computing the mean of a given cell from the actual cases and then using this value as the missing score (Lindquist, 1958). Of the 12 tables used in the data analysis, there were two tables with complete data, four tables requiring at least one synthetic score, five requiring two synthetic scores, and only one table requiring five synthetic scores. In only one instance did a deficient cell require two synthetic scores, with all other deficient cells requiring one.

To enhance the possibility for anxiety effects to emerge, only subjects from the upper 33 per cent and lower 33 per cent of the anxiety score distribution were included in the analysis. The actual scores used for determining levels for each variable are given in Appendix J.

When the 149 female subjects in Economics were classified on all four dimensions, there were many cells with frequencies far short of the five replications desired. Therefore, it was necessary to omit these subjects from the four-way analysis of variance.

Standardization of Grades

Since assignment of grades in Psychology was left to the discretion of each teaching fellow, grades within each section were converted to standard scores with a mean of 50 and a standard deviation of 10. This was not done in Economics because of the use of uniform testing procedures and assigning of grades on the basis of an absolute scale of test score points.

Results

Intercorrelations of Student Characteristics

Independent and dependent variables were intercorrelated prior to carrying out the multivariate analysis. The Pearson product-moment coefficients are reported in Tables 2 and 3. Since sex differences were anticipated, analyses were carried out separately for men and women. Distribution characteristics for these variables can be found in Appendix H. Reliability coefficients are reported in Appendix I.

Teacher Ratings

The experimental hypotheses presupposed that there would be sufficient variability in teacher ratings, from class to class, to permit a clear evaluation of their effect. This did not prove to be the case as can be observed in Tables 4-7. Mean item ratings tended to fall below the neutral point indicating a low level of evaluative stress cues stemming from the teacher's behavior.

The relative frequency of competition for grades was somewhat more promising as mean ratings did span the neutral point with five out

TABLE 2

INTERCORRELATION OF VARIABLES FOR 370 MEN AND
577 WOMEN IN INTRODUCTORY PSYCHOLOGY
(Fall, 1964)

		DA	SATV	Grades
SH	Men	-.235**	.010	.181**
	Women	-.297**	.038	.220**
DA	Men		-.334**	-.251**
	Women		- 268**	-.151**
SATV	Men			.419**
	Women			.402**

*p ≤ .05 (2 tailed) **p ≤ .01 (2 tailed)

TABLE 3

INTERCORRELATION OF VARIABLES FOR 338 MEN AND
149 WOMEN IN INTRODUCTORY ECONOMICS
(Fall, 1964)

		DA	SATV	Grades
SH	Men	-.204**	.022	.205**
	Women	-.177	.053	.312**
DA	Men		-.280**	-.179**
	Women		-.292**	-.202*
SATV	Men			.424**
	Women			.330**

*p ≤ .05 (2 tailed) **p ≤ .01 (2 tailed)

TABLE 4

MEAN ITEM RATINGS FOR ITEM ONE[a]

Rating	Score Range	Number of Sections	
		Psych	Econ
Very often	1.00 - 1.49	0	0
Often	1.50 - 2.49	0	0
Occasionally	2.50 - 3.49	0	0
Seldom	3.50 - 4.49	1	4
Almost never	4.50 - 5.00	38	22
Total		39	26

[a]How often did the instructor warn the class to expect unannounced quizzes?

TABLE 5

MEAN ITEM RATINGS FOR ITEM TWO[b]

Rating	Score Range	Number of Sections	
		Psych	Econ
Four or more times	1.00 - 1.49	0	0
Three times	1.50 - 2.49	0	0
Two times	2.50 - 3.49	0	2
One time	3.50 - 4.49	1	4
At no time	4.50 - 5.00	38	20
Total		39	26

[b]Unannounced quizzes were given in this class. . . .

TABLE 6

MEAN ITEM RATINGS FOR ITEM THREE[c]

Rating	Score Range	Number of Sections	
		Psych	Econ
Very often	1.00 - 1.49	0	0
Quite often	1.50 - 2.49	0	0
Sometimes	2.50 - 3.49	9	0
Hardly ever	3.50 - 4.49	30	25
Almost never	4.50 - 5.00	0	1
Total		39	26

[c]The teacher strongly emphasized grades. . . .

TABLE 7

MEAN ITEM RATINGS FOR ITEM FIVE[d]

Rating	Score Range	Number of Sections	
		Psych	Econ
Almost always occurred	1.00 - 1.49	0	0
Often occurred	1.50 - 2.49	0	0
Occasionally occurred	2.50 - 3.49	4	0
Seldom occurred	3.50 - 4.49	34	24
Almost never occurred	4.50 - 5.00	1	2
Total		39	26

[d]The teacher made it clear that he was continually judging the worth of a student's performance in class.

of 65 sections being rated as having competition often. However, in spite of the increase in relative frequency, variability was still restricted. The distribution of ratings for item four can be seen in Table 8.

TABLE 8

MEAN ITEM RATINGS FOR ITEM FOUR[e]

Rating	Score Range	Number of Sections	
		Psych	Econ
Most of the time	1.00 - 1.49	0	0
Often	1.50 - 2.49	4	1
Occasionally	2.50 - 3.49	32	25
Seldom	3.50 - 4.49	3	0
Hardly at all	4.50 - 5.00	0	0
Total		39	26

[e]Most students in class competed for grades. . . .

It had been assumed that the instructor's demands for achievement via graded tests would be an important source of stress arousal. Therefore, the actual frequency of graded tests was determined. Teaching fellows in Psychology varied in the number of hour examinations assigned. In Economics, hour examinations were given on a departmental basis and thus did not vary in number. However, Economics teaching fellows were given the option of assigning graded quizzes. The frequency distributions for teacher assigned tests and quizzes are given in Table 9.

TABLE 9

FREQUENCY OF TEACHER ASSIGNED TESTS

Frequency	Number of Sections	
	Psych: Hour exams	Econ: Quizzes
Four or more tests	2	4
Three tests	12	7
Two tests	20	8
One test	5	7
None	0	0
Total	39	26

Before summing item ratings to obtain an overall scale score for a teacher, item means were intercorrelated to determine whether or not a single dimension was being tapped. High intercorrelations between items would be taken as justification for deriving a combined score. Because of the extremely restricted range of scores, items one and two were eliminated from further analysis. The item intercorrelations are given in Table 10.

There were only two coefficients significantly different from zero. In Psychology, students were seen as competing for grades in classes taught by instructors rated as tending to emphasize them ($r = .58$). This relationship did not appear in the Economics course ($r = -.08$), although Economics instructors who tended to give more quizzes were also rated as emphasizing grades ($r = .47$).

In view of the low degree of relationship between items as well as the lack of similarity in correlational patterns between the two

TABLE 10

INTERCORRELATION OF ITEM MEANS FOR TEACHER PRODUCED EVALUATIVE STRESS

Item Number		ES4	ES5	Number of tests
ES3[a]	Psychology	.58**	-.03	.28
	Economics	-.08	-.15	.47*
ES4[b]	Psychology		-.28	.19
	Economics		.30	.11
ES5[c]	Psychology			-.22
	Economics			.03

**p = .01 *p = .05

[a]The teacher strongly emphasized grades. . . .

[b]Most students in class competed for grades. . . .

[c]The teacher . . . judged the . . . student's performance in class. . . .

courses, the hypotheses were tested separately for the teacher behaviors assessed. In this way, the motivational effect of each particular behavior could be determined. Because of the narrow range of evaluative stress cues, only classes falling in the upper 33 per cent and lower 33 per cent of each item distribution were used in the statistical analysis.

The Analysis of Variance

The predicted interaction between fear of Failure and teacher produced evaluative stress failed to materialize. No significant main effects for the teacher behaviors were observed either. (Variance estimates, F-ratios, and probability levels for the complete analysis can be found in Appendix K.) However, there were several main effects for the

student measures which were consistent with results found in other studies using these same variables. For example, a significant main effect for ability was observed in 11 of the 12 analyses.[2] Students with high ability earned higher grades than students of low ability, with middle ability students achieving an intermediate level of performance (Table 11). This was true for males and females in Psychology as well as for males in Economics. These effects were consistent with the results of the correlational analysis presented in Tables 2 and 3.

TABLE 11

MEAN GRADE BY LEVEL OF ABILITY FOR STUDENTS IN PSYCHOLOGY AND ECONOMICS[a]
(Fall, 1964)

Sample	Low SATV	Mid SATV	High SATV	df	F	P
Men in Psychology	46.58[b]	50.80	52.90	2,96	5.47	.01
Men in Economics	2.08[c]	2.68	2.95	2,96	10.01	.005
Women in Psychology	48.30[b]	48.83	55.25	2,96	7.38	.01

[a]Results in this table are typical of those observed. Means by level of ability for all 12 analyses can be found in Table 46 of Appendix K.

[b]Grades in Psychology were converted to standard scores within sections (Mean of 50, S.D. of 10).

[c]Means in Economics are based on raw grades (A=4, B=3, C=2, D=1, E=0).

[2]Subjects included in the analysis of variance for one evaluative stress item were returned to the sample before subjects were drawn at random for the next test. Therefore, tests of significance for samples drawn from a given population, e.g., men in Economics, are not independent of each other.

Main effects for study habits and debilitating anxiety also appeared but with less consistency. Students scoring high in study habits earned higher grades than students low in study habits (Table 12). This was more marked for men than for women. Although effects tended to be in the expected direction for women in Psychology, all failed to reach the 5 per cent level of significance (Table 47, Appendix K).

TABLE 12

MEAN GRADE BY LEVEL OF STUDY HABITS FOR STUDENTS IN PSYCHOLOGY AND ECONOMICS[a] (Fall, 1964)

Sample	Low SH	High SH	df	F	P
Men in Psychology	49.07[b]	50.70	1,96	.80	N.S.
Men in Economics	2.30[c]	2.65	1,96	4.71	.05
Women in Psychology	48.22[b]	50.47	1,96	1.55	N.S.

[a]Results in this table are typical of those observed. Means by level of study habits for all 12 analyses can be found in Table 47 of Appendix K.

[b]Grades in Psychology were converted to standard scores within sections (Mean of 50, S.D. of 10).

[c]Means in Economics are based on raw grades (A=4, B=3, C=2, D=1, E=0).

Debilitating anxiety tended to differentiate levels of performance more consistently in Psychology than in Economics. Psychology students low in debilitating anxiety received higher grades on the average than students high in debilitating anxiety (Table 13). This effect was barely discernible for men in Economics as only one of four

TABLE 13

MEAN GRADE BY LEVEL OF DEBILITATING ANXIETY FOR STUDENTS IN PSYCHOLOGY AND ECONOMICS[a] (Fall, 1964)

Sample	Low DA	High DA	df	F	P
Men in Psychology	52.25[b]	48.07	1,96	4.85	.05
Men in Economics	2.50[c]	2.42	1,96	.27	N.S.
Women in Psychology	52.77[b]	49.05	1,96	5.66	.025

[a]Results in this table are typical of those observed. Means by level of debilitating anxiety for all 12 analyses can be found in Table 48 of Appendix K.

[b]Grades in Psychology were converted to standard scores within sections (Mean of 50, S.D. of 10).

[c]Means in Economics are based on raw grades (A=4, B=3, C=2, D=1, E=0).

tests approached the 5 per cent level of significance for these subjects (Table 48, Appendix K).

The analysis of variance generated significant F-ratios for 1 three-way and 2 two-way interactions (Tables 35, 37, and 45 in Appendix K). Means for each interaction were plotted within each course group as well as across independent samples. Interaction patterns were inconsistent both within and between groups. Moreover, there were no theoretical reasons to expect these outcomes. Because of the inconsistencies in the interaction patterns, as well as the large number of nonindependent tests carried out, it was concluded that these interaction effects were most likely due to errors generated by sampling rather than to any underlying psychological phenomena.

In view of the inconclusiveness of results for the evaluative stress dimension, data obtained from the Subjective Report Questionnaire were omitted from the remainder of this particular study.

Discussion

The purpose of this study was to examine the long term effects of anxiety arousal in academic achievement situations. It was predicted that students high in fear of Failure would achieve higher levels of academic achievement under high than under low levels of teacher produced evaluative stress. The results obtained clearly fail to support this hypothesis. However, results observed for the student characteristics of intellectual ability, study habits, and debilitating anxiety provide additional evidence for these measures as predictors of academic achievement.

Ability, as measured by the verbal scales of the ACE and SAT, was found to be the most potent predictor in both Economics and Psychology (Table 11). But some interesting course differences are suggested by the pattern of results obtained for the study habits and debilitating anxiety scales. For example, a study habits main effect appeared more consistently for men in Economics than for men and women in Psychology (Table 47, Appendix K). Contrariwise, a main effect for anxiety tended to be more consistent for Psychology students than for men in Economics (Table 48, Appendix K). These outcomes, though they are not conclusive, suggest the possibility of a course x anxiety x study habits interaction. On the other hand, sampling error could also

have accounted for the lack of consistency in view of the large number of samples drawn from each course and sex group. Therefore, multiple correlation techniques were used in order to study the effect of these variables on academic performance more precisely. Results of this analysis are given in Table 14.

TABLE 14

ZERO-ORDER AND MULTIPLE CORRELATIONS FOR ABILITY, STUDY HABITS, AND DEBILITATING ANXIETY IN PREDICTING GRADES IN PSYCHOLOGY AND ECONOMICS (Fall, 1964)

	Grade·SATV	Grade·SATV, DA	Grade·SATV, SH
Men in Psychology (N = 370)	.42	.44	.45
Men in Economics (N = 338)	.42	.42	.46
Women in Psychology (N = 577)	.40	.40	.45

(All coefficients are significant: p = .01 or less, 2 tailed.)

As can be seen in Tables 2 and 3, verbal ability is the best single predictor of achievement of the student measures used in this study. The noncognitive variable of study habits accounts for additional variance when used in combination with ability (Table 14), but the magnitude of this gain (2.6 to 4.3 per cent) is small compared to the variance accounted for by ability (16 to 17.6 per cent). Anxiety in combination with ability failed to enhance our predictive power for women in Psychology and men in Economics although a small gain was

observed for men in Psychology. On the basis of these results, it seems most reasonable to conclude that inconsistencies observed in the results within and between various samples were due to sampling errors rather than to course differences, such as assessment procedures or the course content itself.

What of our attempt to clarify the relationship between measures of verbal aptitude and specific anxiety scales? Results of the multiple correlational analysis and the analysis of variance provide contradictory evidence in that expected anxiety effects did not emerge with any reassuring consistency when intellectual ability was taken into account. If anxiety is a noncognitive component of intellectual performance, these results suggest that its effect on academic achievement is small and can be mitigated by compensatory behaviors such as good study habits (Economics men and Psychology women, Table 14).

How can we account for the lack of significant effects for the evaluative stress dimension? Although results obtained with this variable failed to support our predictions, it does not necessarily follow that this dimension is irrelevant to academic performance. It may well be that the teacher behaviors used in operationalizing this variable were insensitive indicators as suggested by the lack of variability among teachers as well as the relatively infrequent occurrence of these behaviors in the classrooms surveyed. Moreover, the use of untrained observers (the students) may have resulted in less reliable assessments than those obtainable by more highly trained observers. The fact that students were asked to make one global rating at the conclusion of the

course could reduce the precision of measurement, also. A single rating such as this would be less reliable than repeated judgments based on time sampling procedures.

On the basis of these results, what do we conclude? First of all, our central hypothesis remains unconfirmed as we have not been able to show that there are any long term effects on academic achievement due to anxiety aroused by the teacher's behavior. On the more positive side, it seems clear that measures of ability continue to be the best predictors of academic achievement. Although the noncognitive variables of study habits and debilitating anxiety can account for variance in addition to that which is accounted for by measures of ability, the gain in predictive power is small.

Summary

This study was designed to investigate the long term effects of anxiety arousal in academic achievement situations. Relevant hypotheses were derived from a theoretical model based on the properties of an avoidance motive, the fear of Failure.

Independent samples of subjects were obtained from an introductory Economics and an introductory Psychology course at the University of Michigan. A multivariate analysis was carried out using intellectual ability, study habits, debilitating anxiety, and evaluative stress (as inferred from ratings of teacher behavior) as independent variables. The dependent variable was course grade.

The central hypothesis that students high in fear of Failure would achieve higher grades in classes characterized by high rather than

low evaluative stress was not supported by the data. However, a strong main effect for verbal ability was observed. Results derived from a multiple correlational analysis suggested that the noncognitive variables of study habits and debilitating anxiety enhance accuracy of predicting grades very little when used in combination with measures of ability.

