INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE EFFECT OF STUDENT KNOWLEDGE OF METHOD OF TEST SCORING UPON ACHIEVEMENT TEST PERFORMANCE

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

JOHN WALTER OTEY

Norman, Oklahoma

1975

THE EFFECT OF STUDENT KNOWLEDGE OF METHOD OF TEST SCORING UPON ACHIEVEMENT TEST PERFORMANCE

APPROVED BY

DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

Without the support of many people this dissertation would not have been possible. The author wishes to especially thank Dr. Omer J. Rupiper, major professor and committee chairman, who made many valuable suggestions and provided much encouragement throughout the preparation of this dissertation. Special thanks are extended to Dr. Barbara Nelson for the use of her achievement instrument for collection of data, to Dr. William Graves for his advisement and assistance in the design and statistical analysis and to Dr. Al Smouse for his advice in the design of the study. Appreciation is gratefully extended to all committee members for their time and assistance in collecting the data.

Finally, debts of gratitude are extended to my good friend, Alan Todd, who assisted me in my research and my wife, Christine, who was brave enough to face a dissertation after having been through the author's master's thesis with him.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Statement of the Problem

In virtually any class, moans of discontent and comments of a negative nature are heard when tests are presented. It seems to make little difference whether one is in an early elementary school classroom, a high school class, or a university class. The attitude seems to be about the same since most students appear to not enjoy the task of taking tests. Many students display and express feelings of fear, hostility, and anxiety, especially when it is announced that the responses to the test items will be recorded on separate answer sheets and computer-scored. The experience of this writer has shown that fewer visible signs of emotion and fewer negative expressions are evident when the announcement is made that the responses to the test items will be recorded on the test itself and hand-scored.

The question arises as to why computer-scored achievement tests are less popular with students than hand-scored tests. The answer may be that students dislike the separate answer sheets or they may not trust computer-scoring. Regardless of the reasons given many students claim they receive

lower scores on computer-scored tests than they do on teacher-scored tests. It is this claim of lower grades obtained on computer-scored tests which this study undertook to investigate. Do students actually receive lower grades on achievement tests which they know are to be computer-scored than they do on conventional hand-scored tests? Are such computer-scored tests more anxiety inducing than hand-scored tests? These are the two main questions which this study sought to answer.

Review of the Literature

Indirectly, Anastasi (1968, pp. 37-38) discussed various factors of the testing situation which may either singly or in combination affect test performance. She stated that we must realize the extent to which testing conditions may give unreliable results. "Even apparently minor aspects of the testing situation may appreciably alter performance." One such "minor" aspect may be the issue of computer-scored answer sheets.

Research studies dealing with the area of separate computer-scored answer sheets, their use, and effect upon performance are relatively sparse. Greater concentration has been seen on topics such as: the arrangement of test items (Smouse & Munz, 1968; Smouse & Munz, 1969), the effect of test directions on performance (Sarason, & Palola, 1960; Sarason, 1961; Anastasi, 1968) answer sheet format (Dunlap, 1940; Bell, Hoff, & Hoyt, 1964; Petty, 1964; Thorndike, 1971), and the effect of anxiety on test taking behavior (Alpert &

Haber, 1960; Mandler & Sarason, 1952; Sarason & Mandler, 1952; Sarason, Mandler, & Craighill, 1952; Waite, Sarason, & Lighthall, 1958) were most apparent. Only the most closely related topic of answer sheet format will be reviewed as it more directly concerns computer-scoring. In light of their widespread use, it is most astounding that relatively little research has been done concerning problems associated with the separate test answer sheet.

The most extensive study of separate test answer sheets was undertaken (Dunlap, 1940) when machine-scoring was just coming into practice. Dunlap noted that there were a number of practical and theoretical questions to be considered. Among those were cost, reliability, validity, time, and separate norms for hand-scored and machine-scored answer sheets. These problems were often over looked when it was shown to teachers how much easier and more rapidly answer sheets could be scored with test-scoring machines. Dunlap's study, a series of five separate experiments, compared mean scores on tests having integral answer sheets with tests having separate answer sheets. Results indicated that clerical errors (errors in transferring answers from the test booklet to the separate answer sheet) were few in number and that mean achievement scores, reliability, and validity coefficients for the separate and integral answer sheets were not significanlty different. Scores were said to be "very nearly the same" when the number of items attempted was the same. Dunlap

recommended that more time be allotted for tests with separate answer sheets. As an alternative to increasing the time limit, he said that a practice session might be used to "warm students up" to the test.

Thorndike (1971) also discussed the topic of separate answer sheets wherein he stated that the task of transferring answers from test booklet to answer sheet may be too demanding for young people and people to whom such tests were "strange." The age at which separate answer sheets may be used successfully and the amount of experience in working with tests of this nature were still not definitely determined. Thorndike recommended that young and inexperienced persons write answers directly in the test booklet, however.

Petty (1964) investigated several forms of separate test answer sheets and sought to determine if the method used to mark answers on separate answer sheets had an influence on achievement test scores. He also surveyed student opinions of four forms of answer sheets commonly used on standardized achievement tests. The forms studied were: pin-punch--the student punches out the desired choice on the answer sheet with a wood stylus, cal-card--an answer sheet printed on a standard computer card, separate self-scoring answer sheet--an answer sheet which consisted of two sheets of paper with a carbon sandwiched between, thus enabling the sheet to be self-scoring, and test booklet--where the student answers questions directly in the test booklet. The subjects, fourth,

sixth, eighth, and eleventh graders, all scored equally well regardless of the answer sheet format. However, they did express a preference for the test booklet format saying that it saved time over the other forms and was easier to use.

