8116750

CONKRIGHT, ROBERT KEITH

THE EFFECTS OF PERSONAL PROBABILITY SCORING ON ACHIEVEMENT AND TEST ANXIETY

The University of Oklahoma PH.D. 1981

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE EFFECTS OF PERSONAL PROBABILITY SCORING ON ACHIEVEMENT AND TEST ANXIETY

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

ROBERT KEITH CONKRIGHT

Norman, Oklahoma

1980

The Effects of Personal Probability Scoring on Achievement

and Test Anxiety

APPROVED BY:

William ...

George ... Letchworth

W. Alan ...

... Miller

Gerald ...

DISSERTATION COMMITTEE

Acknowledgements

I would like to thank George Letchworth, Ph.D., Gerald Kowitz, Ph.D., Al Nicewander, Ph.D., and Tom Miller, Ph.D. for their assistance in the development of this dissertation. I am especially grateful to William Graves, Ph.D., my dissertation committee chairman, for his help and direction in my academic career and in the development of this dissertation.

I would also like to thank the psychology staff and students at Austin Community College, Austin, Texas, for their cooperation in the collection of my dissertation data.

Ms. Barbara L. Keir has been a continual source of motivation, information, support, and understanding. Words cannot adequately express my appreciation for her help.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

The Effects of Personal Probability Scoring on Achievement and Test Anxiety

Robert Keith Conkright

The University of Oklahoma

Running Head: Personal Probability Scoring

The Effects of Personal Probability Scoring on Achievement and Test Anxiety

High test anxious (HTA) persons typically score lower on tests than low test anxious (LTA) persons (Wine, 1971), and several theories have been proposed to explain the effects of anxiety on test performance.

One theory focuses on the way examinees process information (Tobias, 1976). Tobias used an information processing model with three stages: input, processing, and output. Anxiety may interfere with performance by interrupting the flow of information at three points: (1) preprocessing between input and processing, (2) processing, and (3) postprocessing between processing and output. Although one may perceive three separate points of interference in Tobias's model, the three points are probably not independent. The detrimental effects of anxiety may also be cumulative (Tobias, 1976). Tobias suggested that reduced anxiety leads to increased performance because of an uninterrupted flow of information during preprocessing, processing, and postprocessing.

De Finetti (1965) described several personal probability scoring systems, and Brunza (1973) suggested that using a personal probability scoring system reduces anxiety and improves the performance of HTA examinees. Brunza used the five-star personal probability scoring system, and although it reduced anxiety and improved performance, it was found less reliable and valid than the traditional binary scoring (BS) system. Lord and Novick (1968) suggested that the five-star scoring system may be too complex for examinees to understand and use effectively.

Another personal probability scoring system described by de Finetti was the rank order scoring (ROS) system. The ROS system requires examinees to rank alternatives with regard to their personal probability of correctness. It was reported that ROS is at least as reliable as BS (Black, 1974) and that examinees can easily understand and effectively use ROS (Conkright & Williams, 1975). Results indicating a relationship between ROS and anxiety have not been reported.

It is hypothesized that ROS will reduce HTA examinees' test anxiety. The reduced anxiety should allow HTA examinees to improve their performance because of reduced interference during the three stages of information processing.

Method

Examinees

Thirty-two students enrolled in two introductory psychology classes at an urban community college served as examinees. The 18 examinees in one class were referred to as Class 1, and the 14 examinees in the other class were referred to as Class 2. Examinees were male and female and ranged in age from 18 to 33 years. On the second day of class the Test Anxiety Scale (TAS) was administered, and results of the TAS were used to classify examinees as either HTA or LTA. Examinees scoring 20 points or more on the TAS were classified as HTA; examinees scoring 17 points or fewer were classified as LTA. In Class 1, eight examinees were classified as HTA and ten as LTA; in Class 2, seven examinees were classified as HTA and seven as LTA.

Procedure

One instructor taught both classes and presented the same information in 33 lectures to both classes. Lecture outlines were distributed before each lecture and all information in both the outlines and the lectures was taken from a designated text for the course. When an examinee was absent from a lecture, he or she was given an outline for the missed lecture at the next class meeting.

Six examinations were administered during the semester. Each examination consisted of 40 multiple choice items, and each item had four alternatives. The correct answer for each item was discussed in the lectures and provided in the outlines. To determine the complexity of each item, two judges constructed a test plan for each examination (Bloom, 1956). The test plan indicated that the majority of items tested knowledge and comprehension (Appendix A).

Class 1 used ROS on the first three examinations and BS on the last three examinations. Class 2 used BS on the first three examinations and ROS on the last three examinations. A norm-referenced testing approach was used (Karmel & Karmel, 1978). Examinations one and four were considered practice trials and their scores were not used in the analysis of results.

Testing instructions. Examinees in both classes were instructed orally and given a written example of the correct way to use their respective scoring systems at six points: (1) during the first class, (2) during the lecture preceeding the first examination, (3) immediately before the first examination, (4) during the lecture preceeding the fourth examination, and (6) immediately before the fourth examination. On the days of the second, third, fifth, and sixth examinations, oral instructions were given to the examinees.

Examinees using BS were instructed to mark the space that corresponded with the alternative they believed to be correct. Examinees using ROS were instructed to write a "3" in the space that corresponded with the alternative they believed most correct, a "2" in the space for their second choice, a "1" in the space for their third choice, and to leave the space blank for their fourth choice. Examinees were told that the ROS tests would be scored by totaling only the points they assigned to the correct alternatives and converting that total to a percentage of the correct score. The binary scored tests were also converted to a percentage of the correct score.

Administration of the TAS. The TAS was administered four times during the semester: (1) during the first lecture in order to classify examinees as HTA or LTA, (2) immediately before the first examination, (3) immediately after the third examination, and (4) immediately after the sixth examination.

Examinees were told that the TAS was designed to measure their level of anxiety before an examination and that TAS scores would not affect their examination scores or final grades.

After completing the fourth TAS, each examinee was asked to complete a course evaluation that contained Likert type items to determine whether examinees believed ROS improved their course grades, ability to comprehend course information, ability to study for examinations, and enjoyment of lectures. Examinees were also asked to indicate whether ROS reduced their test anxiety and whether they always assigned a "3" to the alternative they would have chosen as correct if BS had been used.

Results

A 2 (HTA/LTA) x 2 (BS/ROS) ANOVA for repeated measures was used to analyze the TAS scores. Results of the TAS administered after the last examination using BS and the last examination using ROS were used to analyze the data. Results of the ANOVA indicated that LTA examinees' TAS scores were significantly different from those of HTA examinees, F (1, 28) = 155.51, $p < .0001$. Tukey's HSD was used as an aid to interpret the significant ANOVA results and revealed that the significant results were due to LTA examinees scoring lower than HTA examinees on the TAS when using both BS ($p < .0001$) and ROS ($p < .0001$). The type of scoring main effect was not significant, F (1, 28) $< 1.$, and neither was the interaction F (1, 28) < 1. Figure 1 depicts the interaction, Table 1 summarizes results of the ANOVA, and Table 2 lists means and standard deviations for all the TAS and examination scores.

Insert Table 1, Table 2, and Figure 1 about here

Rank order scores were transformed to make their expected values for random guessing equal to the expected values for random guessing on binary scored examinations. The number-right transformation was used. Items with a "3" assigned to the correct alternative were counted, and the total was used as the number-right score. The number-right score was then converted to a percentage of the correct score.

Two 2 (HTA/LTA) x 2 (BS/ROS) ANOVAs were used to analyze the percentage of correct scores of the examinations. The first ANOVA analyzed the combined scores of examinations 2 and 3, and the second ANOVA analyzed the combined scores for examinations 5 and 6.