Methodological problems in assessing teacher behavior were discussed. It was suggested that the precision of measurement could be enhanced through the use of trained observers and time sampling procedures.

CHAPTER IV

THE EXPERIMENTAL STUDY

Procedure

Student Subjects

Subjects for the evaluative (high stress) condition were obtained through the cooperation of five teaching fellows who agreed to use the Introductory Psychology Criteria Test as part of their final examination. There were students from nine sections involved. A small number of these subjects had been absent during the survey study pretest so that they lacked study habits and anxiety test scores. Complete data were obtained for 96 men and 99 women in the evaluative condition.

Subjects participating in the nonevaluative (low stress) condition, were enrolled in the Introductory Psychology Course also, and were obtained from the departmental subject pool. These subjects were drawn from 14 of the 39 course sections and comprised a reasonably representative sample of subjects enrolled in Introductory Psychology. They were assigned to one of two evening testing sessions on the basis of their time schedules. Of the 223 subjects contacted, 178 were able to participate in the two regular sessions. A make-up session was held and 21 more persons completed the test forms. After a series of follow-up attempts, complete data were obtained on 101 men and 108 women (or 209 of the original 223) for the low stress condition.

Measures

Student characteristics of interest in the experimental study were study habits, debilitating anxiety and intellectual ability. These measures have already been described in Chapter III. The dependent variable was score on the Introductory Psychology Criteria Test, Form X (Milholland, 1964). This test was constructed in order to assess cognitive abilities as defined in the Taxonomy of Educational Objectives (Bloom, 1956). Items are based on the content of Introductory Psychology and are designed to tap intellectual abilities such as interpretation, analysis of relationships, judgments in terms of external criteria, etc. (see Appendix C for a more detailed description of this scale). The fact that complex cognitive abilities were being assessed provided a relatively unique task for subjects to perform. It was anticipated that unfamiliarity with these particular kinds of test problems would bring out differences in the performance of groups differing in level of anxiety and stress.

Data Collection

The nonevaluative or low stress testing sessions were held during the eleventh week of the semester. Subjects were asked to meet in one of two large groups. The testing room was large enough so the approximately 80 subjects could be placed in alternate seats. After three female assistants distributed Criteria Test booklets, IBM answer sheets, and pencils, the following instructions were read aloud by the experimenter:

> The test which you are about to take is based on the content of introductory psychology. This test is more than a test of knowledge, however, as the problems have been designed to test psychological thinking processes such as application of principles, analysis of relationships, and critical thinking. You are asked to do your best and to earn as high a score as possible since your group's results will be compared with the results of another group of students who will also be taking this test. It should be emphasized, however, that test scores obtained this evening will not be used in any way in determining your course grade in psychology. There is no penalty for guessing. Therefore, be sure to answer all 62 questions. You may begin.

At the end of 40 minutes, subjects were reminded that they could take as much time as desired. After finishing the Criteria Test, each subject filled out a questionnaire containing the study habits and the debilitating anxiety scales (Appendix C). Subjects were allowed to leave after completing the second instrument. Almost all subjects were finished with both tests in the space of an hour and 15 minutes with only a few taking the opportunity to stay longer.

Since the stress condition was structured in the context of final examinations, it was assumed that high evaluative stress would be present in view of the importance of these examinations in determining course grades. Therefore, no special instructions to induce stress were given. However, students were told that the Criteria Test was a department examination so that some of the items might not be familiar. In order to reduce the possibility of coaching prior to the final examination, teaching fellows were given the option of scoring only those items which they felt were related to their own objectives. However, all 60 items were scored for this study.

In order to assess the perceived level of stressfulness in both testing conditions two items were added to the Criteria Test in the low

stress condition. These same items were included on a 5-item "test reactionnaire" which was given to students at the conclusion of the final examination. Students were not required to identify themselves on the "test reactionnaire." The items were:

> How much pressure did you feel you were under while taking this test?
>
> How nervous did you feel while you were taking this test?

Each item was rated on a 5-point scale with one signifying a great deal of pressure/nervousness, and five indicating little, if any, pressure/nervousness. The item means were 3.34 for the evaluative condition, and 4.21 for the nonevaluative condition. Although students reported experiencing only an average amount of pressure/nervousness during the final examination, the difference between item means was significant ($t = 9.24$, $p < .01$, 2 tails) and in the expected direction suggesting that the situational cues had been effective in producing different levels of perceived stress.

A treatments x levels (fixed effects) analysis of variance design was used to analyze the data. Problems were encountered in maintaining equal cell frequencies for the four-way analysis of variance similar to those encountered in the survey study. Therefore, procedures used for determining levels described in Chapter III were also used in categorizing subjects for this analysis. Scores used for breaking points on each variable are given in Tables 32-33, Appendix J. Since sex differences were anticipated, separate analyses were done for men and women.

The research design called for five subjects per cell. But because of chance factors in subject selection, it was necessary to

supply missing data in order to maintain the desired number of replications in each cell. Scores for missing subjects were generated by computing the mean of the deficient cell based on the actual cases and then using this value as the missing score (Lindquist, 1958). Only one case of missing data had to be provided for male subjects. There were four deficient cells for females requiring one additional synthetic score, and one cell requiring two scores.

To enhance the opportunity for anxiety effects to emerge, only subjects from the upper 33 per cent and lower 33 per cent of the anxiety score distribution were used in the analysis.

Results

Independent and dependent variables were intercorrelated prior to carrying out the analysis of variance. These Pearson product-moment correlation coefficients are reported in Tables 15 and 16. Distribution characteristics for the variables can be found in Appendix H, and reliability coefficients in Appendix I.

Although it had been found that the verbal score of the Scholastic Aptitude Test was most highly correlated with course grade in the survey study, it was found that the SAT total score tended to be a better predictor of Criteria Test score (Tables 15 and 16). Therefore, SAT total score was used as the measure of ability in the analysis of variance design.

Complete tables of variance estimates, F-ratios, and significance levels for the four-way analysis of variance are reported in Appendix L.

TABLE 15

INTERCORRELATION OF VARIABLES FOR 101 MEN AND 108 WOMEN IN INTRODUCTORY PSYCHOLOGY, CRITERIA TEST GIVEN UNDER NONEVALUATIVE CONDITIONS
(Fall, 1964)

		DA	SATV	SAT Total	Criteria Test Score
SH	Men	-.231*	-.092	-.073	.006
	Women	-.483*	-.004	.047	.082
DA	Men		-.346**	-.383**	-.355**
	Women		-.226*	-.148	-.132
SATV	Men			.890**	.587**
	Women			.746**	.649**
SAT Total	Men				.635**
	Women				.573**

*p ≤ .05 (2 tailed) **p ≤ .01 (2 tailed)

TABLE 16

INTERCORRELATION OF VARIABLES FOR 96 MEN AND 99 WOMEN IN INTRODUCTORY PSYCHOLOGY, CRITERIA TEST GIVEN UNDER EVALUATIVE CONDITIONS
(Fall, 1964)

		DA	SATV	SAT Total	Criteria Test Score
SH	Men	-.179	-.048	.019	.141
	Women	-.246**	.085	.029	.149
DA	Men		-.278**	-.366**	-.341**
	Women		-.147	-.132	-.136
SATV	Men			.864**	.499**
	Women			.811**	.508**
SAT Total	Men				.566**
	Women				.539**

*p ≤ .05 (2 tailed) **p ≤ .01 (2 tailed)

Main effects and interactions found to be significant at or approaching the 5 per cent level are reported below. As can be seen in Table 17, there was a highly significant main effect for ability for both men and women. High ability subjects earned higher Criteria Test scores than low ability subjects with performance of middle ability subjects falling at an intermediate level.

TABLE 17

MEAN CRITERIA TEST SCORES FOR LEVELS OF ABILITY
(Fall, 1964)

	Low SAT Total	Mid SAT Total	High SAT Total	df	F	P
Men	32.95	37.00	40.65	2,96	22.96	.005
Women	31.10	36.95	42.03	2,96	50.41	.005

Results shown in Table 18 reveal a highly significant main effect for level of evaluative stress for both men and women. As can be observed in the Table, students earned higher Criteria Test scores when constrained by penalties for low performance than in the situation where constraint was absent ($p < .005$).

The results in Table 19 show a weak main effect for level of anxiety for men only. Moreover, there was no anxiety x stress interaction which had been expected on the basis of theory.

Although there was no main effect for study habits, an interesting interaction between study habits and debilitating anxiety occurred.

TABLE 18

MEAN CRITERIA TEST SCORES FOR LEVELS OF EVALUATIVE STRESS
(Fall, 1964)

	Low Evaluative Stress	High Evaluative Stress	df	F	P
Men	34.05	39.68	1,96	36.41	.005
Women	34 63	38.75	1,96	20.03	.005

TABLE 19

MEAN CRITERIA TEST SCORES FOR LEVELS OF DEBILITATING ANXIETY AND STRESS
(Fall, 1964)

		Low Evaluative Stress	High Evaluative Stress	df	F	P
				Main effect		
	High DA	33.43	38.50	1,96	3.67	.05<p<.10
Men				Interaction of anxiety x stress		
	Low DA	34.67	40.87	1,96	.65	N.S.
				Main effect		
	High DA	34.93	37.87	1,96	.65	N.S.
Women				Interaction of anxiety x stress		
	Low DA	34.33	39.63	1,96	1.39	N.S.

This interaction was highly significant for women ($p < .01$), but only approached significance for men ($.05 < p < .10$). Table 20 reveals that students who are low in anxiety and high in study habits earned higher Criteria Test scores than students low in anxiety and low in study habits. Level of study habits did not differentiate groups high in debilitating anxiety.

TABLE 20

MEAN CRITERIA TEST SCORES FOR THE INTERACTION OF STUDY HABITS AND DEBILITATING ANXIETY
(Fall, 1964)

		Low DA	High DA	df	F	P
Men	High SH	39.33	35.83	1,96	3.40	.05<p<.10
	Low SH	36.43	36.10			
Women	High SH	38.77	35.93	1,96	7.17	.01
	Low SH	35.20	36.87			

Discussion

The results obtained from the data analysis proved to be more complex than originally anticipated. For example, it had been predicted that students high in fear of Failure would achieve higher levels of performance under conditions of constraint (imposed by penalties for low achievement) than in conditions where such constraint is absent. In view of the main effect for evaluative stress (Table 18) it would seem

that our hypothesis is supported. However, students low in fear of Failure performed significantly lower under the nonevaluative condition, also. The absence of the anticipated interaction between degree of constraint and fear of Failure (Hypotheses IV and V) suggests that the concept of failure avoidance is inadequate to account for the observed outcomes (Tables 18 and 19).

The fact that both high and low debilitating anxiety subjects performed less well under the nonevaluative condition could be explained by situational factors rather than by the arousal of the avoidance motive. It may well be that the instructions used in the nonevaluative condition led to a low degree of ego-involvement for all subjects irrespective of their anxiety level. College examinations are seldom viewed with euphoric anticipation by students. Since the testing situation was heavily loaded with cues associated with the pressures of evaluation (e.g., the test was given in a regular classroom, the performance measure itself closely resembled a typical achievement test, etc.) the instruction that the test papers "would not be graded" may have been an open invitation for mediocre performance since no penalty would be incurred for nonachievement.

Another factor which may have enhanced the conditions main effect is the point in time at which the nonevaluative condition was administered. As mentioned previously, this testing was done in the eleventh week of the semester. In view of the importance of final examinations in determining course grades, it seems reasonable to assume that the quantity and quality of course related activities increases as

the semester draws to a close. Thus, the observed increase in Criteria Test scores during the evaluative condition could also be due to learning increments acquired during the last four weeks of the 15-week term. Whether the gain in performance is due to the high degree of constraint, to actual increases in learning, or to a combination of the two, is indeterminate at this point. A more precise research design is needed to disentangle the relative contribution of these two factors. In any case, our hypotheses predicting that subjects varying in level of fear of Failure would be differentially affected by the degree of social constraint are not supported by the data.

It was also predicted that students low in debilitating anxiety would excel students with high debilitating anxiety to the extent that spontaneous learning is an attribute of the task. Although the data are suggestive of a main effect for debilitating anxiety (Table 19), the F-ratios failed to reach the 5 per cent level of significance. On the other hand, the interaction obtained between study habits and debilitating anxiety suggests that the relationship between performance of difficult tasks and anxiety is more complex than originally anticipated. As can be seen in Table 20, students high in debilitating anxiety performed poorly whether or not they differed in study habits. However, study habits did differentiate levels of performance for students low in debilitating anxiety as high study habits-low anxiety students tended to outperform all others. This effect could be accounted for in part by the disruptive effect of anxiety interacting with characteristics of the task.

Problems in the Criteria Test are somewhat unusual in that complex problem solving skills are required by many of the items. For example, certain problems require reading of one or more paragraphs of information before a solution can even be attempted. Since students had no opportunity to study these items prior to testing, successful completion would require adeptness at unpracticed problem solving. Thus, Criteria Test performance would require a component of spontaneous learning. In Chapter I it was observed that anxiety is most disruptive to performance during early stages of response acquisition. Therefore it seems reasonable to conclude that anxiety would disrupt performance on the Criteria Test to the extent that new response acquisition is an attribute of the task. The results appear to support this interpretation as high debilitating anxiety students received lowest scores on the test and did not seem to be helped by level of study habits. On the other hand, low anxiety-low study habits students also performed poorly either because they lack the persistence or the instrumental skills necessary to earn high scores. The observed superiority of low anxiety-high study habits students would seem to be a logical consequence of the absence of task-irrelevant responses (as inferred from their low level of debilitating anxiety) and the presence of task-relevant skills (good study habits) which can be engaged to enhance performance on intellectual achievement tasks (Mandler and Sarason, 1952). Thus, debilitating anxiety and study habits seem to be noncognitive components of intellectual performance.

In order to assess the tenability of our assumption that Criteria Test items are complex and therefore susceptible to the effect of anxiety

arousal, Criteria Test data were re-analyzed by using separate sub-scale scores for easy and difficult items. (Complete details for this analysis can be found in Appendix M.) Interaction patterns for debilitating anxiety and study habits for easy items tended to be consistent with those observed for the full scale scores although they failed to reach the 5 per cent level of significance (Tables 21-22). However, results for the difficult items were surprising in that no interaction was observed for males while the interaction for females revealed that low study habits-high debilitating anxiety women unexpectedly performed the best (Tables 21-22). In view of the magnitude of the probability estimates for these interactions, the observed outcomes could well be due to chance. Hence, there appears to be but limited support for our assertion that anxiety and study habits interact with characteristics of the task.

Although assessment of perceived stress indicated that differing levels of pressure and nervousness were experienced under the two conditions, no stress x anxiety interaction was observed contrary to what would be expected on the basis of previous studies (Table 19). However, anticipated effects began to emerge when item difficulty was taken into account, although these effects also failed to reach significance (Table 23). This interaction may have been washed out by the fact that no time limits were imposed under either condition. This being the case, subjects would have had sufficient time to overcome some of the debilitating effects of anxiety arousal.

The main effect observed for ability (Table 15) confirms the potent predictive power of this variable on tasks requiring intellectual

TABLE 21

MEAN SCORES FOR THE INTERACTION OF DEBILITATING ANXIETY AND STUDY HABITS FOR EASY AND DIFFICULT CRITERIA TEST ITEMS, MEN IN PSYCHOLOGY (Fall, 1964)

		Low SH	High SH	df	F	P
Easy Items	High DA	7.70	7.50	1,96	2.17	.10<p<.20
	Low DA	7.70	8.20			
Difficult Items	High DA	4.53	4.73	1,96	.00	N.S.
	Low DA	4.63	4.93			

TABLE 22

MEAN SCORES FOR THE INTERACTION OF DEBILITATING ANXIETY AND STUDY HABITS FOR EASY AND DIFFICULT CRITERIA TEST ITEMS, WOMEN IN PSYCHOLOGY (Fall, 1964)

		Low SH	High SH	df	F	P
Easy Items	High DA	7.96	7.80	1,96	3.46	.05<p<.10
	Low DA	7.37	8.13			
Difficult Items	High DA	4.97	4.33	1,96	2.29	.10<p<.20
	Low DA	4.50	4.73			

TABLE 23

MEAN SCORES FOR THE INTERACTION OF DEBILITATING ANXIETY AND EVALUATIVE STRESS FOR DIFFICULT CRITERIA TEST ITEMS, MEN AND WOMEN IN PSYCHOLOGY (Fall, 1964)

		Low Evaluative Stress	High Evaluative Stress	df	F	P
Men in Psychology	High DA	4.43	4.87	1,96	1.22	N.S.
	Low DA	4.33	5.40			
Women in Psychology	High DA	4.43	4.87	1,96	2.65	.11
	Low DA	3.93	5.30			

aptitude. This was especially true for Criteria Test performance as ability accounted for 22 per cent and 41 per cent of the total variance (men and women respectively). Although the interaction of study habits and debilitating anxiety was statistically significant for women ($p < .01$), this effect accounted for but 3 per cent of the variance.[4] The magnitude of the discrepancy between strengths of association for these variables reinforces the need for including ability level as a control whenever the dependent variable loads heavily on the ability factor. In this way effects due to biased sampling procedures can be avoided.