Although statistically significant differences in mean achievement scores on integral and separate answer sheets and among various forms of separate answers were not found by Petty, Bell, Hoff, & Hoyt (1964) found that answer sheet format did make a difference in achievement test scores. They conducted a study of two forms of answer sheets for the Government USES General Aptitude Test Battery using Government adult civilian employees as subjects. Both forms of the test studied were machine scorable forms, differing in format. The differences were described as differences in packaging density. One form of the answer sheet was arranged more compactly in order to squeeze more answers on a sheet. Scores were found to inversely correlate with the packaging density or number of items per page. The sheet with the fewest number of items resulted in the highest scores. No reasons for this difference were posited.

The literature was more sparse concerning computerized testing; the few studies completed to date have dealt primarily with computer assisted instruction (CAI) and the various forms of optical scoring machines. Whitlock (1964) gave an overview of data processing. Much of his book was devoted to showing how the school administrator could automate records. Only

brief mention was made with respect to the use of computers in scoring and analyzing test scores. Whitlock commented in generalities and did little more for the reader than to show what a machine-scorable answer sheet looked like and described how it was to be filled out.

Smith (1963) listed the advantages of automated-scoring, such as, speed and ease of scoring answer sheets, increased accuracy resulting from elimination of human errors, and the ability to receive "extensively analyzed output." What the author meant by "extensively analyzed output" was not explained. Overall, this article was general and not specific on points of procedure.

The literature concerning studies dealing with student attitude toward computerized testing was very meager. In an article about computer assisted instruction (CAI), Mathis, Smith, and Hansen (1970) reported parenthetically that college students seemed to have "positive attitudes" toward computers, generally. However, it was not clear from this article whether students initially had favorable attitudes toward computers or if they developed those feelings after having undergone a period of CAI.

A number of intelligence tests have been administered on computers. Elwood (1969) reported on a project to administer and score the WAIS by computer. A savings of several hours in administration, scoring, and forming a diagnosis was cited. For clinical patients results on the performance scale

were said to correlate greater than r=.90 with an examiner administered test. The Slossen Intelligence Test was likewise computerized by Hedl (1971). Scores were again found to be comparable, however, the computerized version did lead to higher levels of state anxiety. Students said they found the computer version more stressful and less enjoyable. They expressed a preference for examiner administered tests over the computerized version.

Summary of the Literature

Generally, there has been very little research in the area of computer induced anxiety in achievement testing situations. The literature dealt briefly with answer sheet format, computer assisted instruction, and computerized tests.

Several of the studies reviewed supported the conclusion that students scored equally well on tests utilizing hand-scorable and machine-scorable answer sheets. Differences in performance attributable to the various formats of machine-scorable answer sheets had likewise, been shown to be minimal. Attitudes of students toward computerized examinations were mentioned in several of the studies reviewed. One article reported parenthetically that students had "positive attitudes" toward computers in CAI situations; while another study reported that students found computerized examinations to be stressful and more anxiety inducing than conventional examiner administered tests. It thus appears that little can be reliably concluded at this time about student attitudes toward

computerized testing. In addition, none of the studies reviewed investigated the possibility of students becoming anxious and, hence, performaing less well on an examination which they knew would later be scored by computer. A number of studies have shown that test anxiety was produced by sequencing of test items, test directions, and answer sheet format. Conclusions were that varying answer sheet format, altering the sequencing of items and differing test directions did indeed have an affect on levels of state (test) anxiety. Whether or not these conclusions could be generalized to machine-scorable answer sheets was not clear. This investigator was not aware of any study having been conducted on machine-scorable answer sheets to determine their effect on students' test anxiety levels and resultant achievement scores. There were no empirical data at this time which indicated whether students were affected adversely by computer-scoring.

Assumptions Underlying the Hypotheses

The framework upon which this study rested was that of Mandler and Sarason (1952) which stated that students' test performance may be adversely affected by the personality variable test anxiety. This particular type of anxiety results in unreliable estimates of one's knowledge when it is present during testing. Although test anxiety may serve to raise one's performance level on an examination (Sweeney, Smouse, Rupiper, & Munz, 1970), it was more often thought of in the

literature as a detriment to testing. This was felt to be the case particularly when test directions were unusual (Sarason, Mandler, & Craighill, 1952). The particular case which the present study sought to investigate was that of computerized test scoring. The information presented in the directions for taking a test, that is, that it is to be computer-scored, is apparently anxiety inducing to some students. As reported earlier in this paper, such anxious students complain of receiving lower grades on this type of test than they would have received on a hand-scored test. It was thus assumed by this investigator that students who showed lower performance on computer-scored tests did so because of test anxiety generated by the students' knowledge of the method used to score the test.

Hypotheses

The present study undertook to investigate the following hypotheses which were tested at the p=.05 level of significance:

Hypothesis 1: The mean raw score on a hand-scored achievement test is significantly greater than the mean raw score on the same achievement test when it is computer-scored if the method of scoring is presented beforehand in the test directions.

Hypothesis 2: When the Test Anxiety Scale (TAS) (Sarason & Ganzer, 1962) is presented immediately following an achievement test the mean raw TAS score is significantly greater for

the subjects taking a computer-scored form of an achievement test than it is for those taking a hand-scored form of the same achievement test if the method of scoring is presented beforehand in the achievement test directions.

Hypothesis 3: Raw achievement test scores and raw Test Anxiety Scale (TAS) scores will show a statistically significant inverse correlation.

Operational Definitions

Test Anxiety: This is a form of state anxiety which arises in situations in which one is being evaluated. For the purposes of this study test anxiety was defined in terms of a score on the Test Anxiety Scale (TAS). Scores on the TAS (see Appendix D) range from a low of 0, indicating no test anxiety, to a high of 16.

Computer-Scored Achievement Test: This was a thirty-four item diagnostic mathematics achievement test (see Appendix E) which was answered on a separate machine-scorable answer sheet.

Hand-Scored Achievement Test: This was a thirty-four item diagnostic mathematics achievement test (see Appendix E) which was answered on a separate machine-scorable answer sheet.

Separate Answer Sheet: This was a machine scorable answer sheet (IBM No. 503, see Appendix F) upon which students were to mark all achievement test answers with a soft lead pencil. The sheets were scored by an optical test scoring machine.