Results of the first ANOVA are summarized in Table 3. The interaction effect was significant, $F\ (1, 28) = 8.64$, $p < .025$. Figure 2 illustrates the interaction effects.

Insert Figure 2 and Table 3 about here

Using a simple main effects test, the combined scores for examinations 2 and 3 were tested for main effect differences. The test indicated that LTA examinees scored significantly higher than HTA examinees on the binary scored examinations, F (1, 28) = 4.74, $p < .05$. It also revealed that HTA examinees significantly improved their scores on the rank order scored examinations, F (1, 28) = 4.95, $p < .05$. No other simple main effects were significant.

Results of the second ANOVA are summarized in Table 4. The interaction effect was significant, F (1, 28) = 6.32, $p < .025$. Figure 3 illustrates the interaction effect.

Insert Figure 3 and Table 4 about here

Using a simple main effects test, the combined scores for examinations 5 and 6 were tested for main effect differences. The test indicated that LTA examinees scored significantly higher on the binary scored examinations than did HTA examinees, F (1, 28) = 4.62, $p < .05$. Results also indicate that HTA examinees' scores significantly improved on rank order scored examinations, F (1, 28) = 4.23, $p < .05$. No other simple main effects were significant.

The Kuder-Richardson formula 20 (K-R 20) was used to measure internal consistency of the binary scored and rank order scored examinations. Results of the K-R 20 indicate that the transformed rank order scores yielded coefficients no lower than the coefficients for the binary scores for all examinations except 1 and 4. None of the differences were significant. Results of the K-R 20 are listed in

Table 5.

Insert Table 5 about here

Results of the Likert-scale course evaluation were analyzed. Chi square tests were used to analyze yes/no responses and t-tests were used to analyze the Likert-scale responses. Results of the evaluation are listed in Table 6. The two χ^2 tests indicated that examinees prefer ROS to BS, χ^2 (1) = 6.31, $p < .02$, and would prefer that other instructors use ROS instead of BS on multiple choice examinations, χ^2 (1) = 6.31, $p < .02$.

Insert Table 6 about here

Only two of the six t-tests yielded significant results. Results of the t-tests indicated that HTA examinees felt ROS was responsible for a higher course grade and LTA examinees did not, t (30) = 2.54, $p < .05$. Analysis of HTA examinees' responses also indicated that they felt significantly more favorable toward ROS, t (30)= 3.48, $p < .01$, than did the LTA examinees. The HTA examinees indicated that ROS only slightly: (1) improved their comprehension of materials, (2) improved their ability to study for examinations, (3) increased their enjoyment of lectures, and (4) decreased their test anxiety; and LTA examinees indicated that ROS had no effect. There was not a significant difference in the LTA and HTA responses.

Discussion

Results of this study indicate that HTA examinees can improve test performance without significantly reducing test anxiety, and this contrasts with predications made from Tobias's model. Tobias (1976) suggested that reduced anxiety is necessary for HTA examinees to improve performance. Tobias's model indicates that anxiety causes poor performance because it interrupts the flow of information at some point between input and output. Results of this study indicate that anxiety may not interrupt the flow of information.

One explanation of these results would be that the HTA examinees improved their performance because ROS requires additional memory searches. Each alternative on a rank order scored examination requires a separate memory search and each additional memory search could increase the probability of selecting the correct answer.

If an examinee knows the correct answer to an item on a BS examination, then the examinee needs only to identify the correct alternative. If the examinee does not know the answer to an item on a BS examination, then the examinee needs to conduct additional memory searches to try to determine the correct alternative. LTA examinees may conduct additional memory searches on a BS examination whereas HTA examinees may not. This may account for the LTA's higher scores on BS examinations.

If HTA examinees score lower than LTA examinees because of inadequate memory search procedures, then HTA examinees' anxiety could be the result of poor test performance. This may explain why HTA examinees show a significant improvement in test performances and only a slight decrease in anxiety when using ROS.

Compared to the measures of internal consistency for BS, the measures of internal consistency for ROS increased with additional use, although not significantly. The slightly improved K-R 20 coefficients could have resulted when HTA examinees learned to effectively use the additional memory searches required by ROS. It is possible that HTA examinees had trouble using additional memory searches on the first ROS examination and successive use of ROS made the process of conducting memory searches easier.

The LTA examinees' performance was not significantly affected in either class by using ROS. On examinations 5 and 6, however, LTA examinees of that class showed a decrease in performance when using ROS. Although the decrease was not significant, it is large enough to suggest that further research be conducted to determine the effect of ROS on the performance of LTA examinees. Sarason (1975) suggested that many methods that increase HTA examinee performance have a detrimental effect on LTA performance. If ROS increases HTA examinee performance and does not decrease LTA examinee performance then ROS would seem to provide a fairer measure of knowledge than BS.

Additional research is required to determine the validity of these suggestions. If ROS were to be tested for a longer period, an entire semester, for example, then HTA examinees should experience a gradual decrease in anxiety after each rank order scored examination, performance should improve significantly after the second examination, and performance should then remain constant. Research could also be conducted to determine whether or not the use of additional memory searches is part of an overall cognitive style (e.g., field-independent/field-dependent; impulsive/reflective). Levels of test anxiety might be correlated with different cognitive styles.

Black, R. H. A method of discriminating partial knowledge (Doctoral dissertation, University of Oklahoma, 1974). Dissertation Abstracts International, 1975, 36, 181A, (University Microfilms No. 75-15, 246).

Bloom, B. S. Taxonomy of Educational Objectives, Handbook I: Cognitive Domain. New York: McKay, 1956.

Brunza, J. J. The effects of repeated probabilistic testing conditions on state anxiety and test reliability (Doctoral dissertation, Indiana University, 1973). Dissertation Abstracts International, 1974, 34, 5706A, (University Microfilms No. 74-7001).

Conkright, R. K., & Williams, J. D. A comparison of binary and rank order scoring methods: Quantitative and affective data. Paper presented at the meeting of the Southwestern Psychological Association, Houston, May 1975.

de Finetti, B. Methods of discriminating levels of partial knowledge concerning a test item. British Journal of Mathematical and and Statistical Psychology, 1965, 19, 87-123.

Karmel, L. J., & Karmel, M. O. Measurement and Evaluation in the Schools. New York: McMillan, 1978.

Lord, F. M., & Novick, M. F. Statistical theories of mental test scores. Reading, Massachusetts: Addison Wesley, 1968.

Tobias, S. Achievement treatment interactions. Review of Educational Research, 1976, 46, 61-74.

Sarason, I. G. Test Anxiety, Attention, and the General Problem of Anxiety. In C. D. Spielberger & I. G. Sarason (Eds.), Stress and Anxiety (Vol 1). Washington, D.C.: Hemisphere, 1975.

Wine, J. D. Test anxiety and direction of attention. Psychological Bulletin, 1971, 76, 92-104.