[4]The strengths of association reported here are based on "estimated omega^2" (Hays, 1963). In order to use this technique, it was assumed that the samples were representative of the populations from which they were drawn.

Summary

The purpose of this study was to investigate relationships between fear of Failure (as inferred from achievement anxiety), degree of constraint, and level of stress. Students in an introductory psychology course were subjected to the following conditions: (a) a noneval-uative, low stress condition; (b) a high stress condition where their performance would actually be graded. The dependent variable was score on the Introductory Psychology Criteria Test.

Statistically significant results were obtained for the following variables.

1. A main effect for ability with high ability being associated with high performance.
2. A main effect for degree of constraint. Students earned higher Criteria Test scores when they knew they were going to be graded. However, subjects in the evaluative condition had the benefit of four weeks of additional exposure to the course content.
3. An interaction between study habits and debilitating anxiety. Highly anxious students achieved low levels of performance irrespective of study habits. Low anxiety-low study habits students performed similarly to highly anxious subjects. Students with high study habits and low anxiety achieved highest levels of performance.

The results were interpreted as failing to support predictions derived from properties of the motive, fear of Failure. The interaction between study habits and anxiety was discussed in terms of the characteristics of the achievement task.

CHAPTER V

AN OVERVIEW

The aim of this study was to examine the long term and short term effects of anxiety arousal in academic achievement situations. Hypotheses, predicting long term effects for anxiety arousal, were based on a social psychological model which assumed that human behavior is determined by the arousal of latent motives in the presence of relevant situational cues. Although it had been predicted that teachers who stressed evaluation would enhance the academic achievement of students high in fear of Failure, results of a survey type study failed to support this anticipated outcome. Teacher behaviors believed to be indicative of evaluative stress were found to be unrelated to academic achievement. Suggestions derived from this disconfirmation of theory are taken up in the section entitled, "Implications for Future Research."

Predictions for the effect of anxiety arousal during short periods of time, e.g., performance on course examinations, were based on theories proposed by Spence (1958) and Mandler and Sarason (1952). The dimension of social constraint was added to the research design to determine whether penalties for nonachievement would enhance or mitigate the effect of anxiety arousal on achievement test performance. Theory would lead us to predict that high evaluative stress should be more

detrimental to the learning and performance of anxious subjects than low evaluative stress. The results, reported in Chapter IV, failed to support this hypothesis as no anxiety by evaluative stress interaction appeared. The fact that no time limits had been imposed during testing could account for the lack of interaction as unlimited time would provide an opportunity to overcome some of the disruptive effects of anxiety arousal. Social constraint did not interact with achievement anxiety although a main effect did occur. Students earned higher achievement scores when they knew they were being graded irrespective of anxiety level.

Interactions for the noncognitive variables of study habits and debilitating anxiety were observed in the experimental study (Chapter IV). This result was thought to be related to characteristics of the achievement task. For example, students high in debilitating anxiety performed poorly irrespective of level of study habits, as did students low in anxiety and low in study habits. Highest levels of performance were observed for subjects with high study habits and low anxiety. In view of the unusual problems included in the performance measure, the Introductory Psychology Criteria Test, debilitating anxiety appeared to disrupt performance to the extent that unpracticed problem solving was an attribute of the task. The low level of performance for low anxiety-low study habits subjects was considered to be due to either lack of persistence or lack of instrumental problem solving skills.

On the basis of these results, it seems reasonable to conclude that no general statements should be made regarding the effect of

anxiety on performance without also specifying characteristics of the task as well as relevant characteristics of the subject. Specifying that a task is simple or complex is not enough. Whether or not anxiety arousal is disruptive depends upon the amount of learning that must be acquired in order to perform the task. And the amount of learning depends on the subject's previous familiarity with the task. Thus, performance on academic achievement tests is determined not only by intellectual ability and knowledge of content, but also by the interaction of noncognitive variables, e.g., debilitating anxiety and study habits, with the nature of the achievement test.

Implications for Future Research

Results presented in Chapter III failed to support our theoretical expectation that teacher produced evaluative stress would arouse and effectively channel the efforts of students high in fear of Failure toward relevant achievement goals. This disconfirmation of theory could be accounted for by a number of factors which would include: (a) the lack of differences in teaching styles, (b) the dependent variable, course grade, and (c) method of assessing teacher behavior.

With regard to the first named category, how much do college teachers really differ in their teaching styles? Results obtained from the survey study suggest that differences among teachers in stressing evaluation are small. It may well be that college teachers differ very little in other classroom behaviors also. Thus, the restricted range of teacher differences would make it exceedingly difficult, if not

impossible, to detect relationships between student achievement and the teacher's behavior.

If marked behavioral differences among college teachers are few, then it would be advantageous to create relevant differences experimentally. The communication of positive reinforcement and feedback are two examples of teacher behavior, relevant to learning, which could be experimentally manipulated. Actual control on the presence (or absence) of these behaviors in the classroom would provide a more meaningful and more rigorous test of whether or not these specific behaviors, and others like them, do have an effect upon student learning.

The fact that the teacher's behavior was unrelated to academic achievement raises a basic question about the nature of our dependent variable, course grade. Is this measure sensitive enough to reflect changes in students which are associated with differences in teaching style? Because we have been unable to find relationships between teacher behavior and academic performance, it does not necessarily follow that students remain unchanged in other important educational areas. Desirable outcomes in addition to assimilation of content would include internalization of instrumental values such as developing tolerance and respect for opposing points of view, ways of dealing with empirically based as well as ideological issues, and realistic self appraisal of one's efforts and goals. Unfortunately, scales for measuring the outcomes suggested above are still very crude. However, progress in understanding how the teacher influences the educational growth of students may well hinge on the successful development of techniques for assessing these particular kinds of educational outcomes.

As suggested earlier, the inconclusiveness of our results also could be due to the method used for assessing teacher behavior. Aside from the fact that untrained observers made but one global judgment of the teacher at the conclusion of the course, it may well be that this measurement technique was not sensitive enough to tap elements of teacher behavior that really do make a difference. More specifically, the teacher behaviors assessed were of a general nature, e.g., the frequency of emphasizing grades, judging the worth of a student's performance in class, etc. It could be that qualitative aspects of these behaviors are as important as their quantity in terms of channeling (or extinguishing) motives which have been aroused. Channeling of motives may be too subtle to be assessed from the teacher's behavior independent of the social interaction in which these behaviors occurred. Even if a teacher is rated high on judging students' performance in class, it would be important to know whether the judging was done disparagingly or with support and positive feedback. For example, an evaluative statement such as, "My 10 year old son could think of a better question than that," would have a different impact on motivation than another evaluative statement such as, "That's a good question because it suggests that there may be a relationship between factors X and Y." Therefore, actual scoring of the classroom social interaction could give us a better notion of how student motives are being aroused as well as determining what the teacher does to channel or to block them once they have been aroused.

Research by Flanders (1965) is suggestive in this respect as he has been able to demonstrate that functional relationships exist between

certain kinds of teacher behavior and classroom outcome measures such as achievement, attitudes toward learning, and general satisfaction.

Unfortunately, interaction process analysis is a costly methodology. Highly trained observers are required; an adequate sampling of a single classroom often necessitates from 8-10 hours of observer time. However, if the classroom motive arousal model is to be adequately tested, measurements at this level of precision should be given careful consideration.

APPENDICES

APPENDIX A

CHARACTERISTICS OF THE STUDY SAMPLES BY COURSE, YEAR, AND SEX

TABLE 24

CLASS STANDING OF MEN AND WOMEN IN PSYCHOLOGY AND ECONOMICS, THE SURVEY STUDY (Fall, 1964)

	Fresh.	Soph.	Juniors	Seniors	Total
Men in Psychology	188	120	52	10	370
Women in Psychology	353	184	37	3	577
Men in Economics	20	252	51	15	338
Women in Economics	8	89	37	15	149

APPENDIX B

INSTRUCTIONS TO ECONOMICS AND PSYCHOLOGY INSTRUCTORS FOR RESEARCH ON EFFECTIVE TEACHING PRE-TEST

To the instructor:

The following instructions are being communicated in order to assist you in administering the pre-test of the Economics-Psychology teaching study.

You are being given one set of test booklets which should be used in both sections, that is, if you teach more than one. There should be enough introductory letters and answer sheets for all students enrolled in your classes. If you should find a lack of materials please call 764-6310 immediately and ask either Amy Konopacki or Robert Stakenas to provide you with the needed number of items.

The importance of providing standardized testing procedures in gathering reliable data in social science research is probably not new to you. The effect of experimenter attitude and its effect on subjects' behavior during data gathering has been well-documented by Rosenthal. Therefore, these instructions are being provided to assist you in establishing as neutral conditions as possible and to be certain that all desired information is obtained.

It is very important that the testing session be carried out in a businesslike way. Although it may be tempting to inject a note of humor by referring to the students as "human guinea pigs" you are asked to refrain from doing so for obvious reasons. Try to maintain a matter-of-fact, businesslike atmosphere. The suggested procedure is to read directions verbatim to the students. Read those portions denoted by quotation marks. In order to avoid confusion due to time pressures, it is suggested that you administer the test at the beginning of the class. If effectively organized, it should take no more than 20 minutes.

There are 3 items that each student should receive.

1. A letter of introduction
2. IBM answer sheet
3. Test booklet

So that you may have maximum freedom to administer instructions and monitor test-taking, it is suggested that you have 3 students pass out the materials for you in the order indicated above.

INTRODUCTORY STATEMENT TO BE READ ALOUD:

"The Economics department and Psychology department are collaborating in some research on college teaching. Therefore, all students in the introductory course are being asked to participate. Take a few moments now to read the introductory note you have just received since it describes the nature of the research and what activities are being asked of the students."

(Allow 1 minute by the clock for reading time. If there are questions asked, refer them to the researcher's name and the Haven Hall office number as he will be glad to attempt to answer any detailed questions students might have.)

NEXT: FILLING IN INFORMATION ON THE IBM ANSWER SHEET.

It is extremely important that the desired information be <u>accurately obtained</u>. Refer to the attached specimen for additional details as you read the instructions. Space reading of instructions so that students have an opportunity to keep up with you as you read.

"Fill in your full name, no pet names please."

"Indicate your class standing in the space marked 'Grade.' Use Fr. (for Freshman), Soph., Jr., and Sr., etc."

"Mark sex information as indicated."

"Date of birth can be left blank."

"Fill in today's date as indicated."

"In the space marked <u>School</u> write Psych 101."

"If you are enrolled in Econ 101 this semester, please indicate this on your blank by writing the three words <u>Econ 101 also</u> in the space marked "City." To avoid confusion in data processing, fill in this answer sheet even if you have already done so in the other course."

"In the space for <u>Instructor</u>, write my name and the section number which is Section _____."

"For Name of Test write <u>SH-TTA</u> (which stands for Study Habits and Test-Taking attitudes)."

"Write in your student number in the space indicated."

"Please do not mark in the test booklets. You are asked to use pencils to mark the answer sheet. I have some pencils for those who need them. If you want to change an answer, be sure that your final choice is clear. Erase inappropriate answers or mark an X through them."

"Note that numbers on the answer sheet are ordered across the page from left to right."

"Your test booklet uses letters of the alphabet while the answer sheet uses numbers. Therefore use space 1 for category A in your booklet, 2 for category B, 3 for category C, etc."
(If any student verbalizes a reluctance to participate, ask him to look at the test items to see if they seem objectionable to him.)

"Remember, no information collected in this research will be used in determining your Course Grade."

"Now read the directions on the test booklet silently to yourself. When you finish reading them, turn the page and begin answering the items. You should be able to complete the test in 12 minutes or less. You may begin."

Be certain that <u>all</u> test booklets and answer sheets are turned in at the conclusion of the testing. Encourage the students to turn in a stamped, self-addressed envelope for feedback on the results of the research as stated in the introductory letter.

If a student is absent and requests to make up the pre-test, refer him to 6622 Haven Hall and we will be glad to administer the make-up test to him there.

After you have administered the test to all of your classes, return all test booklets, marked answer sheets, and unused materials by Campus Mail to:

Robert Stakenas
Psychology Department
6622 Haven Hall

At this point congratulate yourself on a job well-done. Thank you again for your cooperation.

THE LETTER OF INTRODUCTION

Dear Student:

We are asking the students in Psychology 101 and Economics 101 to participate in some research on effective college teaching. We are attempting to examine relationships between certain student characteristics and the teacher's classroom behavior. The research design, therefore, calls for a pre-test and a post-test.

The pre-test you will be taking today is intended to measure student attitudes toward study habits and test-taking. During the post-test, which will take place near the end of the semester, you will be asked to rate your teacher. On the basis of these tests, we hope to find systematic relationships between student characteristics, the instructor's teaching style, and academic performance.

You are encouraged to consult the results for your standing on the student characteristics mentioned above, and how these characteristics interact with differing teaching styles.

It is emphasized that individual results will be kept in confidence and reported only to the individual to whom they belong. If you are interested in obtaining the above-mentioned information, as well as the results of the study, you are encouraged to leave a stamped, self-addressed envelope with Miss Amy Konopacki, 6622 Haven Hall. The results should be available some time during the winter semester. You may keep this page for future reference if you wish.

Thank you for your cooperation. Your participation is greatly appreciated.

Sincerely yours,

Robert Stakenas, Research Assistant
Research Project on
Effective College Teaching

APPENDIX C

MEASURES OF STUDENT CHARACTERISTICS

DIRECTIONS

The purpose of the first 20 items of this survey is to furnish an inventory of study habits and attitudes that are related to the way different students perform with different kinds of teachers. Please answer all questions frankly, since the utility of the survey depends upon the accuracy of the responses of individual students.

Mark your answers on the separate answer sheet. Please make no marks on this booklet. For each of the first 20 statements a five-point scale is provided for indicating whether you rarely, sometimes, frequently, generally, or almost always do or feel as the statement suggests. You are to rate yourself on each statement by marking the space on your answer sheet that represents your answer choice. Thus, for example, you would mark space A on your answer sheet if you rarely follow the procedure described or if you feel that the statement is rarely true for you. In marking your answer, be sure that the number of the statement agrees with the number on the answer sheet. Make sure that your marks are heavy and black. Make no stray marks on the answer sheet and erase completely any mark that you wish to change.

To aid you in answering the terms have been defined on a percentage basis as follows:

A - RARELY means from 0 to 15 percent of the time.
B - SOMETIMES means from 16 to 35 percent of the time.
C - FREQUENTLY means from 36 to 65 percent of the time.
D - GENERALLY means from 66 to 85 percent of the time.
E - ALMOST ALWAYS means from 86 to 100 percent of the time.

Remember, you are asked to rate yourself, not in accordance with what you think you should do or feel, or as you think others might do or feel, but as you yourself are in the habit of doing and feeling. When you cannot answer a statement on the basis of actual experience, mark the statement according to what you would be most likely to do if the situation should arise. Work as rapidly as you can without being careless, and do not spend too much time on any one statement. Please do not omit any of the statements.

Items 21 through 48 of the questionnaire ask about certain of your personal feelings, attitudes, and experiences about course examinations. Some of the questions refer to your past experiences with examinations; when you answer these, think back to your school examinations of the last couple of years. Obviously there are no "right" or "wrong" answers to any of these kinds of questions. They merely offer an opportunity to express feelings and attitudes with regard to a large range of situations. Again, the value of the results of this questionnaire will depend on how frank you are in stating your feelings and attitudes.

Read each statement and set of alternatives carefully. Then select the answer which best describes your own actual feelings or behavior, and blacken in the space on your answer sheet that corresponds to the answer that you have selected.

1. In preparing reports, themes, term papers, etc., I make certain that I clearly understand what is wanted before I begin work.

2. I memorize grammatical rules, definitions of technical terms, formulas, etc., without really understanding them.

3. In taking class notes, I try to copy down the teacher's exact words as closely as possible.

4. I keep all the notes for each subject together, carefully arranging them in some logical order.

5. Difficulty in assembling ideas with order and clearness within a brief amount of time results in my doing poorly on examinations.