CHAPTER II

DESIGN OF THE STUDY

Introduction

This study was an outgrowth of a pilot study conducted by this experimenter in the Fall of 1973 at the University of Oklahoma. At that time it was decided to verify whether student claims were indeed valid, that is, if students scored lower on computer-scored tests than they did when the tests were hand-scored.

Identical achievement tests were administered to two of the experimenter's sections of undergraduate educational psychology. By means of a tossed coin it was determined which section was to have the computer-scored separate answer sheet and which section was to have the integral answer sheet. Time-of-day effect had been shown on previous examinations to be of negligible consequence. Thus, it was decided that intact classes might be utilized as the experimental groups. The test was a ten question achievement test over class notes and lectures, both sections having received identical lectures and assignments.

The first section was handed the test questions with the instructions that the student was to X -out the desired

choice directly on the test sheet. The second section received the same test sheet in addition to a separate machine scorable answer sheet which was to be computer-scored by the University Test Service. Students were cautioned to use only a number two pencil, and to follow the directions printed on the answer sheet. It was pointed out that ink must not be used to mark answer sheets or the computer would record a score of "0" for that paper. Students were also cautioned about the danger of stray marks on the answer sheet.

Previous review of the literature led to the hypothesis that raw achievement scores on the computer-scored version would be significantly lower than raw achievement scores on the hand-scored version due to the influence of computer induced test anxiety associated with the machine-scored answer sheet. This hypothesis was tests at the p=.05 level of significance by means of the Kolmogorov-Smirnov Test (Hays, 1963, p. 614) which provided a direct comparison between the two sample distributions. The sets of scores were not found to differ significantly, D-Max = .166, N.S.

Experimental Design

Because a difference in the predicted direction was indicated in the pilot study and a further review of the literature produced no new data concerning test anxiety resulting from knowledge of method of scoring, the main study was undertaken utilizing a larger number of subjects and a more reliable achievement measure. In addition, the two examination

forms, computer-scored and hand-scored, were to be administered by instructors other than this investigator, who were not to be aware of which students were to receive which form of the examination

The experimental design for this study was a 2-group post test design (Campbell & Stanley, 1966). Group A was comprised of those students receiving the computer-scored examination while Group B consisted of those students receiving the hand-scored examination.

Subjects comprising the two experimental groups were all those enrolled in the three sections of a graduate education class, Basic Principles of Measurement and Evaluation, Fall Term, 1974, at the University of Oklahoma. Each of the seventy students was randomly assigned to one of the two groups by the instructors who passed out the shuffled test booklets. Booklets were identical in content and appearance but for a single sentence in the directions on the cover sheet which stated how the given test was to be scored.

By utilizing identical appearing test booklets it was possible to randomly assign students from each section to each of the experimental groups. This procedure was an attempt to control the factors of instructor differences including the method of presenting examinations to a class, supplementary verbal instructions, and student-teacher interaction. Also, other factors such as time of day, classroom

differences, and student knowledge of experimental participation, which are generally acknowledged to be important experimental variables, were controlled. Random assignment of subjects to groups was an attempt to control for these variables without the necessity of blocking on each variable or artificially structuring each section so as to eliminate all possible between group differences except that of experimental treatment. This was a process which, according to Glass and Stanley (1970, p. 491), ". . . may reduce "error" considerably, but it does so at the expense of limiting the generalizability of the findings. . . ." Blocking, in addition to complicating the study was a rather doubtful procedure at best. Kerlinger (1973, pp. 309-310) stated that, "Theoretically, randomization is the only method of controlling all possible extraneous variables. All other methods leave many possibilities of inequality."

Students in both groups took the achievement test (see Appendix E) which was immediately followed by the Test Anxiety Scale (see Appendix D). Students were at no time informed that they were experimental subjects. By keeping the experiment totally covert, it was hoped that the experimental conditions would be as life-like as possible which would in turn contribute to the generalizability of the study (Page, 1958).

Instruments

The Text Anxiety Scale (TAS) was selected for use in this study as it fit the following criteria: ease of administration, no time limit, and a measure of state anxiety. Ease of administration was an important consideration because the anxiety measure was to be obtained in an actual classroom situation under realistic testing conditions. Hence, the instrument to be used had to be such that it could be appended to the achievement measure with little inconvenience or alarm on the part of students taking the tests. This criterion eliminated several of the longer anxiety measures such as the various adject[redacted]which Taylor's (1953) is a prime example[redacted]y time consuming to meet this crit[redacted]

More a[redacted]was the Test Anxiety Questionnaire [redacted]1952). The TAQ was a thirty six ite[redacted]r simple to administer and not in itself ti[redacted] However, the TAQ was contained within the 180 item Student Biographical Inventory. It was this feature of the TAQ which made it impractical for use.

The TAS, a self contained instrument with sixteen true/false items, seemed more appropriate. Another factor supporting the TAS was its answer format. Whereas the TAQ was anwered by marking one's choice on a rating scale, the TAS used a much simpler T/F format. Because of its brevity

Instruments

The Text Anxiety Scale (TAS) was selected for use in this study as it fit the following criteria: ease of administration, no time limit, and a measure of state anxiety. Ease of administration was an important consideration because the anxiety measure was to be obtained in an actual classroom situation under realistic testing conditions. Hence, the instrument to be used had to be such that it could be appended to the achievement measure with little inconvenience or alarm on the part of students taking the tests. This criterion eliminated several of the longer anxiety measures such as the various adjective checklists of which Taylor's (1953) is a prime example. Such lists are overly time consuming to meet this criterion.

More appropriate for this study was the Test Anxiety Questionnaire (TAQ) (Sarason & Mandler, 1952). The TAQ was a thirty six item test which was rather simple to administer and not in itself time consuming. However, the TAQ was contained within the 180 item Student Biographical Inventory. It was this feature of the TAQ which made it impractical for use.