Table 1

Results of the 2 x 2 ANOVA for TAS scores

SOURCE	SS	df	MS	F
ANXIETY	4001.12	1	4001.12	155.51*
SCORING	19.83	1	19.83	.77
ANXIETY x SCORING	1.12	1	1.12	.04
WITHIN	720.43	28	25.73	
TRIALS	10.85	1	10.85	2.46
TRIALS x ANXIETY	7.75	1	7.75	1.76
TRIALS x SCORING	17.57	1	17.57	3.99
TRIALS x ANXIETY x SCORING	.37	1	.37	.08
WITHIN	123.49	28	4.41	
TOTALS	4902.65	63		

*$p < .0001$

Table 2

Means and Standard Deviations for All Examinations and TAS Scores [a]

	HTA				LTA			
	BS		ROS		BS		ROS	
	$\bar{X}$	s_d	$\bar{X}$	s_d	$\bar{X}$	s_d	$\bar{X}$	s_d
EXAMINATIONS								
1	57	10.12	71	10.06	70	11.98	74	9.34
2	56	10.20	77	10.95	71	9.72	70	11.56
3	59	10.24	78	14.84	72	13.60	67	15.92
4	60	12.84	71	11.48	74	13.42	70	12.13
5	61	13.31	72	14.87	75	9.59	66	11.73
6	70	13.73	74	15.74	77	12.71	73	14.84
TAS SCORES								
1	29	4.73	27	3.28	12	3.53	11	4.96
2	29	4.75	26	3.29	12	3.68	11	5.12
3	28	4.50	26	3.24	12	3.40	10	4.93
4	26	4.60	26	3.21	10	2.63	11	4.42

[a]Means are rounded to the nearest whole number.

Table 3

Results of the 2 x 2 ANOVA for Examinations Two and Three

Source	SS	df	MS	F
Anxiety	1752.34	1	1752.34	4.83*
Scoring	1074.62	1	1074.62	2.96
Anxiety x Scoring	3130.63	1	3130.63	8.64**
Within	10151.64	28	362.54	
Total	16108.83	31		

*$p < .05$

**$p < .025$

Table 4

Results of the 2 x 2 ANOVA for Examinations Five and Six

Source	SS	df	MS	F
Anxiety	2676.36	1	2676.36	4.51*
Scoring	77.79	1	77.79	.13
Anxiety x Scoring	3760.56	1	3760.56	6.32**
Within	16667.03	28	595.25	
Total	23181.74	31		

*$p < .05$

**$p < .025$

Table 5

KR-20 Results by Type of Scoring for All Examinations

Exam	Class 1		Class 2	
	Binary	Rank Order	Binary	Rank Order
1	.66	-	-	.65
2	.67	-	-	.69
3	.67	-	-	.75
4	-	.65	.68	-
5	-	.71	.69	-
6	-	.77	.71	-

Table 6

Analysis of Students' Evaluation of Rank Order Scoring

Question	Examinees		
	LTA	HTA	ALL
1. Do you prefer ROS to BS? **	41% (7)	87% (13)	63% (20)
2. Did you always assign a "3" to the alternative you would have chosen if BS were being used?	70% (12)	100% (15)	84% (27)
3. How much do you think ROS:			
A. improved your final grade? *	4.1	6.3	5.1
B. improved your ability to study for examinations?	3.8	4.9	4.3
C. increased your enjoyment of lectures?	3.8	4.9	4.3
D. decreased your anxiety over taking tests?	4.0	4.8	4.4
4. Would you like other classes to use ROS on multiple choice examinations? **	41% (7)	87% (13)	63% (20)
5. How much did you like ROS? ***	3.8	6.2	4.9

Note. Questions 1, 2, and 4 were answered yes or no. The percentage answering yes is listed, and the number in parenthesis indicates the number answering yes. All other questions were answered using a Likert format and the mean Likert response is listed.

*$p < .05$

**$p < .02$

***$p < .01$

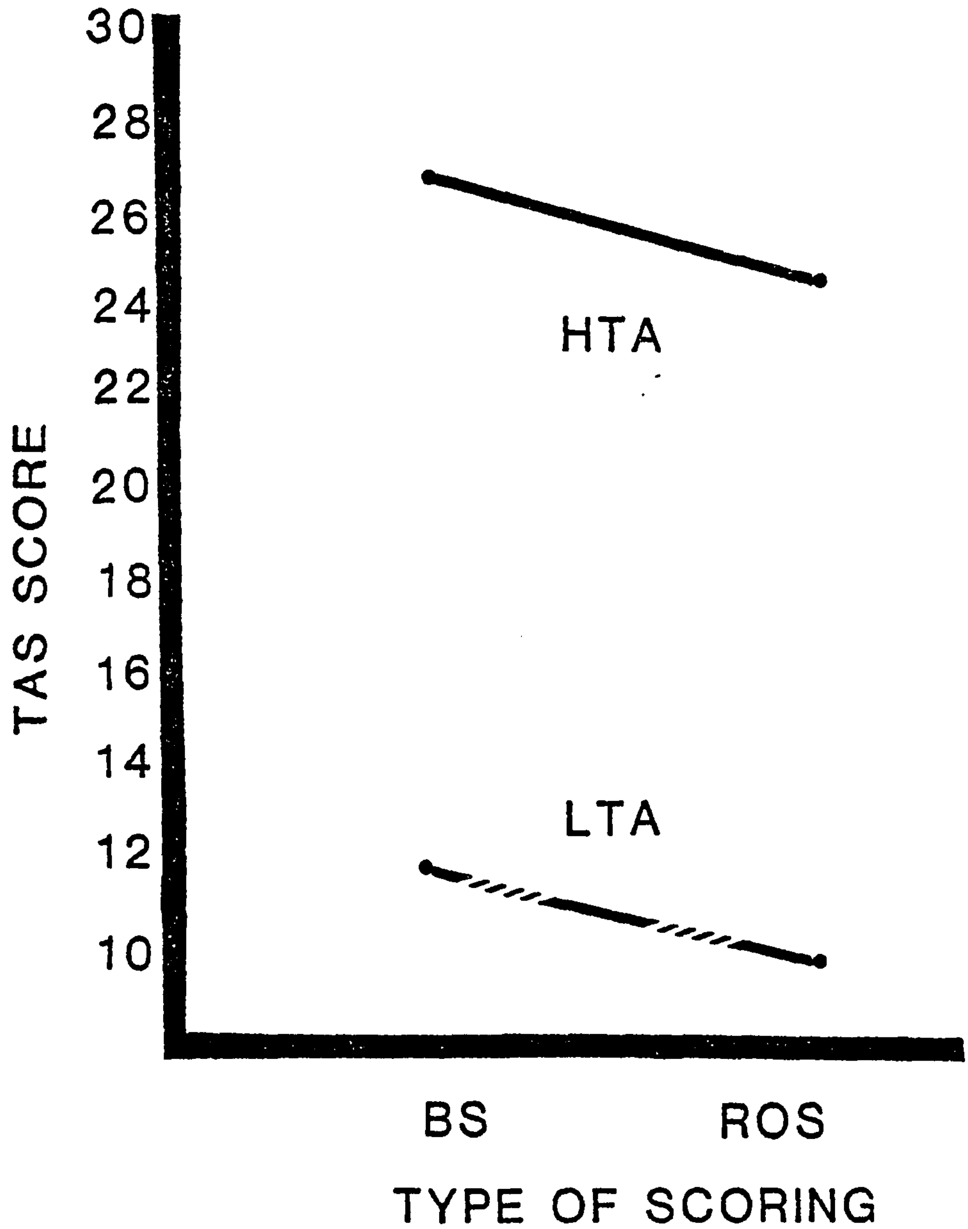

Figure 1. Mean TAS scores for LTA and HTA examinees when using either BS or ROS.