6. I keep my place of study businesslike and cleared of unnecessary or distracting items such as pictures, letters, momentoes, etc.

7. I put off writing themes, reports, term papers, etc., until the last minute.

8. I carefully study the figures, graphs, and tables in a reading assignment.

9. My study is done in a random, unplanned manner and is impelled mostly by the demands of approaching classes.

10. I utilize the vacant hours between classes for studying so as to reduce the evening's work.

11. I have difficulty in picking out the important points of a reading assignment--points that are later asked on examinations.

12. When reading a long assignment, I stop periodically and mentally review the main facts and theories that have been presented.

13. I study three or more hours per day outside of class.

14. Before each study period I set up a goal as to how much material I will cover.

15. I can concentrate on a reading assignment for only a short while before the words become a meaningless jumble.

16. I copy the diagrams, drawings, tables, and other illustrations that the instructor puts on the blackboard.

17. I keep my assignments up to date by doing my work regularly from day to day.

18. I prefer to study my lessons alone rather than with others.

19. When preparing for an examination I arrange facts to be learned in some logical order--order of importance, order of presentation in class or textbooks, order of time in history, etc.

20. If time is available, I take a few minutes to check over my answers before turning in my examination paper.

21. Nervousness while taking an exam or test hinders me from doing well.

    a. Always
    b. Often
DA[1] c. Sometimes
    d. Rarely
    e. Never

22. I work most effectively under pressure, as when the task is very important.

    a. Always
    b. Usually
FA[2] c. Sometimes
    d. Hardly ever
    e. Never

---

[1]Debilitating Anxiety (unmarked items are not scored).

[2]Facilitating Anxiety.

23. In a course where I have been doing poorly, my fear of a bad grade cuts down my efficiency.

    a. Never
    b. Hardly ever
DA c. Sometimes
    d. Usually
    e. Always

24. I keep my grades up mainly by doing well on the big exams rather than on the day-to-day course work.

    a. This is true in almost all of my courses.
    b. This is true in quite a few of my courses.
    c. This sometimes happens.
    d. Usually this is not the case.
    e. This is very rarely true.

25. When I am poorly prepared for an exam or test, I get upset, and do less well than even my restricted knowledge should allow.

    a. This never happens to me.
    b. This hardly ever happens to me.
DA c. This sometimes happens to me.
    d. This often happens to me.
    e. This practically always happens to me.

26. The more important the examination, the less well I seem to do.

    a. Always
    b. Usually
DA c. Sometimes
    d. Hardly ever
    e. Never

27. When I feel confident about my ability to do well on a test,

    a. anxiety does not bother me at all.
    b. anxiety bothers me very rarely.
    c. anxiety may still bother me sometimes.
    d. I still often feel somewhat anxious.
    e. I still feel very anxious.

28. Under pressure I am able to organize my thoughts more clearly than usual.

    a. Definitely not true.
    b. This is hardly ever true.
    c. Sometimes.
    d. This is often the case.
    e. This is true most of the time.

29. While I may (or may not) be nervous before taking an exam, once I start, I seem to forget to be nervous.

    a. I always forget.
    b. Usually.
FA c. Sometimes
    d. I often feel some nervousness.
    e. I am always nervous during an exam.

30. During exams or tests, I block on questions to which I know the answers, even though I might remember them as soon as the exam is over.

    a. This always happens to me.
    b. This often happens to me.
DA c. This sometimes happens to me.
    d. This hardly ever happens to me.
    e. I never block on questions to which I know the answers.

31. Nervousness while taking a test helps me to do better.

    a. It never helps.
    b. It usually doesn't help.
FA c. Now and then it helps.
    d. It generally helps me a little.
    e. It often helps.

32. When I start a test, nothing is able to distract me.

    a. This is always true of me.
    b. This is often true of me.
FA c. This is sometimes true of me.
    d. This is hardly ever true of me.
    e. This is not true of me.

33. During an exam or test, I tend to perspire.

    a. Lots
    b. Quite a bit
    c. Some
    d. Very little (unless there is extreme heat or humidity)
    e. Not at all (unless there is extreme heat or humidity)

34. In courses in which the total grade is based mainly on <u>one</u> exam, I seem to do better than other people.

    a. Never
    b. Hardly ever
FA c. Sometimes
    d. Quite often
    e. Almost always

35. I find that my mind goes blank at the beginning of an exam, and it takes me a few minutes before I can function.

    a. I almost always blank out at first.
    b. I usually blank out at first.
DA c. I sometimes blank out at first
    d. I hardly ever blank out at first.
    e. I never blank out at first.

36. I look forward to exams.

    a. Never
    b. Hardly ever
FA c. Sometimes
    d. Usually
    e. Always

37. I am so tired from worrying about an exam, that I find I almost don't care how well I do by the time I start the test.

    a. I never feel this way.
    b. I hardly ever feel this way.
DA c. I sometimes feel this way.
    d. I often feel this way.
    e. I almost always feel this way.

38. Time pressure on an exam causes me to do worse than the rest of the group under similar conditions.

    a. Time pressure always seems to make me do worse on an exam than others.
    b. Time pressure often seems to make me do worse on an exam than others.
DA c. Time pressure sometimes seems to make me do worse on an exam than others.
    d. Time pressure hardly ever seems to make me do worse on an exam than others.
    e. Time pressure never seems to make me do worse on an exam than others.

39. Although "cramming" under pre-examination tension is not effective for most people, I find that if the need arises, I can learn material immediately before an exam, even under considerable pressure, and successfully retain it to use on the exam.

    a. I am always able to use the "crammed" material successfully.
    b. I am usually able to use the "crammed" material successfully.
FA c. I sometimes can use the "crammed" material successfully.
    d. I hardly ever use the "crammed" material successfully.
    e. I am never able to use the "crammed" material successfully.

40. When I find a question on an exam that I did not expect, I get upset, and don't do as well on the question as my preparation should allow me to do.

    a. Always
    b. Usually
    c. Sometimes
    d. Hardly ever
    e. Never

41. I enjoy taking a difficult exam more than an easy one.

    a. Always
    b. Often
FA c. Sometimes
    d. Rarely
    e. Never

42. When I lack confidence in my ability to do well on a test,

    a. I begin to feel very anxious.
    b. I begin to feel slightly anxious.
    c. I sometimes feel mildly anxious.
    d. It hardly ever upsets me.
    e. It doesn't bother me.

43. I find myself reading exam questions without understanding them, and I must go back over them so that they will make sense.

    a. Never
    b. Rarely
DA c. Sometimes
    d. Often
    e. Almost always

44. The more important the exam or test, the better I seem to do.

    a. This is true of me.
    b. This is true of me much of the time.
FA c. This is sometimes true of me.
    d. This is rarely true of me.
    e. This is not true of me.

45. During an exam or test I become conscious of my heart-beat.

    a. Almost always
    b. Frequently
    c. Sometimes
    d. Hardly ever
    e. Never

46. When I don't do well on difficult items at the beginning of an exam, it tends to upset me so that I block on even easy questions later on.

    a. This never happens to me.
    b. This very rarely happens to me.
DA c. This sometimes happens to me.
    d. This frequently happens to me.
    e. This almost always happens to me.

47. Relative to other students, I seem to have more (or less) than the average amount of harmful nervousness about tests and exams.

    a. Much more
    b. A little more
    c. About average
    d. A little less
    e. Much less

48. I feel that if I were not surrounded by competitors in an exam situation,

    a. I would do much better.
    b. I would do a little better.
    c. It wouldn't make much difference.
    d. I wouldn't do quite as well.
    e. I wouldn't do anywhere near as well.

DIRECTIONS

Be sure to use a pencil when filling in the answer sheets. First of all, fill in your full name, no nick names, please.

Indicate your class standing in the space marked "Grade." Use Fr. (for Freshman), Soph., Jr., or Sr.

Mark sex information as indicated.

Date of birth can be left blank.

Fill in today's date which is _____ as indicated.

In the space marked School write in the course number. This should be either Psychology 101, 102, or 191.

In the space for Instructor, write in his or her name and section number.

For name of Test write C.T., which stands for Criteria Test.

Write in your student number in the vertical column provided. Blacken in the spaces corresponding to the numbers.

Please do not mark in the test booklets. Note that numbers on the answer sheets are ordered across the page from left to right. If you want to change an answer, be sure that your final choice is clear. Erase inappropriate answers or mark an X through them.

The test which you are about to take is based on the content of introductory psychology. This test is more than a test of knowledge, however, as the problems have been designed to test psychological thinking processes such as application of principles, analysis of relationships, and critical thinking. You are asked to do your best and to earn as high a score as possible since your group's results will be compared with the results of another group of students who will also be taking this test. It should be emphasized, however, that test scores obtained this evening will not be used in any way in determining your course grade in psychology. There is no penalty for guessing. Therefore, be sure to answer all 62 questions. You may begin.

INTRODUCTORY PSYCHOLOGY CRITERIA TEST
SAMPLE ITEMS

Questions A through D concern the following:

An experiment was designed to determine whether noise will affect the amount of food rats will eat in a given period. Two groups of twenty male rats, matched according to age, weight, and health, were chosen. For four weeks one group of rats was subjected to a loud noise while feeding and the other group was fed under regular conditions. The amount of food given each rat was carefully weighed before each feeding and the amount uneaten was weighed after each feeding.

Pick the correct answer for each question from the following key:

1. The amount of time the rats took to eat.
2. The loud noise.
3. The amount of food eaten.
4. The emotionality of the rats.
5. The weight of the rats.

A. The independent variable is:

B. The dependent variable is:

C. An uncontrolled variable of possible importance is:

D. A controlled variable is:

E. Indicate which one assumption was implicitly made by the investigator in arriving at his conclusion. (In other words, what assumption would one have to make in order to arrive at the conclusion which the author reached?)

A mental patient was brought to the hospital with symptoms of hallucinations. On the basis of this alone, he was diagnosed schizophrenic by the examining intern:

<u>Assumptions</u>:

1. All schizophrenic show hallucinations.
2. Most people who show hallucinations are schizophrenic.
3. Schizophrenics often show hallucinations.
4. It is necessary to know something about a patient's history in addition to his present symptoms in order to arrive at a diagnosis.

61. How much pressure did you feel you were under while taking this test?

    1. A great deal of pressure
    2. Quite a bit of pressure
    3. About an average amount of pressure
    4. Some pressure
    5. Little, if any pressure

62. How nervous did you feel while you were taking this test?

    1. Very nervous
    2. Quite nervous
    3. An average amount of nervousness
    4. Some nervousness
    5. Little if any nervousness.

There is one more test left to be taken. It will be given to you shortly. In the time remaining, you may wish to check over your answers on this test. So that you can be assured of doing your best, remember that you may take as much time as you wish.

TEST REACTIONNAIRE

Because of the nature of this test, I would like to know how you felt about taking it. Therefore, please circle the one alternative that reflects your feelings most accurately.

1. In general, how difficult was this exam:
    1. Very difficult
    2. Difficult
    3. Of average difficulty
    4. Easy
    5. Very easy

2. How much time pressure did you feel?
    1. A great deal
    2. Some
    3. Very little

3. How fair do you feel this exam is for this course?
    1. Very fair
    2. Fair
    3. So-so
    4. Unfair
    5. Very unfair

4. How much pressure did you feel you were under while taking this exam?
    1. A great deal of pressure
    2. Quite a bit of pressure
    3. About an average amount of pressure
    4. Some pressure
    5. Little if any pressure

5. How nervous did you feel while you were taking this exam?
    1. Very nervous
    2. Quite nervous
    3. An average amount of nervousness
    4. Some nervousness
    5. Little if any nervousness

APPENDIX D

EVALUATIVE STRESS ITEMS AND THE SUBJECTIVE REPORT QUESTIONNAIRE

Research Project on Effective
College Teaching

Student Rating Form

As you know, this semester there is a large-scale research effort devoted to understanding student-teacher interaction in the classroom. You may recall having taken a pre-test on Study Habits and Test Taking Attitudes some weeks ago. We would like to consider your ratings in relation to the information obtained from the pre-test. To do this, students will need to identify themselves. Therefore, please indicate your name, student I.D. number, etc., in the spaces provided at the bottom of this page.

Turn in your completed form to the student proctor who has been assigned to collect and deliver these forms directly to the research project office, 6622 Haven Hall. Instructors will have the opportunity to study these ratings only after this face sheet has been removed to preserve your anonymity, and all final grades have been turned in to the registrar's office.

- -

(Identifying information must be detached from this form before it can be released for inspection.)

Course_____

Instructor's Name_____

Section Number_____

Your Student I.D. No._ _ _ _ _ _ _

Your Name_____

Directions: Circle the one alternative which reflects your feelings most accurately. Be sure to write comments where requested.

1. How often did the instructor warn the class to expect unannounced quizzes?

    1. Very often
    2. Often
    3. Occasionally
    4. Seldom
    5. Almost never

2. Unannounced quizzes were given in this class

    1. Four or more times
    2. Three times
    3. Two times
    4. One time
    5. At no time

3. The teacher strongly emphasized grades.

    1. Very often
    2. Quite often
    3. Sometimes
    4. Hardly ever
    5. Almost never

4. Most students in this class competed for grades.

    1. Hardly at all
    2. Seldom
    3. Occasionally
    4. Often
    5. Most of the time

5. The teacher made it clear that he was continually judging the worth of a student's performance in class.

    1. This almost always occurred
    2. This often occurred
    3. This occasionally occurred
    4. This seldom occurred
    5. This almost never occurred

6. If you had your choice of activities to be graded on, which one would you most prefer? least prefer?

Most prefer	Least prefer
1. Essay test in class	1. Essay test in class
2. Multiple-choice test	2. Multiple-choice test
3. Take-home essay test	3. Take-home essay test
4. Term paper	4. Term paper

In a sentence or two, indicate why you feel this way.

7. In general, I found myself studying

1. very little in this course
2. less than an average amount
3. an average amount in this course
4. more than an average amount
5. very hard in this course

What factors do you think account for the amount of studying you did in this course? (Write a sentence or two)

8. In preparing for exams and quizzes for this course, I studied

1. More than for my other courses
2. About as much as for my other courses
3. Less than for my other courses

9. I found this course to be

1. difficult
2. somewhat on the difficult side
3. of average difficulty
4. on the easy side
5. very easy

What are some of the factors that seem to make you feel this way?

10. As far as reading assignments for this course were concerned I found myself:

    1. trying to catch up with them at the last moment
    2. barely keeping up with them
    3. keeping up but with little time for review
    4. keeping up with some time for review
    5. keeping up with them on a regular basis with plenty of time for review

11. By the way the instructor ran the class, he made me feel

    1. very anxious and uncomfortable in class
    2. quite anxious and uncomfortable
    3. an average amount of anxiousness
    4. quite comfortable
    5. very comfortable in class

    In your own words, what did he do that made you feel this way?

12. The marking in this course was generally

    1. too lenient
    2. on the lenient side
    3. about right
    4. on the hard side
    5. too hard

13. The instructor's marking in this course was

    1. very fair and just
    2. fair enough
    3. arbitrary and unfair on at least one occasion
    4. arbitrary and unfair to students on several occasions

14. In this class, I felt under

    1. little, if any pressure
    2. some pressure
    3. an average amount of pressure
    4. more than an average amount of pressure
    5. a great deal of pressure

15. Would you want to take another course from this instructor?

    1. I certainly would
    2. I would not mind
    3. I would prefer not to
    4. I would refuse to

    Why do you feel this way?

16. The tests in this course were

    1. too easy
    2. on the easy side
    3. about average difficulty
    4. on the difficult side
    5. too hard

17. When this teacher came to class, in general he appeared to be

    1. poorly prepared
    2. quite poorly prepared
    3. somewhat prepared
    4. fairly well prepared
    5. well prepared

    How did his level of preparedness make you feel?

18. The tests in this course were

    1. too many
    2. about right in number
    3. too few

19. To what extent do you believe that good grades in all your college courses will help you do well in your chosen career?

    1. be a great help
    2. be of some help
    3. be of very little help
    4. practically irrelevant

20. To what extent do you think your college grades will be used as a basis for your selection into a career? (Select one)

    1. College grades will be very important to my selection
    2. College grades will be important to my selection
    3. College grades will matter to some extent
    4. College grades will be almost immaterial to my selection

21. To what extent do you believe getting a good grade in the introductory psychology course will help you to do well in your chosen career?