The TAS, a self contained instrument with sixteen true/false items, seemed more appropriate. Another factor supporting the TAS was its answer format. Whereas the TAQ was anwered by marking one's choice on a rating scale, the TAS used a much simpler T/F format. Because of its brevity

and simplicity in scoring the TAS seemed more appropriate for this study. It was a nontimed instrument and could be given at any time. The instructor was not required to have all students begin at the same time or time them individually. Furthermore, the TAS could be stapled to the test booklet and answered by the student at his own pace.

The third criterion was that the anxiety measure be one of state anxiety, specifically, test anxiety. Alpert & Haber (1960) reported that a measure of specific anxiety such as the TAS was "a more satisfactory measure" than one of various measures of general anxiety. The TAS was reported to have a test-retest reliability (ten-week interval) of .82; no validity figures were reported, however (Alpert & Haber, 1960).

The achievement measure utilized in this study was an instructor constructed, thirty-four item multiple-choice test, (see Appendix E) designed to diagnose deficiencies in students' mathematics backgrounds prior to their becoming involved in the course Basic Principles of Measurement and Evaluation. Responses to the test items were made on a separate answer sheet (see Appendix F). A cover sheet directing students to read each question carefully and to respond to the respective items on the separate answer sheet was attached to the front of the test booklet. Students were further given directions on how to fill out the answer sheets and were told of the method of scoring to be used. As a check to verify

that the test directions had been read, students were instructed to place their names on both the test cover sheet and answer sheet. This is contrary to the standard test directions which request one's name on only the answer sheet.

The difference between the achievement test used by the hand-scored group and the computer-scored group was a single sentence on the cover sheet which stated the method of scoring to be used on each test (Appendices B & C), other than that the booklets and answer sheets were identical. Reliability of the instrument (KR-20) was determined to be .85, standard deviation 5.73.

CHAPTER III

ANALYSIS AND INTERPRETATION OF THE DATA

Mean raw achievement test scores were calculated from the data (Appendix A). Upon comparison of these means, the hand-scored group was found to have a higher mean raw score (see Table 1). Although the mean for the hand-scored group did exceed that of the computer-scored group, it was not a meaningful difference. To be meaningful, the mean raw scores would have differed by at least ten per cent or four points.

TABLE 1

MEANS AND STANDARD DEVIATION FOR THE TWO ACHIEVEMENT TEST FORMS

Test Form	N	$\bar{X}$	S.D.
Computer-Scored	36	21.4	6.03
Hand-Scored	34	22.3	5.53

The *t*-test for a difference between two independent means was used to test whether this difference in means was statistically significant. Applicability of the *t*-test rests upon several assumptions which were first met. First, that the samples were randomly selected was insured by directions

to instructors requesting that they shuffle computer and hand-scored test forms together to insure that they were passed out to and within each of the three sections randomly. Thus, the two groups were distributed at random among the three sections. Second, that the population variances were homogeneous was verified by an F max test on the data (Bruning & Kintz, 1968, pp. 107-108). The variance of each group was calculated and the larger value divided by the smaller. The resulting quotient was then compared to the F distribution table (Hays, 1963, pp. 677-679). The obtained F value was not statistically significant, F (35,33) = 1.19, indicating that group variances were homogeneous. The assumption of normality of distributions was not tested as Hays (1963, p. 322) reported "So long as the sample size is even moderate for each group, quite severe departures from normality seems to make little practical difference in the conclusions reached."

TABLE 2

VARIANCES OF THE ACHIEVEMENT TEST FORMS

Test Form	df	Variances	F
Computer-Scored	35	36.37	1.19
Hand-Scored	33	30.56	

After the assumptions were checked the t-test for a difference between two independent means (Bruning & Kintz,

1968, pp. 9-10) was used. The mean achievement score of the hand-scored group was not significantly higher than the mean achievement score of the computer-scored group, t (68) = .65, N.S. (see Table 3). Note also that the standard error of the difference between means exceeds the value of the difference. The power of the above test to detect a meaningful difference of four raw score points was calculated post hoc and determined to be .80 (Glass & Stanley, 1970, pp. 376-377).

TABLE 3

t-TEST SUMMARY TABLE FOR RAW ACHIEVEMENT SCORES

Test Form	df	$\bar{X}$	$\bar{X}$ diff.	S.E. diff.	t
Computer-Scored	35	21.4	0.90	1.39	.65
Hand-Scored	33	22.3			

When mean raw TAS scores were calculated for both groups, very little difference was found (see Table 4). The hand-scored group was found to have a slightly greater mean score than the computer-scored group.

TABLE 4

TAS MEANS AND STANDARD DEVIATIONS FOR THE TWO TEST FORMS

Test Form	N	$\bar{X}$	S.D.
Computer-Scored	36	6.56	3.62
Hand-Scored	34	6.79	4.06

As explained earlier, an F max test was necessary to check for homogenity of variances prior to computation of the t-test, the assumption of randomness had been met previously. The F max test (Bruning & Kintz, 1968, pp. 107-108) yielded the value F (35, 33) = 1.26, N.S. It was concluded that the variances were homogeneous. The data are presented in Table 5.

TABLE 5

TAS VARIANCES OF THE TWO GROUPS

Test Form	df	Variances	F
Computer-Scored	35	13.1	1.26
Hand-Scored	33	16.5	

Once the assumptions had been checked the mean raw TAS scores were compared by use of the t-test for a difference between two independent means (Bruning & Kintz, 1968, pp. 9-10). The test was not significant, t (68) = .26. These data are shown in Table 6.

TABLE 6

t-TEST SUMMARY TABLE FOR TAS RAW SCORES

Test Form	df	$\bar{X}$	$\bar{X}$ diff.	S.E. diff.	t
Computer-Scored	35	6.56	0.23	0.920	.26
Hand-Scored	33	6.79			

Before the TAS raw scores could be correlated with the achievement scores it was first necessary to determine how the two sets of scores were distributed. In graphing the sets of scores it was determined that the relationship between them was nonlinear. This precluded the use of Pearson r. The proper coefficient to use when the relationship between two sets of data is curvilinear is the eta coefficient (Downie & Health, 1970, p. 116; Hays, 1963, pp. 547-548). Eta was found to be .12. As this statistic, when squared, is distributed as F with J - 1 and N - J degrees of freedom eta squared was converted to F (Hays, 1963, p. 548) to test for significance. The resulting F value was not significant, F (1,68) = .966. It was thus concluded that Test Anxiety Scales Scores (TAS) and achievement scores were not significantly related.