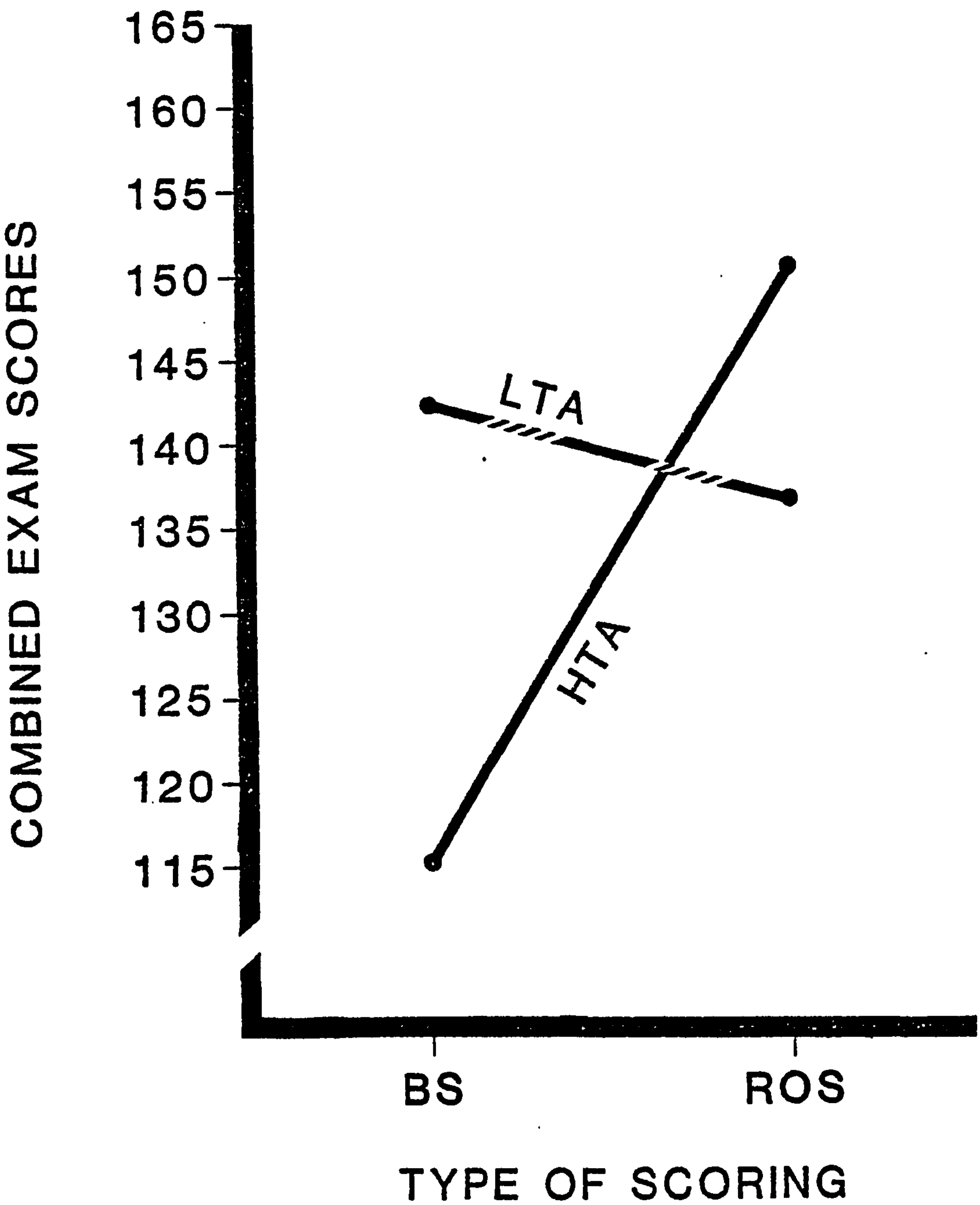

Figure 2. Interaction effect for ANOVA results of examinations two and three.

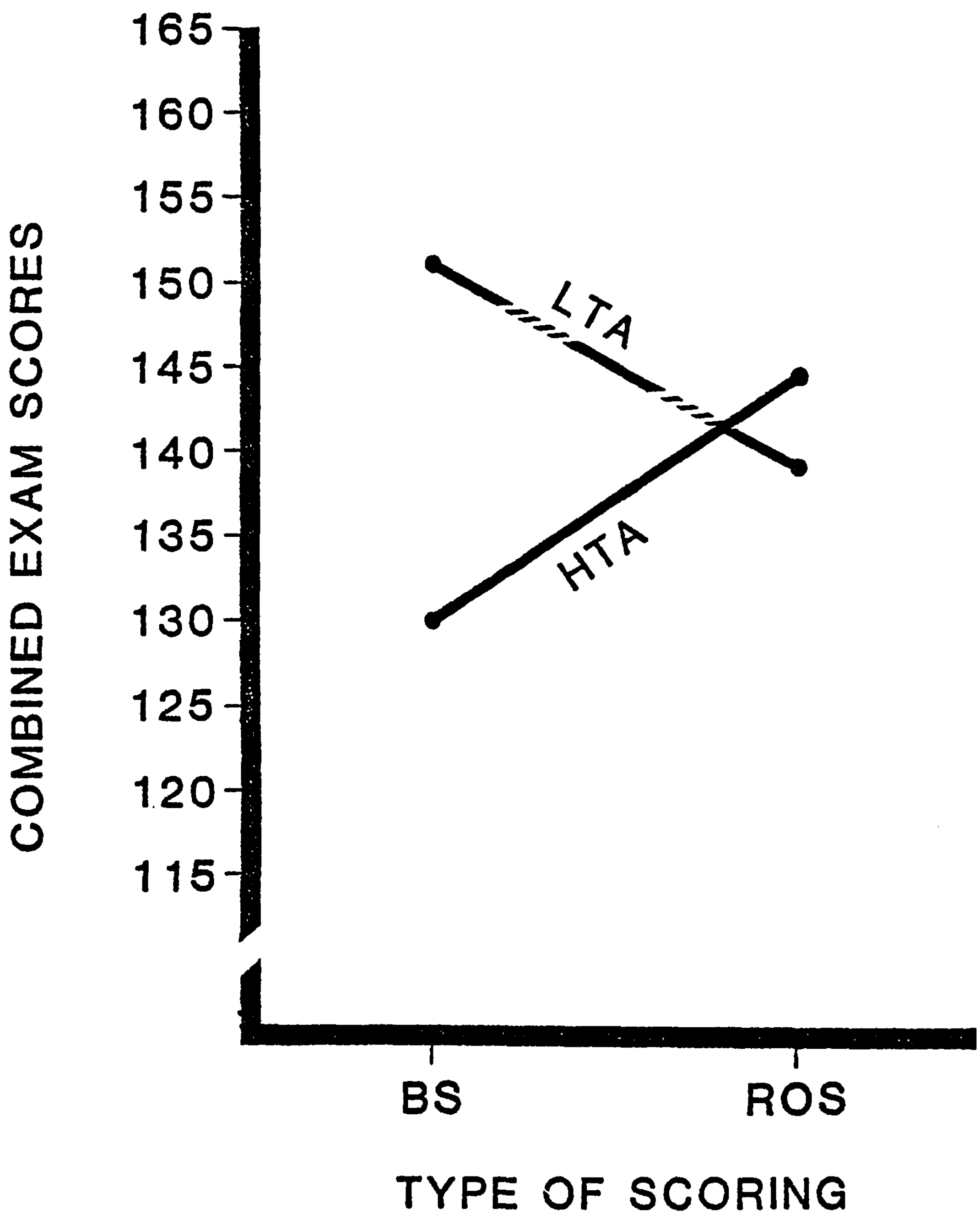

Figure 3. Interaction effect for ANOVA results of examinations five and six.