    1. be a great help
    2. be of some help
    3. be of very little help
    4. practically irrelevant

APPENDIX E

FREQUENCY OF EVALUATION QUESTIONNAIRE

Instructor: _____

Section No: _____, Course: _____

The purpose of this questionnaire is to determine the kind and frequency of activities which were used by you in evaluating your students. The need for this information arises from the goal of this study which is to examine relationships between frequency of evaluative cues and level of fear of Failure on the part of the student. It seems clear that tests and other graded assignments, e.g., term papers, are very salient evaluative cues for students. Therefore, the frequency of these activities may be a very important determinant of student motivation to achieve academically.

This questionnaire requires only 4-5 minutes to complete. Simply circle the alternative which seems most appropriate for your class. If no alternative seems appropriate, feel free to write in comments which are more descriptive of what you did in evaluating your students.

In the blank spaces record the approximate percentage that the designated activity counted in the final grade.

If you taught more than one section but used the same evaluational procedure in both sections, you need fill out only one form.

---

1. Does your department require departmental hour exams for all sections in the introductory course?

    1. Yes
    2. No

2. How many hour exams were there in this class?

My own hour exams	Departmental hour exams
1. None	1. None
2. One	2. One
3. Two	3. Two
4. Three	4. Three
5. Four or more	5. Four or more

3. Scores on all hour exams counted approximately _____ per cent of the final grade.

4. In addition to hour exams, how many quizzes did you require?

1. None
2. One
3. Two
4. Three
5. Four or more

5. Scores on quizzes counted approximately _____ per cent of the final grade.

6. Was a term paper(s) required?

1. Yes
2. No
3. Optional. Students could choose to write one.

7. Score on the term paper(s) counted approximately _____ per cent of the final grade.

8. How often were students required to hand in written assignments other than term papers?

Assignments handed in but not graded	Assignments handed in and graded
1. None required	1. None required
2. About once a month	2. About once a month
3. About twice a month	3. About twice a month
4. Once a week	4. Once a week
5. More often than once a week	5. More often than once a week

9. Hand-in assignments which were graded counted approx. _____ per cent of the final grade.

10. Does your department require a departmental final exam for all sections in the introductory course?

    1. Yes
    2. No

11. Score on the final exam counted approximately _____ per cent of the final grade.

12. How many sections of the introductory course did you teach?

    1. One section 2. Two sections

13. If you taught 2 sections, did you use the same evaluational procedure in both?

    1. Yes
    2. No

14. Please return this completed form to:

    Bob Stakenas
    Psychology Department
    6620 Haven Hall
    Via Campus Mail

Thank you again for your cooperation.

Any comments you might have:

APPENDIX F

CONVERSION TABLE BASED ON NORMS FOR
UNIVERSITY OF MICHIGAN FRESHMAN
(1955)

U/M	ACE[a] (1949)			SAT[b]		
PR	Total	Ling.	Quant.	Total	Verb.	Math.
99	168+	109+	69+	139+	693+	739+
98	164-67	107-08	64-8	135-38	684-89	730-35
97	162-63	105-06	63	133-34	670-79	712-21
96	161	103-04		131-32	661-66	702
95	160	100-02	62	129-30	652-57	693
94	157-59	98-99	61	128	647	684
93	155-56	97	60	126-27	638-43	674
92	154	96	59	125	634	665
91	150-53	95	58	124	625-29	
90	149		57	123	620	656
89	147-48	94		122	615	
88	146	93	56	121	611	646
87	145	92		120	606	637
86	144	91	55	119	602	
85	143			118	597	628
84	142	90	54		593	
83	141			117	588	
82	140	89		116	584	618
81	139	88	53		579	609
80				115	574	
79	138	87		114	570	600
78	137	86	52			

[a]Raw Scores.

[b]Scaled Scores.

U/M	ACE (1949)			SAT		
PR	Total	Ling.	Quant.	Total	Verb.	Math.
77	136	85		113		
76	135				565	590
75	134	84	51	112		
74	133			111		581
73	132	83			556	
72	131	82	50	110	552	572
71	130	81				
70				109	547	
69					542	562
68	129	80		108		
67	128				538	
66				107	533	553
65	127	79	49			
64		78		106	529	
63						544
62	126	77		105	524	
61	125					
60	124	76	48	104	520	534
59	123				515	
58				103	510	525
57	122		47			

U/M	ACE (1949)			SAT		
PR	Total	Ling.	Quant.	Total	Verb.	Math.
56		75		102	506	
55	121	74				516
54		73			501	
53	120			101	497	
52	119		46			506
51		72		100	492	
50						
49	118			99	488	497
48					483	
47				98		
46	117		45		478	488
45				97	474	
44	116	70				
43	115			96	469	479
42	114	69		95	465	
41	114	69		95	465	
40			44			469
39	113				460	
38				94		
37	112	68			456	460
36				93		
35					451	

U/M	ACE (1949)			SAT		
PR	Total	Ling.	Quant.	Total	Verb.	Math.
34	111	67	43	92	446	451
33		66				
32	110	65		91	442	
31	109					441
30				90	437	
29	108	64	42		433	
28	107			89		432
27					428	
26	106	63		88		
25		62	41		424	423
24	105			87		
23					419	
22	104			86		413
21	103	61	40		415	
20	102			85	410	
19			39			404
18	101			84	405	
17		60	38	83	401	
16	100				396	
15		59	37	82	392	395
14				81		

U/M	ACE (1949)			SAT		
PR	Total	Ling.	Quant.	Total	Verb.	Math.
13	99	58	36		387	
12	98	57		80	383	385
11	97				378	
10		56	35	79	373	376
9	96			78	369	
8	94-95	55		77	364	367
7	92-93	54	34	76	360	357
6	91	53	33	75	355	
5	90	52		74	351	348
4	86-89	50-51	32	73	338-46	
3	83-85	48-49	30-31	71-72	329-37	339
2	82	46-47	29	69-70	315-28	329
1	72-81	43-45	21-28	66-68	301-14	320
0	00-71	00-42	00-20	40-65	200-300	200-319

APPENDIX G

CODING CATEGORIES FOR OPEN-ENDED RESPONSES ON THE SUBJECTIVE REPORT QUESTIONNAIRE

Codes for Item (1 in Economics; 7 in the Psychology Form)

Free response given to the question: "What factor accounts for the amount of studying that you did in this course?"

Area	Code	Description
Teacher	0[a]	Teacher willing to help; a good teacher; interesting teacher, etc.
	1	Teacher was demanding; assigned lots of readings, many papers, assignments, etc.
	2	Teacher was <u>not</u> demanding; class discussion and lecture were sufficient to keep up, etc.
Content	3	Concepts in the course were <u>difficult</u> to understand, etc.
	4	Concepts in the course were <u>easy</u>; common-sense; had a previous course in it, etc.
Student	5	I <u>liked</u> the subject; it was interesting; stimulating, etc.
	6	I <u>didn't</u> like the subject; I was <u>not</u> interested in it; the course was <u>not</u> very stimulating.
Misc.	7	Irrelevant responses; I wanted to get a good grade, etc.
	8	No response given.

[a]There is no scale implied by these numbers. Each category is considered discrete.

Decision rules: Whenever there is more than one response, score the first one appearing that is relevant.

Check for a rational relationship between amount of work put in and the reason given (assignments which result in an average amount of studying, code as 7).

Codes for Item (3 in Economics; 9 in the Psychology Form)

Free response given to the question: "What are some of the factors that seem to make you feel . . . (that the course was easy/difficult)?"

Area	Code	Description
Teacher	0[a]	The teacher explained things clearly; he made the course interesting, etc.
	1	Teacher was demanding; he required many readings; grading was hard, etc.
	2	Teacher was not demanding; not much work expected; an easy grader; tests were easy, etc.
Content	3	Concepts in the course were difficult to understand; there were many new concepts; concepts were ambiguous, etc.
	4	Concepts were easy; only required memorization; concepts were common-sense, etc.
Student	5	I liked the subject; it was interesting; stimulating, etc.
	6	I didn't like the subject; I was apathetic, etc.
Misc.	7	Irrelevant responses; I did poorly on tests, etc.
	8	No response given.

[a]There is no scale implied by these numbers. Each category is considered discrete.

Decision rules: Whenever there is more than one response, score the first one appearing that is relevant.

Check for a rational relationship between amount of work put in and the reason given (assignments which result in an average amount of studying, code as 7).

If item 9 is marked "easy" and called a memory course, code as a 2.

Codes for Item (5 in Economics; 11 in the Psychology Form)

Free response given to the question: "What are some of the factors . . . (that made you feel anxious/comfortable)?"

Area	Code	Description
Presentation	0[a]	Clear; knew what he was doing; taught well, etc.
	1	Poor; bad teacher, etc.
Interpersonal Relations	2	Interested in student; humorous; informal; understanding; not-threatening; friendly; permissive; never belittled students; good eye contact, etc.
	3	Critical; looked down on student; threatening; conflicted with students attitudes, etc.
Teacher Expectations	4	Was demanding; emphasized grades, etc.
	5	Undemanding, did not expect much from student, etc.
Misc.	7	Irrelevant
	8	No response

[a]There is no scale implied by these numbers. Each category is considered discrete.

Decision rules: Whenever there is more than one response, score the first one appearing that is relevant.

Check for a rational relationship between level of comfort/anxiety and the reason given.

APPENDIX H

DISTRIBUTION CHARACTERISTICS OF VARIABLES FOR THE SURVEY AND EXPERIMENTAL STUDIES

TABLE 25

MEANS AND STANDARD DEVIATIONS FOR MEASURES OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND GRADES FOR 370 MEN AND 577 WOMEN IN INTRODUCTORY PSYCHOLOGY, THE SURVEY STUDY (Fall, 1964)

Variable		Mean	Standard Deviation
SATV	Men	543.51	91.28
	Women	551.47	77.85
SH	Men	17.45	5.79
	Women	20.48	5.75
DA	Men	27.09	5.59
	Women	27.88	5.75
Grades	Men	2.77	.91
	Women	2.79	.85

TABLE 26

MEANS AND STANDARD DEVIATIONS FOR MEASURES OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND GRADES FOR 388 MEN AND 149 WOMEN IN INTRODUCTORY ECONOMICS, THE SURVEY STUDY
(Fall, 1964)

Variable		Mean	Standard Deviation
SATV	Men	562.96	90.48
	Women	572.21	91.00
SH	Men	18.20	5.64
	Women	20.99	5.77
DA	Men	26.97	5.50
	Women	27.04	5.62
Grades	Men	2.67	.82
	Women	2.75	.77

TABLE 27

MEANS AND STANDARD DEVIATIONS FOR MEASURES OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND CRITERIA TEST SCORES FOR MEN, THE EXPERIMENTAL STUDY
(Fall, 1964)

Variable	Condition	N	Mean	Standard Deviation
SAT Total	Evaluative	96	1139.00	151.30
	Non-Eval.	101	1172.10	169.30
SH	Evaluative	96	16.76	5.83
	Non-Eval.	101	18.15	5.67
DA	Evaluative	96	28.04	4.90
	Non-Eval.	101	25.88	5.06
Criteria Test Score	Evaluative	96	39.14	6.05
	Non-Eval.	101	35.02	7.02

TABLE 28

MEANS AND STANDARD DEVIATIONS FOR MEASURES OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND CRITERIA TEST SCORES FOR WOMEN, THE EXPERIMENTAL STUDY
(Fall, 1964)

Variable	Condition	N	Mean	Standard Deviation
SAT Total	Evaluative	99	1111.60	133.80
	Non-Eval.	108	1124.80	165.10
SH	Evaluative	99	21.16	5.12
	Non-Eval.	108	19.32	6.47
DA	Evaluative	99	27.50	5.15
	Non-Eval.	108	27.88	5.72
Criteria Test Score	Evaluative	99	39.33	6.22
	Non-Eval.	108	34.28	6.59

APPENDIX I

RELIABILITY COEFFICIENTS FOR THE STUDENT MEASURES

TABLE 29

RELIABILITY COEFFICIENTS FOR ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND THE CRITERIA TEST

Variable	Reliability	Method
SATV[a]	.90	Internal Consistency
SATM[a]	.88	Internal Consistency
SH[b]	.72	Test-retest
DA[b]	.79	Test-retest
Criteria Test[c]	.69	Parallel Forms

[a]Reported in College Board Score Reports (Princeton: College Entrance Examination Board, 1964).

[b]These coefficients were derived from data obtained for this study. Time between the test and retest varied from two to six weeks (N = 159).

[c]Reported by Milholland (1964).

APPENDIX J

SCALE SCORES FOR LEVELS OF THE TREATMENTS BY LEVELS DESIGN

TABLE 30

SCALE SCORES FOR LEVELS OF ABILITY, DEBILITATING ANXIETY, AND STUDY HABITS FOR MEN IN PSYCHOLOGY AND ECONOMICS, THE SURVEY STUDY (Fall, 1964)

SATV Level	Score Range	DA Level	Score Range	SH Level	Score Range
High SATV	590-800	High DA	28-37	High SH	20-26
				Low SH	07-19
		Low DA	12-22	High SH	20-26
				Low SH	07-19
Mid SATV	510-589	High DA	28-37	High SH	20-33
				Low SH	06-19
		Low DA	15-23	High SH	20-33
				Low SH	06-19
Low SATV	200-509	High DA	32-40	High SH	19-33
				Low SH	11-18
		Low DA	21-27	High SH	19-33
				Low SH	11-18

TABLE 31

SCALE SCORES FOR LEVELS OF ABILITY, DEBILITATING ANXIETY, AND STUDY HABITS FOR WOMEN IN PSYCHOLOGY, THE SURVEY STUDY (Fall, 1964)

SATV Level	Score Range	DA Level	Score Range	SH Level	Score Range
High SATV	590-800	High DA	29-43	High SH	19-31
				Low SH	05-18
		Low DA	11-23	High SH	20-26
				Low SH	07-19
Mid SATV	520-589	High DA	31-41	High SH	21-33
				Low SH	11-20
		Low DA	15-24	High SH	21-33
				Low SH	11-20
Low SATV	200-519	High DA	33-40	High SH	20-31
				Low SH	09-19
		Low DA	17-27	High SH	20-31
				Low SH	09-19

TABLE 32

SCALE SCORES FOR LEVELS OF ABILITY, DEBILITATING ANXIETY, AND STUDY HABITS FOR MEN IN PSYCHOLOGY, AND EXPERIMENTAL STUDY
(Fall, 1964)

SAT Total Level	Score Range	DA Level	Score Range	SH Level	Score Range
		High DA	27-34	High SH	15-32
				Low SH	06-14
High SAT Total	1240-1530				
		Low DA	12-24	High SH	15-32
				Low SH	06-14
		High DA	28-44	High SH	19-29
				Low SH	08-18
Mid SAT Total	1110-1239				
		Low DA	19-25	High SH	19-27
				Low SH	11-18
		High DA	31-37	High SH	19-27
				Low SH	11-18
Low SAT Total	660-1109				
		Low DA	18-26	High SH	19-27
				Low SH	11-18

TABLE 33

SCALE SCORES FOR LEVELS OF ABILITY, DEBILITATING ANXIETY, AND STUDY HABITS FOR WOMEN IN PSYCHOLOGY, THE EXPERIMENTAL STUDY (Fall, 1964)

SAT Total Level	Score Range	DA Level	Score Range	SH Level	Score Range
		High DA	29-43	High SH	18-33
				Low SH	06-17
High SAT Total	1180-1450				
		Low DA	14-25	High SH	18-33
				Low SH	06-17
		High DA	29-40	High SH	22-33
				Low SH	10-21
Mid SAT Total	1060-1179				
		Low DA	18-24	High SH	22-33
				Low SH	10-21
		High DA	31-41	High SH	18-28
				Low SH	05-17
Low SAT Total	780-1059				
		Low DA	16-26	High SH	18-28
				Low SH	05-17

APPENDIX K

THE COMPLETE ANALYSIS OF VARIANCE FOR THE SURVEY STUDY

TABLE 34

ANALYSIS OF VARIANCE FOR ITEM THREE[a]
MEN IN PSYCHOLOGY
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SATV (A)	1874.47	2	937.23	9.31	.005
SH (B)	80.03	1	80.03	--	--
DA (C)	288.30	1	288.30	2.86	.05<p<.10
ES3 (D)	50.70	1	50.70	--	--
A x B	205.27	2	102.63	--	--
A x C	194.60	2	97.30	--	--
A x D	67.40	2	33.70	--	--
B x C	4.03	1	4.03	--	--
B x D	.83	1	.83	--	--
C x D	4.03	1	4.03	--	--
A x B x C	24.07	2	12.03	--	--
A x B x D	195.47	2	97.73	--	--
A x C x D	125.07	2	62.53	--	--
B x C x D	140.83	1	140.83	--	--
A x B x C x D	.47	2	.23	--	--
Within Cells	9662.80	96	100.65		
Total	12918.37	119			

[a]The teacher strongly emphasized grades.