Summary of Results

Hypothesis 1, which predicted that the mean raw score on the hand-scored achievement test would be greater than the mean raw score on the computer-scored achievement test, was not supported. Although the hand-scored mean was greater than the computer-scored mean, the standard error of the difference was fifty per cent larger than the difference between the means which more than accounted for the observed difference. It would be reasonable to say that no statistically significant difference existed between the two treatments. Knowledge of how the achievement test was to be

scored did not appreciably affect the groups' performance.

Hypothesis 2 had predicted that students taking the computer-scored achievement test would experience more anxiety than students taking the hand-scored achievement test and thus score higher on the Test Anxiety Scale. Evidently, this was not true as the mean raw TAS score for the computer-scored group was not the higher. The hand-scored group received a higher mean TAS score, although, this difference was not significant.

Hypothesis 3 stated that raw achievement test scores and raw TAS scores would be significantly correlated. This was not demonstrated since the obtained value of eta was .12 which was not significant. On this achievement test, higher levels of test anxiety were not associated with lower achievement test performance.

CHAPTER IV

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Much research dealing with the personality variable of anxiety has been conducted in the last twenty years. Sarason and Mandler have pioneered research in test anxiety and it was this area of anxiety and its effect upon students' achievement test performance which was the topic of this investigation.

The present study undertook to test whether students received lower grades on an achievement test which they knew was to be computer-scored than they did on a like examination which was hand-scored. In order to test student claims of being penalized by computer-scored examinations, two randomly assigned groups of subjects were given the same mathematics achievement test which differed only in a part of the instructions as the treatment. Group A received the sentence: "This examination will be computer-scored so . . ." in its test instructions; while Group B received the following: "This examination will be hand-scored but . . ." in its test instructions. Each group received the same achievement test followed by the Test Anxiety Scale.

It had been hypothesized that those students taking the computer-scored form would have higher test anxiety (TAS scores) and correspondingly lower achievement scores than the group taking the hand-scored form. It was also predicted that TAS and achievement scores would show a significant inverse correlation. None of these predictions were borne out by this study since all statistical tests were nonsignificant.

Results of the study showed that there was very little, if indeed any, real effect on students' performance as a result of knowing that the exam to be taken would be computer-scored. Mean raw scores for the hand-scored and computer-scored achievement tests were not significantly different which indicated no significant increase in test anxiety as a result of knowledge about test scoring means. This was evidenced by the fact that Test Anxiety Scale scores for the two groups, computer-scored and hand-scored, were not significantly different. Had students been more disturbed about having the test computer-scored they should have manifested this on the TAS in terms of higher scores.

When Test Anxiety Scale scores were correlated with achievement test scores the result was likewise nonsignificant. This indicated that test anxiety, at least as measured by the TAS and as generated by computer-scoring, had little relation to students' achievement test scores.

Conclusion

Evidently student claims heard by this examiner and reported in the Introduction, Chapter One, are but superficial in nature and serve possibly as face saving devices to protect a student from disgrace should a poor grade be received. Another possible reason for the lack of expected results could be that test anxiety, as measured by the TAS, is not of such a magnitude as to be influential on graduate students. Or, as Mandler & Sarason (1952) stated, influence of familiarity with test material must be considered in interpreting test anxiety results. It is quite likely that many of the subjects of this study were already at either such a high or low level of anxiety as suggested by the inverted-u hypothesis that the treatments (type of scoring) had little effect. Previous experiences and performances in similar achievement test situations, also, may have clouded the results (Mandler & Sarason, 1952). A questionnaire following the examination could have possibly provided this information. Students could have been grouped as to similiarity of background and personality variables in a counter balanced design opposed to the present study's technique for grouping.

In conclusion, this research did not support students' claims that their grades were lower when they took computer-scored examinations. No significant differences in achievement or test anxiety measures were found. Since it was demonstrated in this study that scores did not differ significantly

between computer-scored and hand-scored achievement tests, instructors need not mention whether tests are to be computer scored or hand scored when the tests are administered. Scoring instructions did not appear to raise anxiety levels as measured by the TAS, nor did they adversely affect test performance.

Implications

On the basis of the findings of this investigation, the following recommendations for future research are suggested:

1. Research is needed to determine if level of experience at answering tests on machine scorable answer sheets is a variable which should be included in future studies.

2. The present study should be replicated utilizing an actual classroom achievement test such as a midterm examination. It is possible that all students were not totally involved with the present investigation's achievement measure as they would be in a test upon which their grades depended.

3. Replication of the present investigation with the Test Anxiety Scale administered prior to the achievement measure should be considered. This could possibly provide a truer estimation of anxiety manifested during achievement testing situations.

REFERENCES

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

Anastasi, A. Psychological testing, 3rd edition. New York: Macmillan, 1968.

Bell, F., Hoff, H., & Hoyt, K. Answer sheets do make a difference. Personnel Psychology, 1964, 17, 65-71.

Bruning, J., & Kintz, B. Computational handbook of statistics. Glenview: Scott, Foresman, 1968.

Campbell, D., & Stanley, J. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1966.

Downie, N., & Heath, R. Basic statistical methods, (3rd ed.). New York: Harper & Row, 1970.

Dunlap, J. Problems arising from the use of a separate answer sheet. Journal of Psychology, 1940, 10, 3-48.

Elwood, D. Automation of psychological testing. American Psychologist, 1969, 24, 287-289.

Fischer, W. Theories of anxiety. New York: Harper & Row, 1970.