Appendix A

Test Plan for Determining Complexity of Examinations Items

	Number of Items Tested						Total Items Tested
	Exam #1	Exam #2	Exam #3	Exam #4	Exam #5	Exam #6	
Knowledge							
-Terminology	7	5	5	9	8	4	38
-Specific facts	10	9	7	8	12	9	55
-Classifications and Categories	4	4	1	3	4	5	17
-Methodology	4	4	4	4	-	3	15
-Principles and generalizations	-	2	-	2	-	-	4
-Theories and structures	2	-	4	3	1	1	11
Total knowledge items	27	24	21	29	25	22	148
Comprehension							
-Translation	-	-	-	2	-	2	4
-Interpretation	6	7	8	5	7	6	39
-Extrapolation	7	7	7	4	8	7	40
Total comprehension items	13	14	15	11	15	15	83
Application							
Total application items	-	2	4	-	-	3	9

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

EFFECTS OF PERSONAL PROBABILITY SCORING ON TEST ANXIETY

A DISSERTATION PROPOSAL

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

ROBERT KEITH CONKRIGHT

1980

The Effects of Personal Probability Scoring on Test Anxiety

APPROVED BY

DISSERTATION COMMITTEE

TABLE OF CONTENTS

Abstract

The proposed study will investigate the effects of a personal probability scoring system on test anxiety. Rank order scoring will be compared to binary scoring. Students enrolled in two introductory psychology classes will serve as examinees and will be classified as high test anxious or low test anxious on the basis of Test Anxiety Scale results. The first class will use rank order scoring on the first three examinations and binary scoring on the last three examinations. The second class will use binary scoring on the first three examinations and rank order scoring on the last three examination. The Test Anxiety Scale will be administered before the first and fourth examinations and after the sixth examination. Three 2 x 2 ANOVAs will be used to determine whether rank order scoring reduced anxiety for high test anxious examinees and improved their performance. It is expected that high test anxious examinees will report a decrease in test anxiety and show improved test performance when using rank order scoring.

The Effects of Personal Probability Scoring on Test Anxiety

Multiple choice tests are common in American society, and the results of these tests often affect an individual's life style, job, or educational accomplishments. Multiple choice tests usually contain incomplete statements, or items, that are followed by lists of statements called alternatives. The alternatives for an item usually contain one statement that most accurately completes the item, and examinees must select the alternative that best completes each item. Each item is judged as correct or incorrect, and a scoring system is used to assign a grade.

Psychologists and measurement experts have attempted to determine what factors, excluding ability, affect performance on multiple choice tests. One factor that affects performance is the amount of anxiety a person experiences before, during, and after testing. It has been shown that high test anxious (HTA) individuals typically score lower on tests than low test anxious (LTA) individuals even if the levels of ability of both are equivalent (Wine, 1971). Many theories have been proposed to explain the effects of anxiety on performance.

One theory focuses on examinees' information processing (Tobias, 1976). Tobias suggested that reduced anxiety leads to increased performance because of an uninterrupted flow of information when examinees are processing information.

Another factor that has been shown to affect performance on multiple choice tests is the scoring system applied to assign a final grade (Brunza, 1973). A binary scoring (BS) system is the most commonly used system to assign a grade to multiple choice tests (Littlefield, 1978). The BS system assigns a score of one to each correctly answered item and a score of zero to each incorrectly answered item. With this system, examinees are assumed to have either complete knowledge or zero knowledge of each item.

Lord and Novick (1968) indicated that the binary scoring system has both advantages and disadvantages. The advantages include the small scoring costs and the administration of a large number of items in a fixed period. The disadvantage is that the scorer cannot distinguish between complete knowledge, partial knowledge, guessing, and misinformation.

De Finetti (1965) described several scoring systems that allow examinees to demonstrate complete or partial knowledge of a correct answer. These partial knowledge scoring systems are based on the theory of personal probability, which states that examinees assign a subjective probability of correctness to each alternative and the alternative with the highest probability of correctness is selected as the correct answer.

Brunza (1973) suggested that using a personal probability scoring system reduces anxiety and improves performance of HTA individuals. Brunza used the five-star personal probability scoring system described by de Finetti. Although the five-star scoring system reduced anxiety and improved performance, it was found less reliable and valid than the binary scoring system.

Another scoring system described by de Finetti was the rank order scoring (ROS) system. It was reported that ROS may be as reliable and valid as the BS system (Black, 1974), but results indicating a relationship between ROS and test anxiety have not been reported.

Review of Literature

Mandler and Sarason (1952) developed the Test Anxiety Questionnaire (TAQ) to measure people's anxiety associated with taking a test. The TAQ was a 37-item Likert-type questionnaire asking examinees to rate the amount of anxiety experienced in situations described by each item. The items referred to responses that examinees typically reported experiencing before, during, and after examinations (Wine, 1971).

Mandler and Sarason (1952) used the TAQ to classify examinees as either HTA or LTA. After all examinees were tested, it was concluded that the academic performance of HTA examinees was significantly lower than that of LTA examinees. Further, when examinees were administered tests under stressful, ego-involving conditions, they scored lower than HTA examinees administered the same test under conditions less stressful and ego-involving. The decrease in HTA examinees' performance in stressful conditions was believed the result of a learned response drive. It was postulated that testing situations evoke two types of learned drives: (1) a task drive that stimulates responses relevant to the testing situation and results in improved performance, and (2) a general anxiety drive leading to decreased performance because it stimulates responses both relevant and irrelevant to the tesing situation. It was therefore hypothesized that HTA examinees have a greatly enhanced general anxiety drive and that irrelevant responses aroused by the general anxiety drive lead to a decreased performance (Mandler & Sarason, 1952; Sarason & Mandler, 1952).

Sarason (1958) concluded that Mandler and Sarason failed to account for attentional factors in HTA examinees and proposed an attentional theory of test anxiety. Sarason hypothesized that HTA examinees focus their attention on internal, self-orienting responses.

> This interpretation, briefly put, states that Ss scoring high and low in anxiety differ in the response tendencies activated by personally threatening conditions. Whereas low scoring Ss may react to such conditions with an increased effort and attention to the task at hand, high scoring Ss respond to the threat with self-orienting, personalized responses. (Sarason, 1960, p. 405)

In 1972, Sarason elaborated that worry over the threat of evaluation causes HTA persons to emit self-orienting responses and search the environment for respondable cues. The search of the environment would divert attention from the task at hand and cause decreased performance.

Sarason (1958) developed the Test Anxiety Scale (TAS) to study his attentional theory of test anxiety. The TAS originally consisted of 21 items taken from the TAQ and was later expanded to include all 37 TAQ items with a true-false format instead of the Likert scale of the TAQ. Results of the TAS have been highly correlated (r = .93) with the results of the TAQ (Wine, 1971).

Spielberger (1970) concluded that both Mandler and Sarason (1952) and Sarason (1958) failed to consider personality characteristics of examinees and therefore developed the state-trait theory of test anxiety. Spielberger described two components of test anxiety: (1) a personality characteristic and (2) a situational factor.

Spielberger (1972) defined the situational factor as state anxiety (A-state), a "transitory emotional state of the human organism that varies in intensity and fluctuates over time..." (p.36). Trait anxiety (A-trait) is the personality characteristic Spielberger described as "a relatively stable individual difference in anxiety proneness, that is, to differences in disposition to perceive a wide range of stimulus situations as dangerous or threatening..." (p. 39).

According to the state-trait theory: (1) the perceived stress of the evaluative situation evokes both emotional and task irrelevant responses, (2) the evaluative situation is sufficiently stressful to produce differential A-state reactions in persons with different levels of A-trait, and (3) A-state (emotional reaction) exerts a greater influence on test performance than A-trait (worry). Spielberger, O'Neil, and Hanson (1972) cited research supporting the three hypotheses of the state-trait theory.