TABLE 35

ANALYSIS OF VARIANCE FOR ITEM FOUR[a]
MEN IN PSYCHOLOGY
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SATV (A)	1032.27	2	516.13	5.47	.01
SH (B)	202.80	1	202.80	--	--
DA (C)	790.53	1	790.53	8.38	.005
ES4 (D)	4.80	1	4.80	--	--
A x B	31.20	2	15.60	--	--
A x C	460.87	2	230.43	--	--
A x D	99.80	2	49.90	--	--
B x C	172.80	1	172.80	--	--
B x D	360.53	1	360.53	3.82	N.S.*
C x D	22.53	1	22.53	--	--
A x B x C	108.20	2	54.10	--	--
A x B x D	108.07	2	54.03	--	--
A x C x D	2.87	2	1.43	--	--
B x C x D	58.80	1	58.80	--	--
A x B x C x D	281.40	2	140.70	--	--
Within Cells	9056.40	96	94.34		
Total	12793.87	119			

[a]Most students in class competed for grades.

*Although an F-ratio of 3.82 (1,96 df) has an associated probability level of .06, this effect was called nonsignificant. In view of the large number of nonindependent tests carried out, and lack of theoretical expectancies, it was concluded that this interaction was due to errors generated by sampling rather than to any underlying psychological phenomena.

TABLE 36

ANALYSIS OF VARIANCE FOR ITEM FIVE[a]
MEN IN PSYCHOLOGY
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SATV (A)	670.05	2	335.02	3.59	.05
SH (B)	616.53	1	616.53	6.60	.025
DA (C)	381.63	1	381.63	4.09	.05
ES5 (D)	7.50	1	7.50	--	--
A x B	143.12	2	71.56	--	--
A x C	91.02	2	45.51	--	--
A x D	66.15	2	33.08	--	--
B x C	16.13	1	16.13	--	--
B x D	53.33	1	53.33	--	--
C x D	24.30	1	24.30	--	--
A x B x C	218.82	2	109.41	--	--
A x B x D	52.02	2	26.01	--	--
A x C x D	132.65	2	66.32	--	--
B x C x D	213.33	1	213.33	--	--
A x B x C x D	77.12	2	38.56	--	--
Within Cells	8961.60	96	93.35		
Total	11725.30	119			

[a]The teacher made it clear that he was continually judging the worth of a student's performance in class.

TABLE 37

ANALYSIS OF VARIANCE FOR FREQUENCY OF TEACHER ASSIGNED TESTS
MEN IN PSYCHOLOGY
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SATV (A)	1835.82	2	917.91	10.10	.005
SH (B)	374.53	1	374.53	4.12	.05
DA (C)	440.83	1	440.83	4.85	.05
TESTS (D)	.00	1	.00	--	--
A x B	54.72	2	27.36	--	--
A x C	654.12	2	327.06	3.60	N.S.*
A x D	7.85	2	3.92	--	--
B x C	22.53	1	22.53	--	--
B x D	4.03	1	4.03	--	--
C x D	224.13	1	224.13	--	--
A x B x C	82.02	2	41.01	--	--
A x B x D	116.22	2	58.11	--	--
A x C x D	126.02	2	63.01	--	--
B x C x D	28.03	1	28.03	--	--
A x B x C x D	186.32	2	93.16	--	--
Within Cells	8726.80	96	90.90		
Total	12883.97	119			

*Although an F-ratio of 3.60 (2,96 df) has an associated probability level of .05, this effect was called nonsignificant. In view of the large number of nonindependent tests carried out, and lack of theoretical expectancies, it was concluded that this interaction effect was due to errors generated by sampling rather than to any underlying psychological phenomena.

TABLE 38

ANALYSIS OF VARIANCE FOR ITEM THREE[a]
MEN IN ECONOMICS
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SATV (A)	17.45	2	8.72	11.18	.005
SH (B)	3.67	1	3.67	4.71	.05
DA (C)	.01	1	.01	--	--
ES3 (D)	.07	1	.07	--	--
A x B	.65	2	.33	--	--
A x C	.12	2	.06	--	--
A x D	1.85	2	.93	--	--
B x C	.01	1	.01	--	--
B x D	2.41	1	2.41	--	--
C x D	.41	1	.41	--	--
A x B x C	2.92	2	1.46	--	--
A x B x D	.32	2	.16	--	--
A x C x D	2.12	2	1.06	--	--
B x C x D	.41	1	.41	--	--
A x B x C x D	.32	2	.16	--	--
Within Cells	75.20	96	.78		
Total	107.93	119			

[a]The teacher strongly emphasized grades.

TABLE 39

ANALYSIS OF VARIANCE FOR ITEM FOUR[a]
MEN IN ECONOMICS
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SATV (A)	16.02	2	8.01	10.01	.005
SH (B)	1.63	1	1.63	--	--
DA (C)	2.70	1	2.70	3.38	.05<p<.10
ES4 (D)	.83	1	.83	--	--
A x B	2.72	2	1.36	--	--
A x C	.05	2	.03	--	--
A x D	.22	2	.11	--	--
B x C	.13	1	.13	--	--
B x D	.53	1	.53	--	--
C x D	.53	1	.53	--	--
A x B x C	.22	2	.11	--	--
A x B x D	.52	2	.26	--	--
A x C x D	.12	2	.06	--	--
B x C x D	.83	1	.83	--	--
A x B x C x D	2.02	2	1.01	--	--
Within Cells	76.40	96	.80		
Total	105.47	119			

[a]Most students in class competed for grades.

TABLE 40

ANALYSIS OF VARIANCE FOR ITEM FIVE[a]
MEN IN ECONOMICS
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SATV (A)	16.65	2	8.32	10.67	.005
SH (B)	5.21	1	5.21	6.68	.025
DA (C)	.41	1	.41	--	--
ES5 (D)	.21	1	.21	--	--
A x B	4.72	2	2.36	3.03	.05<p<.10
A x C	.62	2	.31	--	--
A x D	.02	2	.01	--	--
B x C	.67	1	.67	--	--
B x D	.07	1	.07	--	--
C x D	.21	1	.21	--	--
A x B x C	.35	2	.18	--	--
A x B x D	.95	2	.47	--	--
A x C x D	1.12	2	.56	--	--
B x C x D	.01	1	.01	--	--
A x B x C x D	1.32	2	.66	--	--
Within Cells	74.80	96	.78		
Total	107.33	119			

[a]The teacher made it clear that he was continually judging the worth of a student's performance in class.

TABLE 41

ANALYSIS OF VARIANCE FOR FREQUENCY OF TEACHER ASSIGNED TESTS
MEN IN ECONOMICS
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SATV (A)	21.12	2	10.56	13.37	.005
SH (B)	3.01	1	3.01	3.81	.05<p<.10
DA (C)	.21	1	.21	--	--
TESTS (D)	.21	1	.21	--	--
A x B	.12	2	.06	--	--
A x C	2.22	2	1.11	--	--
A x D	1.12	2	.56	--	--
B x C	.01	1	.01	--	--
B x D	1.01	1	1.01	--	--
C x D	.67	1	.67	--	--
A x B x C	1.02	2	.51	--	--
A x B x D	4.12	2	2.06	--	--
A x C x D	.35	2	.18	--	--
B x C x D	.67	1	.67	--	--
A x B x C x D	.35	2	.17	--	--
Within Cells	75.60	96	.79		
Total	111.79	119			

TABLE 42

ANALYSIS OF VARIANCE FOR ITEM THREE[a]
WOMEN IN PSYCHOLOGY
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SATV (A)	422.47	2	211.23	2.15	N.S.
SH (B)	151.87	1	151.87	--	--
DA (C)	7.01	1	7.01	--	--
ES3 (D)	1.01	1	1.01	--	--
A x B	54.60	2	27.30	--	--
A x C	64.07	2	32.03	--	--
A x D	26.47	2	13.23	--	--
B x C	226.87	1	226.87	2.29	N.S.
B x D	37.41	1	37.41	--	--
C x D	170.41	1	170.41	--	--
A x B x C	68.60	2	34.30	--	--
A x B x D	268.87	2	134.43	--	--
A x C x D	283.27	2	141.63	--	--
B x C x D	4.41	1	4.41	--	--
A x B x C x D	16.47	2	8.23	--	--
Within Cells	9409.20	96	98.01		
Total	11212.99	119			

[a]The teacher strongly emphasized grades.

TABLE 43

ANALYSIS OF VARIANCE FOR ITEM FOUR[a]
WOMEN IN PSYCHOLOGY
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SATV (A)	1198.12	2	599.06	7.38	.01
SH (B)	.07	1	.07	--	--
DA (C)	476.01	1	476.01	5.86	.025
ES4 (D)	35.21	1	35.21	--	--
A x B	9.05	2	4.52	--	--
A x C	51.62	2	25.81	--	--
A x D	68.82	2	34.41	--	--
B x C	60.21	1	60.21	--	--
B x D	88.41	1	88.41	--	--
C x D	175.21	1	175.21	--	--
A x B x C	70.82	2	35.41	--	--
A x B x D	177.22	2	88.61	--	--
A x C x D	11.32	2	5.66	--	--
B x C x D	1.87	1	1.87	--	--
A x B x C x D	1.05	2	.52	--	--
Within Cells	7792.80	96	81.18		
Total	10217.79	119			

[a]Most students in class competed for grades.

TABLE 44

ANALYSIS OF VARIANCE FOR ITEM FIVE[a]
WOMEN IN PSYCHOLOGY
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SATV (A)	1412.72	2	706.36	7.55	.01
SH (B)	208.03	1	208.03	--	--
DA (C)	9.63	1	9.63	--	--
ES5 (D)	40.83	1	40.83	--	--
A x B	292.82	2	146.41	--	--
A x C	59.82	2	29.91	--	--
A x D	141.82	2	70.91	--	--
B x C	172.80	1	172.80	--	--
B x D	48.13	1	48.13	--	--
C x D	136.53	1	136.53	--	--
A x B x C	25.65	2	12.82	--	--
A x B x D	4.12	2	2.06	--	--
A x C x D	3.72	2	1.86	--	--
B x C x D	168.03	1	168.03	--	--
A x B x C x D	115.22	2	57.61	--	--
Within Cells	8977.60	96	93.52		
Total	11817.47	119			

[a]The teacher made it clear that he was continually judging the worth of a student's performance in class.

TABLE 45

ANALYSIS OF VARIANCE FOR FREQUENCY OF TEACHER ASSIGNED TESTS
WOMEN IN PSYCHOLOGY
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SATV (A)	1544.62	2	772.31	10.55	.005
SH (B)	170.41	1	170.41	--	--
DA (C)	414.41	1	414.41	5.66	.025
TESTS (D)	14.01	1	14.01	--	--
A x B	303.52	2	151.76	--	--
A x C	82.12	2	41.06	--	--
A x D	47.62	2	23.81	--	--
B x C	8.01	1	8.01	--	--
B x D	95.41	1	95.41	--	--
C x D	29.01	1	29.01	--	--
A x B x C	135.62	2	67.81	--	--
A x B x D	680.12	2	340.06	4.65	N.S.*
A x C x D	32.72	2	16.36	--	--
B x C x D	.08	1	.08	--	--
A x B x C x D	3.95	2	1.97	--	--
Within Cells	7026.40	96	73.19		
Total	10587.99	119			

*Although an F-ratio of 4.65 (2,96 df) has an associated probability level of .025, this effect was called nonsignificant. In view of the large number of nonindependent tests carried out, and lack of theoretical expectancies, it was concluded that this interaction effect was due to errors generated by sampling rather than to any underlying psychological phenomena.

TABLE 46

MEAN GRADE BY LEVEL OF ABILITY FOR STUDENTS IN PSYCHOLOGY AND ECONOMICS (Fall, 1964)

Sample	Item	Low SATV	Mid SATV	High SATV	df	F	P
Men in Psychology	ES3	44.65[a]	50.80	54.20	2,96	9.31	.005
	ES4	46.58	50.80	52.90	2,96	5.47	.01
	ES5	47.88	50.43	53.65	2,96	3.59	.05
	Tests	45.15	50.65	54.68	2,96	10.10	.005
Men in Economics	ES3	1.98[b]	2.55	2.90	2,96	11.18	.005
	ES4	2.08	2.68	2.95	2,96	10.01	.005
	ES5	1.90	2.65	2.73	2,96	10.67	.005
	Tests	1.93	2.50	2.95	2,96	13.37	.005
Women in Psychology	ES3	47.53[a]	48.58	51.93	2,96	2.16	N.S.
	ES4	48.30	48.83	55.25	2,96	7.38	.01
	ES5	47.88	49.43	55.00	2,96	7.55	.01
	Tests	46.20	51.63	54.90	2,96	10.55	.005

[a]Grades in Psychology were converted to standard scores within sections (mean of 50, S.D. of 10).

[b]Means in Economics are based on raw grades (A = 4, B = 3, C = 2, D = 1, E = 0).

TABLE 47

MEAN GRADE BY LEVEL OF STUDY HABITS FOR STUDENTS IN PSYCHOLOGY AND ECONOMICS (Fall, 1964)

Sample	Item	Low SH	High SH	df	F	P
Men in Psychology	ES3	49.07[a]	50.70	1,96	.80	N.S.
	ES4	48.57	51.62	1,96	2.15	N.S.
	ES5	48.38	52.92	1,96	6.60	.025
	Tests	48.57	51.75	1,96	4.12	.05
Men in Economics	ES3	2.30[b]	2.65	1,96	4.71	.05
	ES4	2.45	2.68	1,96	2.04	N.S.
	ES5	2.22	2.63	1,96	6.68	.025
	Tests	2.33	2.58	1,96	3.81	.05<p<.10
Women in Psychology	ES3	48.22[a]	50.47	1,96	1.55	N.S.
	ES4	50.77	50.82	1,96	.00	N.S.
	ES5	49.28	52.25	1,96	2.22	N.S.
	Tests	49.72	52.10	1,96	2.33	N.S.

[a]Grades in Psychology were converted to standard scores within sections (mean of 50, S.D. of 10).

[b]Means in Economics are based on raw grades (A = 4, B = 3, C = 2, D = 1, E = 0).

TABLE 48

MEAN GRADE BY LEVEL OF DEBILITATING ANXIETY FOR STUDENTS IN PSYCHOLOGY AND ECONOMICS (Fall, 1964)

Sample	Item	Low DA	High DA	df	F	P
Men in Psychology	ES3	51.43[a]	48.33	1,96	2.86	.05<p<.10
	ES4	52.05	48.13	1,96	8.38	.005
	ES5	52.43	48.87	1,96	4.09	.05
	Tests	52.25	48.07	1,96	4.85	.05
Men in Economics	ES3	2.48[b]	2.47	1,96	.01	N.S.
	ES4	2.72	2.42	1,96	3.38	.05<p<.10
	ES5	2.50	2.35	1,96	.53	N.S.
	Tests	2.50	2.42	1,96	.27	N.S.
Women in Psychology	ES3	49.58[a]	49.10	1,96	.07	N.S.
	ES4	52.78	48.80	1,96	5.86	.025
	ES5	50.73	50.80	1,96	.10	N.S.
	Tests	52.77	49.05	1,96	5.66	.025

[a]Grades in Psychology were converted to standard scores within sections (mean of 50, S.D. of 10).

[b]Means in Economics are based on raw grades (A = 4, B = 3, C = 2, D = 1, E = 0).