Forhetz, J. Anxiety differences in difficult and easy test conditions. U.S. Dept. of HEW, 1969. (ERIC Document No. ED 053 421)

Glass, G., & Stanley, J. Statistical methods in education and psychology. Englewood Cliffs: Prentice-Hall, 1970.

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

Hedl, J. An evaluation of a computer-based intelligence test. Office of Naval Research, 1971. (ERIC Document No. ED 061 722).

Kerlinger F. Foundations of behavioral research. New York: Holt, Rinehart and Winston, 1973.

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

Mathis, A., Smith, T., & Hansen, D. College students' attitudes toward computer-assisted instruction. Journal of Educational Psychology, 1970, 61, 46-51.

Munz, D. The effects of item difficulty sequences on achievement testing scores under two conditions of anxiety. Unpublished master's thesis, University of Oklahoma, 1968.

Page, E. Educational research: replicable or generalizable? Phi Delta Kappan, 1958, 39, 302-304.

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

Petty, E. The effect of four methods of recording answers to standardized test items on mean achievement at four grade levels. Unpublished doctoral dissertation, University of Oklahoma, 1964.

Sarason, I. The effects of anxiety and threat on the solution of a difficult task. Journal of Abnormal and Social Psychology, 1961, 62, 165-168.

Sarason, I. Personality: an objective approach. New York: John Wiley & Sons, 1966.

Sarason, I., & Ganzer, V. Anxiety, reinforcement, and experimental instructions in a free verbalization situation. Journal of Abnormal and Social Psychology, 1962, 65, 300-307.

Sarason, I., & Paoloa, E. The relationship of test and general anxiety, difficulty of task and test performance. Journal of Experimental Psychology, 1960, 59, 185-191.

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

Sarason, S., Mandler, G., & Graighill, F. The effect of different instructions on anxiety and learning. Journal of Abnormal and Social Psychology, 1952, 47, 561-565.

Silberman, H. The digital computer in education. Phi Delta Kappan, 1962, 43, 346.

Smith, R. Examination by computer. Behavioral Science, 1963, 8, 76-79.

Smouse, A., & Munz, D. The effects of anxiety and item difficulty sequences on achievement testing scores. The Journal of Psychology, 1968, 68, 181-84.

Smouse, A., & Munz, D. Item difficulty sequencing and response style: a follow up analysis. Educational and Psychological Measurement, 1969, 29, 469-472.

Spielberger, C. (Ed.) Anxiety: current trends in theory and research. New York: Academic Press, 1972.

Sweeney, C. The effects of stress and achievement anxiety on academic test performance. Unpublished doctoral dissertation, University of Oklahoma, 1968.

Sweeney, C., Smouse, A., Rupiper, O., & Munz, D. A test of the inverted - u hypothesis relating achievement anxiety and academic test performance. The Journal of Psychology, 1970, 74, 267-273.

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

Thorndike, R. (Ed.) Educational measurement. New York: American Council on Education, 1971.

Waite, R., Sarason, S., Lightall, F., & Davidson, K. A study of anxiety and learning in children. Journal of Abnormal and Social Psychology, 1958, 57, 267-270.

Whitlock, J. Automatic data processing in education. New York: Macmillan, 1964.

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

Zuckerman, M. The development of an adjective check list for the measurement of anxiety. Journal of Consulting Psychology, 1960, 24, 457-462.

Zuckerman, M., & Lubin, B. Normative data for the multiple affect adjective check list. Psychological Reports, 1965, 16, 438.

APPENDIX A

ACHIEVEMENT AND TEST ANXIETY RAW SCORES

Subject	Test Form	Achievement Score	TAS Score
1	C	21	3
2	C	16	7
3	C	23	0
4	C	20	11
5	H	27	16
6	H	21	1
7	H	25	3
8	H	26	4
9	H	21	8
10	H	17	9
11	C	19	7
12	C	13	9
13	H	15	4
14	C	28	1
15	H	28	3
16	H	23	2
17	C	22	5
18	C	27	5
19	H	31	1
20	C	26	8
21	H	21	3
22	C	17	10
23	H	26	11
24	C	16	2
25	H	28	10
26	C	28	8
27	C	23	7
28	H	20	6
29	H	24	3
30	C	17	7
31	H	23	2
32	H	25	10
33	H	13	12
34	H	22	11
35	C	24	9

ACHIEVEMENT AND TEST ANXIETY RAW SCORES

Subject	Test Form	Achievement Score	TAS Score
36	C	27	3
37	C	27	8
38	C	14	8
39	C	19	8
40	H	3	3
41	H	18	5
42	C	21	8
43	H	29	3
44	C	20	8
45	H	21	9
46	H	27	8
47	H	17	8
48	C	28	5
49	C	19	14
50	C	25	7
51	H	30	10
52	C	24	4
53	H	20	7
54	C	14	13
55	H	17	10
56	C	10	9
57	C	28	9
58	H	20	2
59	H	24	9
60	C	26	3
61	C	19	5
62	H	20	11
63	C	4	15
64	C	19	4
65	H	26	14
66	C	27	3
67	H	25	10
68	H	25	3
69	C	31	1
70	C	29	2

APPENDIX B

Name_____
(print)

College of Education

The University of Oklahoma

On the following pages you will find some questions which deal with your knowledge of mathematics.

Read each question carefully and respond as directed on the separate answer sheet. This examination will be computer-scored so make all marks dark and within the spaces provided. Be sure to erase completely any answers you wish to change. Check to see that your name is on both this page and the answer sheet.

APPENDIX C

Name_____
(print)

College of Education

The University of Oklahoma

On the following pages you will find some questions which deal with your knowledge of mathematics.

Read each question carefully and respond as directed on the separate answer sheet. This examination will be hand-scored but make all marks dark and within the spaces provided. Be sure to erase completely any answers you wish to change. Check to see that your name is on both this page and the answer sheet.