Spielberger (1975) attributed the decreased performance of HTA examinees to the task irrelevant responses associated with elevated levels of A-state. The task irrelevant responses are elicited by intrinsic characteristics of the task at hand and make HTA examinees more prone to errors.

Wine (1971) suggested that the previously described theories of test anxiety fail to account for cognitive activities (i.e., thought and attentional processes of examinees) and proposed an attentional theory of test anxiety.

Wine's theory is similar to Sarason's (1958) except that it places greater emphasis on attention to distractive, cognitive processes of the HTA examinees. Wine stated that HTA examinees divide their attention between characteristics of the task and task-irrelevant cognitive activities. Worry is therefore debilitating because it requires attention and distracts from the task at hand. "The highly test-anxious person responds to evaluative testing conditions with ruminative, self-evaluative worry and, thus, cannot direct adequate attention to task-relevant variables" (Wine, 1971, p. 99).

Tobias (1976) suggested that Wine's theory places too much emphasis on attentional responses of examinees and not enough emphasis on information processing. Tobias described an adaptation of an information processing model to clarify the interactive effects of different instructional methods and test anxiety. The model divided the flow of information into three stages: input, processing, and output.

Input, in the Tobias model, refers to instructional methods or content. This is the way information is organized and presented to students.

The processing stage refers to cognitive operations performed by the examinees on input information. Operations in the processing stage include attending, registering, storing, coding, deducing, inducing, and any other operation on input information which renders the information more accessible to retrieval from memory.

The output stage refers to behaviors used to determine whether or not instructional objectives have been met. This stage is usually considered a dependent variable or the criterion used to evaluate the effects of test anxiety on performance (e.g., test results).

Tobias (1977) suggested that anxiety may affect three points in the model to decrease performance. The first point is preprocessing, between the input and processing stages. Anxiety may reduce or restrict the flow of input to the processing stage and may result in examinees' failure to attend to or represent information. Oosthoek and Ackers (1972) used HTA and LTA examinees to study the effect of conventional and taped lectures on anxiety during preprocessing.

Examinees in a conventional lecture section heard lectures only once, and examinees in a taped lecture section heard lectures as often as desired. The HTA examinees in the taped lecture section performed significantly better than HTA examinees in the conventional lecture section. No significant difference in performance was noted between LTA examinees in the taped lecture section and LTA examinees in the conventional lecture section. Tobias (1977) suggested that examinees were able to reinstitute information on the taped lecture often enough to compensate for anxiety's interference with preprocessing.

The second point in Tobias's model that may be affected by anxiety is the processing stage. Anxiety may interfere with memory functions used to process information (Tobias, 1977). Sieber (1970) used a multi-stage task and allowed examinees to review portions of a previously completed task (i.e., "memory supports"). Sieber reported that "memory supports" significantly improved HTA examinees' performance, but did not significantly affect LTA examinees' performance. Tobias (1977) suggested that "memory supports" compensate for anxiety's effect on processing by allowing examinees to review previously learned material without relying on memory functions.

A third point in Tobias's model that may be affected by anxiety is postprocessing, between the processing and output stages (Tobias, 1977). Anxiety may affect postprocessing by causing examinees to to lose information before they verbalize or record it. Tobias cited phenomenological evidence to support this hypothesis.

> It seems possible that anxiety may interfere with the accurate rendering of that response in output. Phenomenologically, we are all familiar with occasions in which the solution of a problem has been grasped and then lost. For example, one may hit upon an intuitive answer to an arithmetic problem, or the sudden identification of a melody, quotation or poem, which is inexplicably 'forgotten' before one is able to verbalize of record it. (Tobias, 1977, p. 227)

Although one may perceive three separate points of interference in Tobias's model, the three points are probably not independent. The effects of anxiety may also be cumulative. If information is lost in preprocessing, examinees must rely on the processing stage to determine the lost information, and the efficiency of processing is therefore reduced. Interference during postprocessing also increases the loss of information (Tobias, 1977).

Tobias's model suggested that reduced anxiety leads to increased performance by allowing an uninterrupted flow of information to preprocessing, processing, and postprocessing. Allen (1972) and Wine (1971) reviewed research indicating that reduced anxiety leads to improved performance for HTA examinees.

Data indicate that the type of system used to score an examination may affect anxiety and performance (Brunza, 1973). Brunza used a personal probability scoring system to score multiple choice tests, and significantly reduced HTA examinees' anxiety. The HTA examinees also demonstrated significantly improved performance. Brunza attributed the reduced anxiety and improved performance to "conceptual differences" between the binary scoring and personal probability scoring systems. The personal probability scoring system makes the testing situation less threatening.

De Finetti (1965) described four classes of personal probability scoring procedures: (1) rank order, (2) flexible schemes, (3) five-star scheme with fixed probability, and (4) flexible five-star scheme. These scoring procedures are used to test for partial knowledge. When rank order scoring procedures are used examinees are required to rank all alternatives with respect to their personal probability of correctness (e.g., examinees assign the number "4" to the answer most likely correct, the number "3" to the second most likely answer, "2" to the third most likely answer, and "1" to the answer least likely correct). When using rank order scoring, the examinees' strategy should be to assign the highest rank to the alternative with the highest probability of correctness (de Finetti, 1965).

Compared to binary scoring, rank order scoring allows the scorer to better distinguish between complete and different levels of partial knowledge. The rank order and binary scoring systems do not allow the scorer to estimate examinees' underlying personal probabilities.

Flexible schemes and five-star schemes also allow the scorer to better distinguish between complete and different levels of partial knowledge. Unlike the rank order procedures, they allow the scorer to estimate examinees' personal probabilities. Both procedures have been criticized as being too complex for examinees to fully understand (Diamond, 1975; Littlefield, 1978; Lord & Novick, 1968).

Unlike de Finetti's other procedures, examinees indicated that rank order scoring is easy to understand (Conkright & Williams, 1975). Examinees also indicated that rank order scoring made examinations less stressful, class lectures more enjoyable, and improved their ability to comprehend material presented. Although Conkright and Williams (1975) did not report reliability and validity data in their experiment, Black (1974) reported reliability and validity coefficients for rank order scored items that were equal to or greater than those of binary scored items.

Statement of Problem

At some point between the input and output of information, anxiety can interfere with information processing and result in poor test performance. If examinees perform poorly because of anxiety experienced when processing information, then decreasing anxiety should improve examinees' performance.

Research indicates that the type of system used to score an examination can significantly improve HTA examinees' performance. Unlike the traditional binary scoring system, personal probability scoring systems have effectively improved HTA examinees' performance. Brunza used a five star personal probability scoring system that was difficult for examinees to use, yet HTA examinees demonstrated significantly improved scores when using the system. Brunza attributed the improvement to decreased anxiety.

The intention of this investigation is to determine the effect of rank order scoring, an easy-to-use personal probability scoring system, on HTA examinees' anxiety and performance. If rank order scoring successfully reduces HTA examinees' anxiety, then (1) their performance should improve, (2) they should prefer rank order scoring to binary scoring, and (3) they should assign the highest possible rank to alternatives they would have chosen correct if using binary scoring.

Method

Examinees

Students enrolled in two introductory psychology courses at an urban community college will serve as examinees. The examinees in one course will be referred to as Class 1 and the examinees in the other course will be referred to as Class 2. Examinees will be administered the TAS (Appendix A) and results will be used to classify examinees as either HTA or LTA. Brunza's procedures for classifying HTA and LTA examinees will be used (Brunza, 1973). Examinees scoring 20 points or more will be classified as HTA, and examinees scoring 17 points or less will be classified as LTA.