APPENDIX L

THE COMPLETE ANALYSIS OF VARIANCE FOR THE EXPERIMENTAL STUDY

TABLE 49

ANALYSIS OF VARIANCE FOR EXPERIMENTAL CONDITIONS MEN IN PSYCHOLOGY (Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SAT TOTAL (A)	1193.52	2	596.76	22.96	.005
SH (B)	46.87	1	46.87	--	--
DA (C)	95.41	1	95.41	3.67	.05<p<.10
CONDITIONS (D)	946.41	1	946.41	36.41	<.005
A x B	21.05	2	10.53	--	--
A x C	26.52	2	13.26	--	--
A x D	115.02	2	57.51	--	--
B x C	88.41	1	88.41	3.40	.05<p<.10
B x D	.01	1	.01	--	--
C x D	16.87	1	16.87	--	--
A x B x C	78.52	2	39.26	--	--
A x B x D	45.02	2	22.51	--	--
A x C x D	38.15	2	19.08	--	--
B x C x D	2.41	1	2.41	--	--
A x B x C x D	59.82	2	29.91	--	--
Within Cells	2494.80	96	25.99		
Total	5268.79	119			

TABLE 50

ANALYSIS OF VARIANCE FOR EXPERIMENTAL CONDITIONS
WOMEN IN PSYCHOLOGY
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SAT TOTAL (A)	2396.32	2	1198.16	50.41	<.005
SH (B)	63.07	1	63.07	--	--
DA (C)	15.41	1	15.41	--	--
CONDITIONS (D)	476.01	1	476.01	20.03	<.005
A x B	4.55	2	2.28	--	--
A x C	47.62	2	23.81	--	--
A x D	41.82	2	20.91	--	--
B x C	170.41	1	170.41	7.17	.01
B x D	.21	1	.21	--	--
C x D	33.07	1	33.07	--	--
A x B x C	35.12	2	17.56	--	--
A x B x D	31.52	2	15.76	--	--
A x C x D	15.05	2	7.52	--	--
B x C x D	6.07	1	6.07	--	--
A x B x C x D	108.15	2	54.08	--	--
Within Cells	2281.60	96	23.77		
Total	5725.99	119			

APPENDIX M

ANALYSIS OF CRITERIA TEST DATA WITH ITEM DIFFICULTY CONTROLLED

Although a stress x anxiety interaction had been anticipated in the experimental study, results presented in Chapter IV failed to confirm this expectation. Since task complexity has been known to interact with anxiety level, these factors could have combined in such a way so as to mask anticipated effects when total score was used as the basis for analysis. Therefore, Criteria Test data were reanalyzed with control on item difficulty level in order to determine effects for task complexity more precisely.

Procedure

Item difficulty was defined in terms of the percentage of students passing an item. Easy items would be those passed by almost everyone; difficult items would be passed by only a few. Since item difficulties for Form X of the Criteria Test had been determined during standardization of the test (Milholland, 1964), these values were used for item selection for this analysis, also. The 10 items with the highest and 10 items with the lowest per cent passing were chosen. These items are identified in Table 51.

Answer sheets were re-scored and a sub-scale score for each level of difficulty was determined. The data, based on these sub-scores, were analyzed by means of a four-way analysis of variance with control on ability, study habits, debilitating anxiety, and evaluative stress.

TABLE 51

PER CENT PASSING FOR THE TEN EASIEST AND TEN MOST DIFFICULT ITEMS ON THE INTRODUCTORY PSYCHOLOGY CRITERIA TEST, FORM X
(Spring, 1962; N = 64-67)

Easy Items		Difficult Items	
Item Number	Per Cent Passing	Item Number	Per Cent Passing
06.	79	07.	44
10.	89	17.	45
15.	85	18.	29
33.	84	19.	40
34.	83	25.	42
40.	95	29.	38
43.	83	30.	34
44.	87	47.	34
52.	83	55.	41
54.	79	58.	37
Mean	84.7		38.4

Results

Significant main effects were observed for intellectual ability and evaluative stress, and were consistent with results based on full scale scores (Tables 52 and 53). Again, high ability was associated with highest levels of performance irrespective of difficulty level.

TABLE 52

MEAN SCORES BY LEVEL OF ABILITY FOR TEN EASY AND TEN DIFFICULT CRITERIA TEST ITEMS, MEN AND WOMEN IN PSYCHOLOGY (Fall, 1964)

	Items	Low SAT Total	Mid SAT Total	High SAT Total	df	F	P
Men	Easy	7.25	7.73	8.35	2,96	7.20	.01
	Difficult	4.05	4.90	5.33	2,96	6.85	.01
Women	Easy	7.15	7.75	8.55	2,96	10.44	.005
	Difficult	3.93	4.85	5.13	2,96	6.43	.01

TABLE 53

MEAN SCORES BY CONDITIONS FOR TEN EASY AND TEN DIFFICULT CRITERIA TEST ITEMS, MEN AND WOMEN IN PSYCHOLOGY (Fall, 1964)

	Items	Nonevaluative	Evaluative	df	F	P
Men	Easy	7.32	8.23	1,96	14.92	.005
	Difficult	4.38	5.13	1,96	6.85	.01
Women	Easy	7.47	8.17	1,96	7.78	.01
	Difficult	4.18	5.08	1,96	9.88	.01

Students also earned higher scores when they were being graded than under the nonevaluative condition.

As can be seen in Tables 52 and 53, there was also a main effect for difficulty level (Men: $F = 109.38$, $p < .005$; Women: $F = 122.41$, $p < .005$). This outcome is not surprising in view of the criteria used for item selection; these outcomes can be considered as a replication of the standardization group's results.

An interesting interaction between evaluative stress and level of ability emerged (Table 54). Men tended to perform at uniformly high levels on easy items during the evaluative condition irrespective of

TABLE 54

MEAN CRITERIA TEST SCORES FOR THE INTERACTION OF ABILITY AND CONDITIONS, MEN AND WOMEN IN PSYCHOLOGY (Fall, 1964)

		Low SAT Total	Mid SAT Total	High SAT Total	df	F	P
Men (Easy items)	Evaluative	8.10	8.20	8.40	2,96	3.79	.05
	Nonevaluative	6.40	7.25	8.30			
Women (Difficult items)	Evaluative	4.90	5.10	5.25	2,96	3.42	.05
	Nonevaluative	2.95	4.60	5.00			

ability level. However, performance for low and middle ability men tended to fall off under the nonevaluative condition. A similar

interaction was observed for women's performance on the difficult items. How difficulty level enters this interaction is not entirely clear.

Even with level of item difficulty controlled, interactions for anxiety and stress failed to reach significance. These results are described in Tables 21-23 of Chapter IV. Variance estimates, F-ratios, and probability estimates for the complete analysis are given in Tables 55-58 (see pages 157-160).

Discussion

Even with level of task difficulty controlled, the results tend to be highly consistent with those based on scores for the full 60-item Criteria Test. The fact that an ability main effect continues to appear for both easy and difficult items suggests that the Criteria Test is indeed complex and requires high levels of intellectual ability for high performance.

The interaction of ability and evaluative stress is interesting in view of the differential levels of performance attained under the two conditions. When students were being graded, levels of performance were uniformly high although there is still a slight main effect for ability (Table 54). Under the nonevaluative condition, levels of performance become much more disparate for middle and low ability students while high ability students performed almost equally well irrespective of evaluative stress. This outcome suggests that the middle and low ability students are prone to be satisfied with mediocre performance when penalties for nonachievement are not present. On the other hand,

high ability students seem to be more intrinsically motivated as their performance tended to be relatively similar under both conditions. It may well be that test performance is viewed as stimulating and challenging by high ability students and as an opportunity to exercise their mental powers. On the other hand, low ability students would find themselves under constant pressure to maintain satisfactory levels of performance due to their limited potential and would be more apt to be task avoidant in order to escape the feeling of pressure whenever the opportunity arises, as it did here under the nonevaluative condition. Why this interaction effect should appear for men on easy items and for women on difficult items is difficult to understand on any theoretical grounds. Hence, no explanation for this effect is offered at this time.

Although the anticipated anxiety x stress x task complexity interaction effects began to emerge when item difficulty was controlled, these effects failed to reach significance. As mentioned in Chapter IV, the effect of anxiety arousal may have been attenuated by the lack of time limits in performing the test. Thus, persistent effort could be utilized to overcome some of the debilitating effects of anxiety arousal.

On the basis of these results it seems clear that the Criteria Test is a difficult achievement task in view of the relatively high degree of association between performance level and intellectual ability. The fact that anxiety did not interact with level of stress or with item difficulty seems to be due to the absence of time constraints during test performance. If our explanation for this lack of interaction is

valid, than this outcome suggests that the academic test performance of highly anxious students could be enhanced by removing time pressures from course examinations. However, the student must be willing to persist at the task in order to take advantage of the time factor.

TABLE 55

ANALYSIS OF VARIANCE FOR TEN EASY CRITERIA TEST ITEMS
MEN IN PSYCHOLOGY
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SAT TOTAL (A)	24.35	2	12.17	7.20	.01
SH (B)	.67	1	.67	--	--
DA (C)	3.67	1	3.67	--	--
CONDITIONS (D)	25.21	1	25.21	14.92	.005
A x B	10.55	2	5.27	--	--
A x C	9.15	2	4.57	--	--
A x D	12.82	2	6.41	3.79	.05
B x C	3.67	1	3.67	--	--
B x D	1.01	1	1.01	--	--
C x D	.67	1	.67	--	--
A x B x C	1.95	2	.97	--	--
A x B x D	1.22	2	.61	--	--
A x C x D	.15	2	.07	--	--
B x C x D	.41	1	.41	--	--
A x B x C x D	1.02	2	.51	--	--
Within Cells	162.40	96	1.69		
Total	258.93	119			

TABLE 56

ANALYSIS OF VARIANCE FOR TEN DIFFICULT CRITERIA TEST ITEMS
MEN IN PSYCHOLOGY
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SAT TOTAL (A)	33.72	2	16.86	6.85	.01
SH (B)	.67	1	.67	--	--
DA (C)	1.41	1	1.41	--	--
CONDITIONS (D)	16.87	1	16.87	6.85	.01
A x B	3.95	2	1.97	--	--
A x C	2.92	2	1.46	--	--
A x D	.65	2	.33	--	--
B x C	.01	1	.01	--	--
B x D	1.41	1	1.41	--	--
C x D	3.01	1	3.01	--	--
A x B x C	1.82	2	.91	--	--
A x B x D	5.02	2	2 51	--	--
A x C x D	3.72	2	1.86	--	--
B x C x D	1.41	1	1.41	--	--
A x B x C x D	1.02	2	.51	--	--
Within Cells	236.40	96	2.46		
Total	313.99	119			

TABLE 57

ANALYSIS OF VARIANCE FOR TEN EASY CRITERIA TEST ITEMS
WOMEN IN PSYCHOLOGY
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SAT TOTAL (A)	39.47	2	19.73	10.44	.005
SH (B)	2.70	1	2.70	--	--
DA (C)	.53	1	.53	--	--
CONDITIONS (D)	14.70	1	14.70	7.78	.01
A x B	.60	2	.30	--	--
A x C	.47	2	.23	--	--
A x D	3.20	2	1.60	--	--
B x C	6.53	1	6.53	3.46	.05<p<.10
B x D	.83	1	.83	--	--
C x D	1.20	1	1.20	--	--
A x B x C	2.47	2	1.23	--	--
A x B x D	2.87	2	1.43	--	--
A x C x D	.20	2	.10	--	--
B x C x D	1.20	1	1.20	--	--
A x B x C x D	3.80	2	1.90	--	--
Within Cells	181.20	96	1.89		
Total	261.97	119			

TABLE 58

ANALYSIS OF VARIANCE FOR TEN DIFFICULT CRITERIA TEST ITEMS
WOMEN IN PSYCHOLOGY
(Fall, 1964)

Source of Variation	Sums of Squares	df	Variance Estimate	F	P
SAT TOTAL (A)	31.62	2	15.81	6.43	.01
SH (B)	1.20	1	1.20	--	--
DA (C)	.03	1	.03	--	--
CONDITIONS (D)	24.30	1	24.30	9.88	.01
A x B	8.75	2	4.37	--	--
A x C	.42	2	.21	--	--
A x D	16.85	2	8.42	3.42	.05
B x C	5.63	1	5.63	2.29	.10<p<.20
B x D	.03	1	.03	--	--
C x D	6.53	1	6.53	2.65	.10
A x B x C	1.02	2	.51	--	--
A x B x D	.72	2	.36	--	--
A x C x D	2.92	2	1.46	--	--
B x C x D	.13	1	.13	--	--
A x B x C x D	1.32	2	.66	--	--
Within Cells	236.40	96	2.46		
Total	337.87	119			

APPENDIX N

RAW DATA

This Appendix contains means based on the raw data used in each analysis of variance. Scale scores used for determining levels of ability, study habits, and debilitating anxiety can be found in Appendix H.

MEAN GRADE BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND THE TEACHER'S EMPHASIS ON GRADES (ES3), MEN IN PSYCHOLOGY (Fall, 1964)

SATV	SH	DA	Low ES3	High ES3
High	High	High	53.80[a]	51.80
		Low	58.00	54.60
	Low	High	52.60	51.60
		Low	52.40	58.80
Mid	High	High	49.40	49.80
		Low	51.00	51.00
	Low	High	54.00	50.40
		Low	47.80	53.00
Low	High	High	38.80	50.20
		Low	48.80	51.20
	Low	High	38.60	39.00
		Low	45.60	45.00

[a]Grades converted to standard scores (mean = 50, S.D. = 10).

MEAN GRADE BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND STUDENT COMPETITION FOR GRADES (ES4), MEN IN PSYCHOLOGY (Fall, 1964)

SATV	SH	DA	Low ES4	High ES4
High	High	High	49.80[a]	50.00
		Low	57.60	56.60
	Low	High	47.60	50.40
		Low	52.20	59.00
Mid	High	High	50.20	49.80
		Low	59.40	49.80
	Low	High	51.20	52.60
		Low	40.20	53.20
Low	High	High	46.00	38.60
		Low	48.60	53.00
	Low	High	40.40	41.00
		Low	48.20	46.80

[a]Grades converted to standard scores (mean = 50, S.D. = 10).

MEAN GRADE BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND TEACHER'S JUDGMENT OF PERFORMANCE IN CLASS (ES5), MEN IN PSYCHOLOGY
(Fall, 1964)

SATV	SH	DA	Low ES5	High ES5
High	High	High	55.00[a]	55.00
		Low	59.20	59.80
	Low	High	50.00	43.00
		Low	52.00	55.20
Mid	High	High	46.40	50.20
		Low	55.40	53.40
	Low	High	53.00	49.00
		Low	51.20	44.80
Low	High	High	47.00	51.00
		Low	52.00	50.60
	Low	High	46.00	40.80
		Low	43.60	52.00

[a]Grades converted to standard scores (mean = 50, S.D. = 10).

MEAN GRADE BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND FREQUENCY OF TEACHER ASSIGNED TESTS, MEN IN PSYCHOLOGY (Fall, 1964)

SATV	SH	DA	One-two Tests	Three-four Tests
High	High	High	53.00[a]	49.20
		Low	60.20	60.80
	Low	High	50.80	45.00
		Low	53.40	65.00
Mid	High	High	49.20	51.40
		Low	51.80	54.20
	Low	High	53.80	47.00
		Low	49.80	48.00
Low	High	High	49.20	44.40
		Low	48.20	49.40
	Low	High	40.60	43.20
		Low	43.60	42.60

[a]Grades converted to standard scores (mean = 50, S.D. = 10).

MEAN GRADE BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND THE TEACHER'S EMPHASIS ON GRADES (ES3), MEN IN ECONOMICS (Fall, 1964)

SATV	SH	DA	Low ES3	High ES3
High	High	High	3.40[a]	2.60
		Low	3.40	2.80
	Low	High	2.80	2.80
		Low	2.60	2.80
Mid	High	High	2.80	2.80
		Low	2.60	2.40
	Low	High	1.80	2.60
		Low	2.60	2.80
Low	High	High	2.40	1.80
		Low	2.00	2.80
	Low	High	1.80	2.00
		Low	1.20	1.80

[a]Raw grades (A = 4, B = 3, C = 2, D = 1, E = 0).

MEAN GRADE BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND STUDENT COMPETITION FOR GRADES (ES4), MEN IN ECONOMICS
(Fall, 1964)

SATV	SH	DA	Low ES4	High ES4
High	High	High	2.60[a]	3.00
		Low	3.00	3.40
	Low	High	2.60	3.00
		Low	3.20	2.80
Mid	High	High	2.00	3.00
		Low	2.80	2.80
	Low	High	2.60	2.40
		Low	2.80	3.00
Low	High	High	2.00	2.40
		Low	2.80	2.40
	Low	High	1.80	1.60
		Low	1.60	2.00

[a]Raw grades (A = 4, B = 3, C = 2, D = 1, E = 0).

MEAN GRADE BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND TEACHER'S JUDGMENT OF PERFORMANCE IN CLASS (ES5), MEN IN ECONOMICS
(Fall, 1964)

SATV	SH	DA	Low ES5	High ES5
High	High	High	3.00[a]	3.20
		Low	3.60	3.00
	Low	High	2.20	2.60
		Low	2.20	2.00
Mid	High	High	2.60	2.40
		Low	3.00	2.60
	Low	High	2.60	2.40
		Low	2.60	3.00
Low	High	High	2.00	1.80
		Low	2.00	2.40
	Low	High	1.80	1.80
		Low	2.00	1.40

[a]Raw grades (A = 4, B = 3, C = 2, D = 1, E = 0).