APPENDIX D

Test Anxiety Scale

1.	While taking an important examination, I perspire a great deal.	T	F
2.	I get to feeling very panicky when I have to take a surprise exam.	T	F
3.	During tests, I find myself thinking of the consequences of failing.	T	F
4.	After important tests I am frequently so tense that my stomach gets upset	T	F
5.	While taking an important exam I find myself thinking of how much brighter other students are than I am.	T	F
6.	I freeze up on things like intelligence tests and final exams.	T	F
7.	If I were to take an intelligence test I would worry a great deal before taking it.	T	F
8.	During course examinations, I find myself thinking of things unrelated to the actual course material.	T	F
9.	During a course examination, I frequently get so nervous that I forget facts I really know.	T	F
10.	If I knew I was going to take an intelligence test, I would feel confident and relaxed beforehand.	T	F
11.	I usually get depressed after taking a test.	T	F
12.	I have an uneasy, upset feeling before taking a final examination.	T	F
13.	When taking a test, my emotional feelings do not interfere with my performance.	T	F
14.	Getting a good grade on one test doesn't seem to increase my confidence on the second.	T	F
15.	After taking a test I always feel I could have done better than I actually did.	T	F
16.	I sometimes feel my heart beating very fast during important tests.	T	F

APPENDIX E

Name_____
(print)

Indicate the number of the correct answer on the answer sheet.

1. $6 \times (3 + 2) =$

   1. 11
   2. 20
   3. 30
   4. None of the above

2. $10 + 2 + 3 =$

   1. ½
   2. 2
   3. 8
   4. None of the above

3. $4^2 =$

   1. 2
   2. 4
   3. 8
   4. 16
   5. None of the above

4. $26 \times ½ \times 0 =$

   1. 0
   2. 13
   3. 26
   4. None of the above

5. $\frac{36 \times 0}{12}$

   1. 0
   2. 3
   3. 60
   4. 120
   5. None of the above

6. (Omitted, no answer or none of the above)

7. $\frac{6!}{3!\,2!}$

1. 0
2. 1
3. 30
4. 60
5. None of the above

Arrange the following numbers in order from lowest to highest as they appear on the number line. (problem 8-10)

8. -2, +4, -7, +1, 0

1. 0, +1, -2, 4, -7
2. -7, -2, 0. +1, +4
3. -2, -7, 0, +1, +4
4. 0, -2, -7, +1, +4
5. None of the above

9. .5, -1.5, 3.00, -.75, .00

1. -.75, -1.5, .00, .5, 3.00
2. .00, .5, -.75, -1.5, 3.00
3. -1.5, -.75, .00, .5, 3.00
4. .00, .5, .80, 1.5, 2.50
5. None of the above

10. 3/8, 3/4, 1/2, 4/6, 5/12

1. 1/2, 3/4, 3/8, 4/6, 5/12
2. 3/8, 5/12, 4/6, 1/2, 3/4
3. 3/4, 4/6, 1/2, 5/12, 3/8
4. 1/2, 3/4, 4/6, 3/8, 5/12
5. None of the above

Perform the following operations. Express each answer in the simplest form. (problem 11-28)

11. (+12) - (-3) - (+4)

1. 5
2. 11
3. 13
4. 19
5. None of the above

12. 9/24 x 10/12

1. 9/20
2. 5/16
3. 3/8
4. 19/36
5. None of the above

13. (-3) x (-6) =

1. -18
2. 18
3. 1/2
4. -9
5. None of the above

14. -6 - (4-2) - (6+3) =

1. -1
2. 2
3. -11
4. -17
5. None of the above

15. $\frac{(+8) \times (-2) \times (-3)}{(-4) \times (+1)}$

1. -1
2. 12
3. -12
4. 10/13
5. None of the above

16. $\frac{(-3) \times (+12)}{-4}$

1. 0
2. -9/4
3. -9
4. 9/4
5. None of the above

17. 1/4 x 3/4 =

1. 1
2. 3/6
3. 3
4. 1/3
5. None of the above

18. $3^3 + 2^1 =$

1. 0
2. 11
3. 12
4. 28
5. None of the above

19. $(1/3)^3 =$

1. 1
2. 1/9
3. 1/27
4. 3/9
5. None of the above

20. 1/2 + 1/11 + 3/22 =

1. 5/35
2. 15/22
3. 5/22
4. 4/7
5. None of the above

21. 15 x /25 =

1. 1.25
2. 12.5
3. .0125
4. .00125

22. $\overline{/64}$ =

1. 8
2. 16
3. 32
4. 64
5. None of the above

23. .21 $\overline{)6.54}$ =

1. 314
2. 31.4
3. 3.14
4. .314
5. None of the above

24. $\frac{(.6)\ (.04)}{(.03)\ (2)} =$

1. .004
2. .04
3. .4
4. 4
5. None of the above

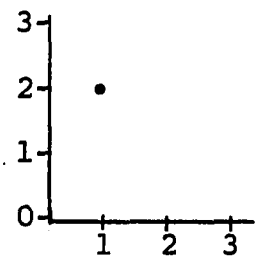
25. $(2/3)^2 =$

1. 4/3
2. 1/3
3. 1 1/3
4. 4/9
5. None of the above

26. What is 16 per cent of 32?

1. 2
2. 5.12
3. .5
4. 32
5. None of the above

27. What are the coordinates of the dot?



1. (1,2)
2. (0,3)
3. (2,3)
4. (1,3)
5. None of the above

28. Which of the following is equation for a straight line?

1. $a = 2$
2. $b = 3c + 1$
3. $c = f^2 + g$
4. $1 = a^2 + b^2$
5. None of the above

Solve for X (problems 29 - 32)

29. $2X + 18 = 0$

1. 9
2. -9
3. 0
4. 1
5. None of the above

30. $\frac{2X}{3} = 12$

1. 4
2. 6
3. 8
4. 18
5. None of the above

31. $4X + 6 = 2X + 2$

1. -2
2. 0
3. 1
4. X + 1
5. None of the above

32. $\frac{3}{4} = \frac{5}{6\ (x - 1)}$

1. 1/9
2. 9
3. 15/24
4. 24/15
5. None of the above

Find the following values when X and Y indicate any score and $\sum$ means to sum. (problems 33 - 36)