Procedure

One instructor will present the same information in lectures to both classes. Lecture outlines will be distributed before each lecture and all information in both the outlines and lectures will be taken from a designated text for the course. When an examinee is absent from a lecture, he or she will be given an outline for the missed class at the next class meeting.

Six examinations will be administered during the semester, and each examination will consist of 40 multiple choice items with four alternatives per item. The correct answer for each item will be discussed in the lectures and provided in the outlines. To determine the complexity of each item, a test plan will be constructed for each examination (Bloom, 1956).

Class 1 will use rank order scoring on the first three examinations and binary scoring on the last three examinations. Class 2 will use binary scoring on the first three examinations and rank order scoring on the last three examinations. A norm-referenced testing approach will be used. Examinations one and four will be considered trial runs, and the scores will not be used in the analysis of results.

Testing instructions. Examinees in both classes will be intructed orally and given a written example of the correct way to use their respective scoring systems at six points: (1) during the first class, (2) during the lecture preceeding the first examination,

(3) immediately before the first examination, (4) during the lecture immediately after the third examination, (5) during the lecture preceeding the fourth examination, and (6) immediately before the fourth examination. On the days of the second, third, fifth, and sixth examinations, oral instructions will be given to the examinees.

Examinees using binary scoring will be instructed to mark the space that corresponds with the alternative they think is correct. Examinees using rank order scoring will be instructed to write a "3" in the space that corresponds with the alternative they think is most correct, a "2" in the space for their second choice, a "1" in the space for their third choice, and to leave the space blank for their fourth choice. Examinees will be told that the rank order scored tests will be scored by totaling only the points they assign to the correct alternatives and converting that total to a percentage of the correct score. The binary scored tests will also be converted to a percentage of the correct score.

Administering the TAS. The TAS will be administered four times during the semester: (1) during the first lecture in order to classify examinees as HTA or LTA, (2) immediately before the first examination, (3) immediately after the third examination, and (4) immediately after the sixth examination.

Examinees will be told that the TAS is designed to measure their anxiety level before an examination and that TAS scores will not effect their examination scores or final grades.

After completing the fourth TAS, each examinee will be asked to complete a nine-item Likert scale evaluation (Appendix B) designed to determine their preference for binary scoring or rank order scoring. The evaluation will be designed to determine whether examinees think rank order scoring improved their course grades, their ability to comprehend course information, their ability to study for examinations, and their enjoyment of lectures. Examinees will be asked to indicate whether rank order scoring reduced their test anxiety and whether they always assigned a "3" to the alternative they would have chosen as correct if binary scoring had been used.

Data Management

Results of the TAS administrations will be analyzed to test the hypothesis that rank order scoring will reduce HTA examinees' anxiety. A 2 (HTA/LTA) x 2 (type of scoring) ANOVA will be used to analyze the data. The TAS scores for HTA and LTA examinees using rank order scoring will be used as data as will the TAS scores for HTA and LTA examinees using binary scoring. Tukey's Honest Significant Difference (HSD) method of multiple comparison will be used to assist in interpreting significant ANOVA results (Hopkins & Glass, 1978). It is expected that the anxiety level main effect will reveal a significant effect. The significant effect should indicate that anxiety decreased for HTA examinees when using rank order scoring and increased when using binary scoring.

The hypothesis that rank order scoring improves HTA examinees' performance will be tested with the aid of two factorial

2 (HTA/LTA) x 2 (type of scoring) ANOVAs. The first ANOVA will analyze the combined results of the second and third examinations. The second ANOVA will analyze the combined results of the fifth and sixth examinations. The first and fourth examinations will be considered trial runs and will not be analyzed. Tukey's HSD will be used to assist in interpreting significant ANOVA results. It is expected that significant anxiety level main effects and the anxiety level by type of scoring interaction will result. Significant results should indicate that HTA examinees improve their performance when using rank order scoring and decrease their performance when using binary scoring.

A transformation procedure will be used to make the expected values of rank order and binary scored examinations equal. The transformation procedure to be used will be the "number-right" procedure used by Diamond (1975). All rank order scored examinations will be rescored using a binary scoring examination. Each answer with a "3" assigned to it will be assumed to be the answer that would have been chosen as correct using the binary scoring system. The answers with a "3" assigned to the correct answer will be counted and the total will be used as the "number-right" score.

The Kuder-Richardson formula 20 (K-R 20) will be used to determine whether rank order scoring is as internally consistent as the binary scoring system.

Results of the nine-item Likert scale evaluations will be presented in tabular form. Tables of the results will be constructed for HTA examinees, LTA examinees, and all examinees combined. Results of the evluations are expected to indicate that 1) examinees preferred rank order scoring to binary scoring, 2) rank order scoring increased examinees' enjoyment of lecture, 3) rank order scoring improved examinees' ability to comprehend course material, 5) examinees obtained higher course grades, 6) test anxiety decreased, and 7) examinees always assign a "3" to the alternative they would have chosen as correct when using binary scoring. If additional analysis of the evaluation should be necessary, a one-way ANOVA will be used (Hopkins & Glass, 1978).

References

Allen, G. J. The behavioral treatment of test anxiety: Recent research and future trends. Behavior Therapy, 1972, 3, 243-262.

Black, R. H. A method of discriminating partial knowledge (Doctoral dissertation, University of Oklahoma, 1974). Dissertation Abstracts International, 1975, 36, 181A, (University Microfilms No. 75-15, 246).

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

Brunza, J. J. The effects of repeated probabilistic testing conditions on state anxiety and test reliability (Doctoral dissertation, Indiana University, 1973). Dissertation Abstracts International. 1974, 34, 5706A, (University Microfilms No. 74-7001).

Conkright, R. K., & Williams, J. D. A comparison of binary and rank order scoring methods: Quantitative and affective data. Paper presented at the meeting of the Southwestern Psychological Association, Houston, May 1975.

de Finetti, B. Methods of discriminating levels of partial knowledge concerning a test item. British Journal of Mathematical and Statistical Psychology, 1965, 19, 87-123.

Diamond, J. J. A preliminary study of the reliability and validity of a scoring procedure based upon confidence and partial information. Journal of Education Measurement, 1975, 12. 129-133.

Hopkins, K. D., & Glass, G. V. Basic statistics for the behavioral sciences. Englewood Cliffs, New Jersey: Prentice-Hall, 1978.

Littlefield, B. The effects of corrected for guessing and partial information scoring procedures and number of alternatives on reliability, validity, and difficulty in multiple choice test items (Doctoral dissertation, University of Maryland, 1978). *Dissertation Abstracts International*, 1979, *39*, 3693A, (University Microfilms No. 79-00916).

Lord, F. M., & Novick, M. F. *Statistical theories of mental test scores*. Reading, Massachusetts: Addison Wesley, 1968.

Mandler, G., & Sarason, S. A study of anxiety and learning. *Journal of Abnormal and Social Psychology*, 1952, *47*, 166-173.

Oosthoek, H., & Ackers, G. The evaluation of an audio-taped course (II). *British Journal of Educational Technology*, 1973, *4*, 55-73.

Sarason, I. G. Effects of verbal learning, reassurance, and meaningfulness of material. *Journal of Experimental Psychology*, 1958, *56*, 472-477.

Sarason, I. G. Empirical findings and theoretical problems in the use of anxiety scales. *Psychological Bulletin*, 1960, *57*, 403-413.