MEAN GRADE BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND FREQUENCY OF TEACHER ASSIGNED QUIZZES, MEN IN ECONOMICS
(Fall, 1964)

SATV	SH	DA	One-two Quizzes	Three-four Quizzes
High	High	High	3.40[a]	3.20
		Low	3.60	2.40
	Low	High	2.60	3.20
		Low	2.20	3.00
Mid	High	High	2.60	2.60
		Low	2.80	2.60
	Low	High	2.20	2.00
		Low	2.60	2.60
Low	High	High	1.40	2.20
		Low	2.20	2.40
	Low	High	1.60	2.00
		Low	1.80	1.80

[a]Raw grades (A = 4, B = 3, C = 2, D = 1, E = 0).

MEAN GRADE BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND THE TEACHER'S EMPHASIS ON GRADES (ES3), WOMEN IN PSYCHOLOGY (Fall, 1964)

SATV	SH	DA	Low ES3	High ES3
High	High	High	56.20[a]	53.60
		Low	51.00	53.60
	Low	High	52.20	45.80
		Low	50.20	53.00
Mid	High	High	48.60	53.60
		Low	49.20	49.00
	Low	High	46.40	47.80
		Low	47.40	47.60
Low	High	High	54.40	43.20
		Low	46.20	47.00
	Low	High	43.40	44.20
		Low	45.80	56.00

[a]Grades converted to standard scores (mean = 50, S.D. = 10).

MEAN GRADE BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND STUDENT COMPETITION FOR GRADES (ES4), WOMEN IN PSYCHOLOGY (Fall, 1964)

SATV	SH	DA	Low ES4	High ES4
High	High	High	55.60[a]	54.00
		Low	53.00	57.00
	Low	High	53.60	51.00
		Low	57.40	60.40
Mid	High	High	48.40	46.20
		Low	50.60	50.60
	Low	High	48.60	46.60
		Low	48.80	50.80
Low	High	High	48.80	44.20
		Low	50.40	51.00
	Low	High	41.80	46.80
		Low	46.00	57.40

[a]Grades converted to standard scores (mean = 50, S.D. = 10).

MEAN GRADE BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND TEACHER'S JUDGMENT OF PERFORMANCE IN CLASS (ES5), WOMEN IN PSYCHOLOGY
(Fall, 1964)

SATV	SH	DA	Low ES5	High ES5
High	High	High	56.80[a]	58.40
		Low	57.60	55.60
	Low	High	51.00	54.20
		Low	50.40	56.00
Mid	High	High	56.00	53.80
		Low	51.20	48.60
	Low	High	42.40	49.20
		Low	49.60	44.60
Low	High	High	47.20	46.80
		Low	45.00	50.00
	Low	High	40.60	53.20
		Low	50.60	49.60

[a]Grades converted to standard scores (mean = 50, S.D. = 10).

MEAN GRADE BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND FREQUENCY OF TEACHER ASSIGNED TESTS, WOMEN IN PSYCHOLOGY
(Fall, 1964)

SATV	SH	DA	One-two Tests	Three-four Tests
High	High	High	55.80[a]	57.40
		Low	55.60	58.20
	Low	High	50.60	48.20
		Low	57.60	55.80
Mid	High	High	49.20	57.20
		Low	49.60	61.40
	Low	High	52.40	44.40
		Low	50.40	48.40
Low	High	High	44.20	39.20
		Low	50.80	46.60
	Low	High	43.00	47.00
		Low	47.60	51.20

[a]Grades converted to standard scores (mean = 50, S.D. = 10).

MEAN CRITERIA TEST SCORES BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND EVALUATIVE STRESS, MEN IN PSYCHOLOGY
(Fall, 1964)

SAT Total	SH	DA	Nonevaluative Condition	Evaluative Condition
High	High	High	39.60	40.80
		Low	41.20	44.80
	Low	High	35.60	41.20
		Low	40.80	41.20
Mid	High	High	31.60	41.00
		Low	33.20	42.40
	Low	High	34.00	39.60
		Low	33.40	40.80
Low	High	High	29.40	32.60
		Low	33.60	39.40
	Low	High	31.60	35.80
		Low	25.80	36.60

MEAN CRITERIA TEST SCORES BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND EVALUATIVE STRESS, WOMEN IN PSYCHOLOGY
(Fall, 1964)

SAT Total	SH	DA	Nonevaluative Condition	Evaluative Condition
High	High	High	42.00	43.00
		Low	41.80	43.20
	Low	High	42.80	42.40
		Low	36.80	44.20
Mid	High	High	33.80	38.60
		Low	35.00	43.40
	Low	High	35.20	38.20
		Low	34.00	37.40
Low	High	High	28.40	29.80
		Low	30.60	38.60
	Low	High	27.40	35.20
		Low	27.80	31.00

MEAN SCORES FOR TEN EASY CRITERIA TEST ITEMS BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND EVALUATIVE STRESS, MEN IN PSYCHOLOGY
(Fall, 1964)

SAT Total	SH	DA	Nonevaluative Condition	Evaluative Condition
High	High	High	8.00	8.40
		Low	8.40	8.80
	Low	High	8.80	8.80
		Low	8.00	7.60
Mid	High	High	6.20	7.40
		Low	7.40	8.80
	Low	High	7.00	8.20
		Low	8.40	8.40
Low	High	High	6.60	8.40
		Low	7.20	8.60
	Low	High	5.80	7.60
		Low	6.00	7.80

MEAN SCORES FOR TEN DIFFICULT CRITERIA TEST ITEMS BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY AND EVALUATIVE STRESS, MEN IN PSYCHOLOGY
(Fall, 1964)

SAT Total	SH	DA	Nonevaluative Condition	Evaluative Condition
High	High	High	5.40	5.40
		Low	5.60	6.20
	Low	High	4.00	5.40
		Low	5.20	5.40
Mid	High	High	4.20	4.80
		Low	4.20	6.00
	Low	High	4.80	5.20
		Low	4.60	5.40
Low	High	High	4.60	4.00
		Low	3.40	4.20
	Low	High	3.60	4.40
		Low	3.00	5.20

MEAN SCORES FOR TEN EASY CRITERIA TEST ITEMS BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY, AND EVALUATIVE STRESS, WOMEN IN PSYCHOLOGY
(Fall, 1964)

SAT Total	SH	DA	Nonevaluative Condition	Evaluative Condition
High	High	High	8.00	9.00
		Low	8.80	8.60
	Low	High	8.80	9.00
		Low	8.00	8.20
Mid	High	High	7.20	8.00
		Low	7.60	9.00
	Low	High	7.20	8.80
		Low	6.80	7.40
Low	High	High	7.40	7.20
		Low	7.20	7.60
	Low	High	6.00	8.00
		Low	6.60	7.20

MEAN SCORES FOR TEN DIFFICULT CRITERIA TEST ITEMS BY LEVEL OF ABILITY, STUDY HABITS, DEBILITATING ANXIETY AND EVALUATIVE STRESS, WOMEN IN PSYCHOLOGY
(Fall, 1964)

SAT Total	SH	DA	Nonevaluative Condition	Evaluative Condition
High	High	High	4.60	5.20
		Low	5.20	5.60
	Low	High	5.40	5.20
		Low	4.80	5.00
Mid	High	High	5.20	4.80
		Low	4.40	5.60
	Low	High	5.00	4.80
		Low	3.80	5.20
Low	High	High	2.60	3.60
		Low	2.40	5.20
	Low	High	3.80	5.60
		Low	3.00	5.20

REFERENCES

REFERENCES

Atkinson, J. W. The achievement motive and recall of interrupted and completed tasks. Journal of Experimental Psychology, 1953, 46, 381-390.

Atkinson, J. W. Motivational determinants of risk-taking behavior. Psychological Review, 1957, 63, 359-372.

Atkinson, J. W. (Ed.). Motives in Fantasy, Action, and Society. New York: Van Nostrand, 1958.

Atkinson, J. W. and Litwin, G. H. Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure. Journal of Abnormal and Social Psychology, 1960, 60, 52-63.

Atkinson, J. W. and Reitman, W. R. Performance as a function of motive strength and expectancy of goal attainment. Journal of Abnormal and Social Psychology, 1956, 53, 361-366.

Beam, J. C. Serial learning and conditioning under real-life stress. Journal of Abnormal and Social Psychology, 1955, 51, 543-551.

Bloom, B. S. (Ed.). Taxonomy of Educational Objectives, Handbook I: Cognitive Domain. New York: Longmans, Green, 1956.

Brown, W. F. and Holtzman, W. H. A study-attitude questionnaire for predicting academic success. Journal of Educational Psychology, 1955, 46, 75-84.

Carney, R. E. An analysis of university student behavior with measures of ability, attitude, performance, and personality. Unpublished doctoral dissertation, University of Michigan, 1961.

Carrier, Neil A. Stress, personality, and performance on course examinations. Unpublished doctoral dissertation, University of Michigan, 1956.

Child, I. L. Personality. Annual Review of Psychology. 1954, 5, 149-170.

Deese, J., Lazarus, R. S., and Keenan, J. Anxiety, anxiety reduction, and stress in learning. Journal of Experimental Psychology, 1953, 46, 55-60.

Dixon, W. J. and Massey, F. J. Introduction to Statistical Analysis (2nd ed.). New York: McGraw-Hill, 1957.

Dollard, J. and Miller, N. E. Personality and Psychotherapy. New York: McGraw-Hill, 1950.

Eriksen, C. W. and Paul G. L. Effects of test anxiety on "real-life" examinations. Journal of Personality, 1964, 32, 480-494.

Farber, I. E. and Spence, K. W. Complex learning and conditioning as a function of anxiety. Journal of Experimental Psychology, 1953, 45, 2, 120-125.

Farber, I. E. and Spence, K. W. Effects of anxiety, stress and task variables and reaction time. Journal of Personality, 1956, 35, 1-18.

Flanders, N. A. Teacher-pupil contacts and mental hygiene. Journal of Social Issues, 15, 1, 1959, 30-38.

Flanders, N. A. Teacher Influence, Pupil Attitudes, and Achievement. Cooperative Research Monograph No. 12. Washington: U.S. Gov't. Printing Office, 1965.

French, J. W. A study of emotional states aroused during examinations. Princeton, N.J.: Educational Testing Service Research Bulletin R B-61-6, March, 1961.

Freud, S. The Problem of Anxiety. New York: Norton, 1936.

Gordon, E. M. and Sarason, S. B. The relationship between test anxiety and other anxieties. Journal of Personality, 1955, 23, 317-323.

Grice, G. R. Discrimination reaction time as a function of anxiety and intelligence. Journal of Abnormal and Social Psychology, 1955, 50, 71-74.

Haber, R. N. and Alpert, R. Anxiety in academic achievement situations. Journal of Abnormal and Social Psychology, 1960, 61, 207-215.

Hays, W. L. Statistics for Psychologists. New York: Holt, Rinehart, and Winston, 1963.

Kagan, J., Sontag, L. W., Baker, C. T., and Nelson, V. L. Personality and I. Q. Change. Journal of Abnormal and Social Psychology, 1958, 56, 261-266.

Klugh, H. E. and Bendig, A. W. The manifest anxiety and ACE scales and college achievement. Journal of Consulting Psychology, 1955, 19, 487.

Lazarus, R. and Eriksen, C. W. Effects of failure stress upon skilled performance. Journal of Experimental Psychology, 1952, 43, 100-105.

Lin, Y. G. Student characteristics related to achievement (main effects). In Research on the Characteristics of Effective College Teaching. Final Report of Cooperative Research Project No. OE 850, Ann Arbor, Michigan, 1964.

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

McClelland, D. C., Atkinson, J. W., Clark, R. A., and Lowell, E L. The Achievement Motive. New York: Appleton-Century-Crofts, 1953.

McKeachie, W. J. Motivation, teaching methods, and college learning. Nebraska Symposium on Motivation. Lincoln: University of Nebraska Press, 1961.

McKeachie, W. J. Research on teaching at the college and university level. In Gage, N. L. (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

McKeachie, W. J., Pollie, D., and Speisman, J. Relieving anxiety in classroom examinations. Journal of Abnormal and Social Psychology, 1955, 50, 93-98.

Maltzman, I., Fox, J., and Morrisett, L. Some effects of manifest anxiety on mental set. Journal of Experimental Psychology, 1953, 46, 50-54.

Mandler, G. and Sarason, S. B. A study of anxiety and learning. Journal of Abnormal and Social Psychology, 1953, 47, 166-173.

Milholland, J. E. Measuring cognitive abilities. In Research on the Characteristics of Effective College Teaching. Final Report of Cooperative Research Project No. OE 850, Ann Arbor, Michigan, 1964.

Moldowsky, S., and Moldowsky, P. C. Digit span as an anxiety indicator. Journal of Consulting Psychology, 1952, 2, 115-118.

Montague, E. K. The role of anxiety on serial rote learning. Journal of Experimental Psychology, 1953, 45, 91-95.

Montague, E. K. and Schneider, L An empirical study of the effect of anxiety on the learning and retention of nonsense syllables. Unpublished study, University of Iowa, 1949.

O'Conner, P. A. The representation of the motive to avoid failure in thematic apperception. Unpublished doctoral dissertation, University of Michigan, 1959.

Ramond, C. K. Anxiety and task as determiners of verbal performance. Journal of Experimental Psychology, 1953, 46, 2, 120-124.

Sarason, I. G. Effect of anxiety and two kinds of motivating instructions on verbal learning. Journal of Abnormal and Social Psychology, 1957, 54, 166-171.

Sarason, I. G. Test anxiety and the intellectual performance of college students. Journal of Educational Psychology, 1961, 52, 201-206.

Sarason, I. G. Test anxiety and intellectual performance. Journal of Abnormal and Social Psychology, 1963, 66, 73-75.

Sarason, I. G. Test anxiety, general anxiety, and intellectual performance. Journal of Consulting Psychology, 1957, 21, 485-490.

Sarason, S. B. and Mandler, G. Some correlates of test anxiety. Journal of Abnormal and Social Psychology, 1953, 47, 728-731.

Sarason, S. B., Mandler, G., and Craighill, P. C. The effect of differential instructions on anxiety and learning. Journal of Abnormal and Social Psychology, 1952, 47, 561-565.

Sarason, I. G. and Palola, E. The relationship of test and general anxiety, difficulty of task, and experimental instructions to performance. Journal of Experimental Psychology, 1960, 59, 185-191.

Siegel, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.

Sinick, D. Two anxiety scales correlated and examined for sex differences. Journal of Clinical Psychology, 1956, 12, 394-395.

Smith, D. E. P., Wood, R. L., Downer, J. W., and Raygor, A. L. Reading improvement as a function of student personality and teaching method. Journal of Educational Psychology, 1956, 47, 47-59.

Spence, K. W. A theory of emotionally based drive (D) and its relation to performance in simple learning situations. American Psychologist, 1958, 13, 131-141.

Sperber, Z. Test anxiety and performance under stress. Journal of Consulting Psychology, 1961, 25, 3, 226-233.

Spielberger, C. D. On the relationship between manifest anxiety and intelligence. Journal of Consulting Psychology, 1958, 22, 220-224.

Spielberger, C. D. and Katzenmeyer, W. G. Manifest anxiety, intelligence, and college grades Journal of Consulting Psychology, 1959, 23, 278.

Stakenas, R. G. An investigation of interactions between student anxiety and teacher produced anxiety cues Unpublished paper, University of Michigan, 1964.

Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. Journal of Experimental Psychology, 1951, 41, 81-92.

Taylor, J. A. A personality scale of manifest anxiety. Journal of Abnormal and Social Psychology, 1953, 48, 285-290.

Taylor, J. A. and Spence, K. W. The relationship of anxiety level to performance in serial learning. Journal of Experimental Psychology, 1952, 44, 2, 61-64.

Tyler, Leona. The Psychology of Human Differences (2nd ed.). New York: Appleton-Century-Crofts, 1956.

Wallen, N. E. and Travers, R. M. Analysis and investigation of teaching methods. In Gage, N. L (Ed.), Handbook of Research on Teaching. Chicago: Rand, McNally, 1963.

Westrope, M. R. Relations among Rorschach indices, manifest anxiety, and performance under stress Journal of Abnormal and Social Psychology, 1953, 48, 4, 515-524.