X	Y
3	1
6	3
12	2
2	6
4	11

33. $\sum X$ =

1. 13
2. 15
3. 27
4. 29
5. None of the above

34. $\sum XY$ =

1. 17
2. -10
3. 209
4. 285
5. None of the above

35. $(\sum X)^2$ =

1. 15
2. 125
3. 209
4. 220
5. None of the above

APPENDIX F

NAME ____________ LAST FIRST MIDDLE GRADE ________ SEX ______ M OR F DATE OF BIRTH ____________ YEAR MONTH DAY

DATE ____________ YEAR MONTH DAY AGE ______ SCHOOL ____________ CITY ____________

INSTRUCTOR ____________

NAME OF TEST ____________

BE SURE TO MAKE YOUR MARKS HEAVY AND BLACK

ERASE COMPLETELY ANY ANSWERS YOU WISH TO CHANGE

IDENTIFICATION NUMBER

0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9

1	1 2 3 4 5	2	1 2 3 4 5	3	1 2 3 4 5	4	1 2 3 4 5
5	1 2 3 4 5	6	1 2 3 4 5	7	1 2 3 4 5	8	1 2 3 4 5
9	1 2 3 4 5	10	1 2 3 4 5	11	1 2 3 4 5	12	1 2 3 4 5
13	1 2 3 4 5	14	1 2 3 4 5	15	1 2 3 4 5	16	1 2 3 4 5
17	1 2 3 4 5	18	1 2 3 4 5	19	1 2 3 4 5	20	1 2 3 4 5
21	1 2 3 4 5	22	1 2 3 4 5	23	1 2 3 4 5	24	1 2 3 4 5
25	1 2 3 4 5	26	1 2 3 4 5	27	1 2 3 4 5	28	1 2 3 4 5
29	1 2 3 4 5	30	1 2 3 4 5	31	1 2 3 4 5	32	1 2 3 4 5
33	1 2 3 4 5	34	1 2 3 4 5	35	1 2 3 4 5	36	1 2 3 4 5
37	1 2 3 4 5	38	1 2 3 4 5	39	1 2 3 4 5	40	1 2 3 4 5
41	1 2 3 4 5	42	1 2 3 4 5	43	1 2 3 4 5	44	1 2 3 4 5
45	1 2 3 4 5	46	1 2 3 4 5	47	1 2 3 4 5	48	1 2 3 4 5
49	1 2 3 4 5	50	1 2 3 4 5	51	1 2 3 4 5	52	1 2 3 4 5
53	1 2 3 4 5	54	1 2 3 4 5	55	1 2 3 4 5	56	1 2 3 4 5
57	1 2 3 4 5	58	1 2 3 4 5	59	1 2 3 4 5	60	1 2 3 4 5
61	1 2 3 4 5	62	1 2 3 4 5	63	1 2 3 4 5	64	1 2 3 4 5
65	1 2 3 4 5	66	1 2 3 4 5	67	1 2 3 4 5	68	1 2 3 4 5
69	1 2 3 4 5	70	1 2 3 4 5	71	1 2 3 4 5	72	1 2 3 4 5
73	1 2 3 4 5	74	1 2 3 4 5	75	1 2 3 4 5	76	1 2 3 4 5
77	1 2 3 4 5	78	1 2 3 4 5	79	1 2 3 4 5	80	1 2 3 4 5
81	1 2 3 4 5	82	1 2 3 4 5	83	1 2 3 4 5	84	1 2 3 4 5
85	1 2 3 4 5	86	1 2 3 4 5	87	1 2 3 4 5	88	1 2 3 4 5
89	1 2 3 4 5	90	1 2 3 4 5	91	1 2 3 4 5	92	1 2 3 4 5
93	1 2 3 4 5	94	1 2 3 4 5	95	1 2 3 4 5	96	1 2 3 4 5
97	1 2 3 4 5	98	1 2 3 4 5	99	1 2 3 4 5	100	1 2 3 4 5
101	1 2 3 4 5	102	1 2 3 4 5	103	1 2 3 4 5	104	1 2 3 4 5
105	1 2 3 4 5	106	1 2 3 4 5	107	1 2 3 4 5	108	1 2 3 4 5
109	1 2 3 4 5	110	1 2 3 4 5	111	1 2 3 4 5	112	1 2 3 4 5
113	1 2 3 4 5	114	1 2 3 4 5	115	1 2 3 4 5	116	1 2 3 4 5
117	1 2 3 4 5	118	1 2 3 4 5	119	1 2 3 4 5	120	1 2 3 4 5
121	1 2 3 4 5	122	1 2 3 4 5	123	1 2 3 4 5	124	1 2 3 4 5
125	1 2 3 4 5	126	1 2 3 4 5	127	1 2 3 4 5	128	1 2 3 4 5
129	1 2 3 4 5	130	1 2 3 4 5	131	1 2 3 4 5	132	1 2 3 4 5
133	1 2 3 4 5	134	1 2 3 4 5	135	1 2 3 4 5	136	1 2 3 4 5
137	1 2 3 4 5	138	1 2 3 4 5	139	1 2 3 4 5	140	1 2 3 4 5
141	1 2 3 4 5	142	1 2 3 4 5	143	1 2 3 4 5	144	1 2 3 4 5
145	1 2 3 4 5	146	1 2 3 4 5	147	1 2 3 4 5	148	1 2 3 4 5
149	1 2 3 4 5	150	1 2 3 4 5	151	1 2 3 4 5	152	1 2 3 4 5
153	1 2 3 4 5	154	1 2 3 4 5	155	1 2 3 4 5	156	1 2 3 4 5
157	1 2 3 4 5	158	1 2 3 4 5	159	1 2 3 4 5	160	1 2 3 4 5