Sarason, S., & Mandler, G. Some correlates of test anxiety. *Journal of Abnormal and Social Psychology*, 1952, *47*, 810-817.

Sieber, J. E. The effects of memory support of the problem-solving ability of test anxious children. *Journal of Educational Research*, 1970, *61*(2), 159-168.

Spielberger, C. D. *State-trait anxiety inventory* (STAI), 1970.

Spielberger, C. D. *Anxiety: Current trends in theory and research* (2 Vols.). New York: Academic Press, 1972.

Spielberger, C. D. Anxiety: State-trait process. In C. D. Spielberger & I. Sarason (Eds.), Stress and anxiety (Vol. 1). Washington, D. C.: Hemisphere, 1975.

Spielberger, C. D., O'Neil, H. F., & Hansen. D. N. Anxiety drive theory and computer-assisted learning. In B. A. Maher (Ed.), Progress in experimental personality research. New York: Academic Press, 1977.

Tobias, S. A model for research on the effects of anxiety of instruction. In J. E. Sieber (Ed.), Anxiety, learning and instruction. New York: Halstead Press, 1977.

Tobias, S. Achievement treatment interactions. Review of Educational Research, 1976, 46, 61-74.

Wine, J. D. Test anxiety and direction of attention. Psychological Bulletin, 1971, 76, 92-104.

Appendix A

TEST ANXIETY SCALE

INSTRUCTIONS: Mark each statement true or false.

_____1. While taking an important exam, I find myself thinking of how much brighter the other students are than I am.

_____2. If I were to take an intelligence test I would worry a great deal before taking it.

_____3. While taking an important exam I perspire a great deal.

_____4. If I knew I was going to take an intelligence test, I would feel confident and relaxed before hand.

_____5. During course exams I find myself thinking of things unrelated to the actual course material.

_____6. I feel very panicky when I take a surprise exam.

_____7. During tests I find myself thinking of the consequences of failing.

_____8. After an important exam, I am frequently so tense that my stomach gets upset.

_____9. I freeze up on things like intelligence tests and course exams.

_____10. Getting a good grade on one test doesn't seem to increase my confidence on the second.

_____11. I sometimes feel my heart beating very fast during an exam.

_____12. After taking a test I always feel that I could have done better than I actually did.

_____13. I usually get depressed after taking a test.

_____14. I have an uneasy, upset feeling before taking a final exam.

_____15. When taking a test my emotional feelings do not interfere with my performance.

_____16. During exams I frequently get so nervous that I forget facts I really know.

_____17. I seem to defeat myself while working on important tests.

_____18. As soon as an exam is over I try to stop worrying about it but I just can't.

_____19. The harder I work at taking a test or studying for one, the more confused I get.

_____20. During an exam, I sometimes wonder if I will ever get through college.

_____21. I would rather write a paper than take an exam for my grade.

_____22. I wish examinations didn't bother me.

_____23. I think I could do better on tests if I could take them alone and not be bothered by time limits.

_____24. Thinking about the grade I may get in a course interferes with my performance on tests.

_____25. If exams could be done away with, I think I would actually learn more.

_____26. On exams I take the attitude "If I don't know it now, there's no point in worrying about it."

_____27. I really don't see why some people get so upset about exams.

____28. Thoughts of doing poorly interfere with my performance on tests.

____29. I don't study any harder for final exams than I do for the rest of my course work.

____30. Even when I am well prepared for a test, I feel anxious about it.

____31. I don't enjoy eating before taking an important exam.

____32. Before an important exam I feel my hands and arms trembling.

____33. I seldom feel the need for "cramming" before an exam.

____34. The university ought to recognize that some students are more nervous than others about tests and that this affects their performance.

____35. It seems to me that exam periods ought not be made the tense situations which they are.

____36. I start feeling very uneasy just before getting a test papaer back.

____37. I dread courses where the professor has the habit of giving pop quizzes.

Appendix B

EVALUATION OF SCORING PROCEDURES

1) Of the two scoring procedures used this semester, which one did you like the best?

   A) Traditional (circle the correct answer)

   B) Rank Order (rank the alternatives)

2) When using the rank order scoring procedure, did you always assign a 3 to the answer you would have circled if the traditional scoring procedure were being used? YES NO

3) Do you think that the rank order scoring procedure:

   A) improved your grade in the course?

1	2	3	4	5	6	7
greatly lowered grade			grade same			greatly improved grade

   B) improved your ability to study for tests?

1	2	3	4	5	6	7
harder to study			same			easier to study

   C) improved your knowledge of the course material?

1	2	3	4	5	6	7
decreased knowledge			same			increased knowledge

D) increased your enjoyment of the class lectures?

1 2 3 4 5 6 7

decreased same increased

enjoyment enjoyment

E) Decreased your anxiety over taking tests?

1 2 3 4 5 6 7

decreased same increased

anxiety anxiety

4) Would you like other coursed to use the rank order scoring procedure?

________YES ________NO

5) How much did you like the rank order scoring procedure?

1 2 3 4 5 6 7

did not same as liked it

like it traditional very much

Oral Instructions for ROS

When you take this test, use the rank order scoring system that we have discussed. Put a "3" beside the alternative you think is most correct. Put a "2" beside the second best answer, a "1" beside the third best answer, and nothing beside the alternative you think is least likely to be correct. Do you remember the examples we talked about earlier in the semester? Does anyone have any questions? Do you understand how to use the rank order scoring system? Remember, write a "3" beside the best answer, a "2" beside the second best answer, a "1" beside the third best answer, and don't write anything beside the answer you think is least likely to be correct.

Written Instructions for ROS

On three of the tests you take in this course, you will be using a new scoring system. Using this new scoring system will require you to mark your tests differently. Instead of marking only one correct answer per item as you do when using the traditional scoring system, you are to RANK each alternative in terms of its correctness. Beside the alternative you think is most correct, write a "3". Write a "2" beside the alternative you think is the second best answer, write a "1" beside the third best answer, and don't write anything beside the alternative you think is least likely to be correct. Here is an example:

The Governor of Texas is:

______ a) a politician

______ b) the chief executive of the state

______ c) a great guy

______ d) a bad guy

The best answer is alternative B. You should write a "3" beside B. The second best answer is A and you should place a "2" beside it. The third best answer is C and you should write a "1" beside it. The space next to alternative D should be left blank because it is the answer least likely to be correct.

Here is another example for you to work:

The Dallas Cowboys and the Houston Oilers are:

______ a) a bunch of big men

______ b) a bunch of little people

______ c) two professional football teams

______ d) two likely Super Bowl participants

Your examination score will be the total number of points you assigned to the correct alternative for each item. If you put a "3" beside the correct alternative, you will get three points for that item. If you put a "2" beside the correct alternative, you will get 2 points for that item. If you put a "1" beside the correct alternative, you will get one point for that item. If you don't write anything beside the correct alternative, you get no points for that item.

If you have questions about this scoring system, please ask them in class or talk to me after class.

Vitae

Robert Keith Conkright was born on January 28, 1949 in Minden, Louisana. His parents are Emogene Rogers Conkright and William Franklin Conkright, and his brother is William Freeman Conkright. Robert moved to Houston in 1958 and graduated from Robert E. Lee High School in Houston on May 30, 1967. He attended Stephen F. Austin State University in Nacogdoches, Texas and was graduated with B.S. and M.A. degrees in 1971 and 1973 respectively. He enrolled in The University of Oklahoma in 1973. Robert lives with his wife, Barbara Lynn Keir, in Austin, Texas.

Permanent address:

P.O. Box 125
Helotes, Texas 78023