DOCUMENT RESUME

ED 076 060 EM 011 071

AUTHOR O'Neil, Harold F., Jr.; And Others
TITLE The Effects of Anxiety Reduction Techniques on Anxiety and Computer-Assisted Learning and Evaluation of College Students. Final Report.
INSTITUTION Texas Univ., Austin. Computer-Assisted Instruction Lab.
SPONS AGENCY National Center for Educational Research and Development (DHEW/OE), Washington, D.C.
BUREAU NO BR-1-0210
PUB DATE Dec 72
GRANT OEG-4-71-0027
NOTE 136p.

EDRS PRICE MF-$0.65 HC-$6.58
DESCRIPTORS *Anxiety; *College Students; *Computer Assisted Instruction; Curiosity; Intelligence Tests; Progressive Relaxation; *Student Behavior; *Testing

ABSTRACT

The goal of this project was to examine various anxiety reduction techniques on the state anxiety levels and performance of college students. These techniques ranged from instructional to experimental treatments and were investigated in a range of computer-based situations. The state-trait anxiety inventory developed by Spielberger, Gorsuch, and Lushene (1970) was used to measure both trait and state anxiety in these studies. The materials were presented by an IBM 1500 computer-assisted instruction system, which also presented the state anxiety scales and recorded student responses. Six studies were conducted to accomplish the project goals. The first study focused on the effect of stimulating curiosity as an anxiety-reducer. Next, a series of five studies examined the impact of anxiety on computer-based intelligence testing, and methods were investigated for reducing test anxiety. Prior terminal experience was shown to reduce anxiety for students during a computer-administered intelligence test. The comparative effectiveness of these test anxiety reduction treatments across the several studies was discussed. (Author)

THE EFFECTS OF ANXIETY REDUCTION TECHNIQUES ON ANXIETY AND COMPUTER-ASSISTED LEARNING AND EVALUATION OF COLLEGE STUDENTS

Final Report

Grant No. OEG-4-71-0027

Project No. 1-0210

Project Director

Harold F. O'Neil, Jr.

With the Assistance of:

John J. Hedl, Jr.

Barbara Leherissey McCombs

Richard D. Grant, Jr.

Jody L. Fitzpatrick

Wilson A. Judd

Computer-Assisted Instruction Laboratory

The University of Texas at Austin

December 1972

THE EFFECTS OF ANXIETY REDUCTION TECHNIQUES ON ANXIETY AND COMPUTER-ASSISTED LEARNING AND EVALUATION OF COLLEGE STUDENTS

Abstract

The goal of this grant was to examine various anxiety reduction techniques on state anxiety levels and performance. These techniques ranged from instructional to experimental treatments and were investigated in a range of computer-based situations. The State-Trait Anxiety Inventory developed by Spielberger, Gorsuch, and Lushene (1970) was used to measure both trait and state anxiety in these studies. The materials were presented by an IBM 1500 Computer-Assisted Instruction system, which also presented the state anxiety scales and recorded student repponses.

Six studies were conducted to accomplish the project goals. The first study focused on the effect of stimulating curiosity as an anxiety reducer. Next, a series of five studies examined the impact of anxiety on computer-based intelligence testing, and methods were investigated for reducing test anxiety. Prior terminal experience was shown to reduce anxiety for students during a computer-administered intelligence test. The comparative effectiveness of these test anxiety reduction treatments across the several studies was discussed.

TABLE OF CONTENTS

LIST OF TABLES

Table		Page

LIST OF FIGURES

Figure		Page

THE EFFECTS OF ANXIETY REDUCTION TECHNIQUES ON ANXIETY AND COMPUTER-ASSISTED LEARNING AND EVALUATION OF COLLEGE STUDENTS

INTRODUCTION

The funded research focused on th . of anxiety within learning and evaluation. The importance of this topi e..ected in the extensive literature which indicates that anxiety can interfere with the learning process and test performance (Sarason, 1960; Spielberger, 1966). As a result, the student's level of achievement is not commensurate with his intellectual aptitude, and his confidence in his own abilities is seriously undermined. Moreover, the valid assessment of his ability may be impaired by his anxiety. The funded research further extended the knowledge of the effects of anxiety on learning and evaluation by considering the impact of various anxiety reduction techniques on anxiety and performance. These techniques ranged from instructional to clinical treatments and were investigated in a variety of computer-based situations. The importance of extending research in these directions relates to the need for testing Trait-State Anxiety Theory (Spielberger, 1972) as a viable theoretical explanation for a wide number of school learning behaviors.

Computer-Assisted Instruction experimental situations were chosen as the basis for this research in that they provided a context which permits the presentation of stimulus materials under carefully controlled conditions that are more relevant to the real life needs of the student than is generally possible with traditional laboratory tasks. An additional research advantage for computer-assisted instruction is that it is also possible to measure anxiety and student responses during the actual learning and testing process (O'Neil, Spielberger, & Hansen, 1969). The measurement of anxiety as learning progresses enables the investigator to determine in finer detail the exact nature of the relationship between anxiety and performance. These capacities of the computer-assisted instruction approach help to bridge the gap between laboratory research on anxiety and learning and applications of learning principles in the classroom.

Hypotheses about the effects of anxiety on learning were derived from Trait-State Anxiety Theory (Spielberger, 1972). This theory provided the conceptual framework within which research on anxiety and computer-assisted learning as well as computer-based testing were examined.

Research on anxiety and learning has suffered from ambiguity with regard to the status of anxiety as a theoretical concept. Spielberger (1966) has emphasized the necessity of distinguishing between anxiety conceptualized as a transitory state and as a relatively stable personality trait. According to Spielberger (1966, pp. 16-17):

> Anxiety states (A-State) are characterized by subjective consciously perceived feelings of apprehension and tension, accompanied by or associated with activation or arousal of the autonomic nervous system. Anxiety as a personality trait (A-Trait) would seem to imply a motive or acquired behavioral disposition that predisposes an individual to perceive a wide range of objectively nondangerous circumstances as threatening, and to respond to these with A-State reactions disproportionate in intensity to the magnitude of the objective danger.

In the next section, Trait-State Anxiety Theory will be utilized as a conceptual framework for evaluating and interpreting findings from the anxiety literature that are most relevant to computer-assisted learning.

The Effect of Anxiety on Computer-Assisted Learning

Computer-Assisted Instruction (CAI) implies a set of procedures in which a computer is employed to control the selection, sequencing, and evaluation of instructional materials (Fishman, Keller, & Atkinson, 1968). In CAI systems, a computer in an instructional role interacts with a student.

In three recent studies, specific hypotheses derived from Trait-State Anxiety Theory were tested with computer-assisted learning materials. O'Neil, Spielberger, and Hansen (1969) investigated the relation between A-State and performance for college students who learned meaningful mathematical materials presented by an IBM 1440 CAI system (IBM, 1965). The state anxiety measures were changes in systolic blood pressure and scores on the state anxiety scale of the State-Trait Anxiety Inventory (STAI) (Spielberger, Gorsuch, & Lushene, 1970). Both state anxiety measures increased while students worked on difficult learning materials and decreased when they responded to easy materials. Moreover, students with high state anxiety scores (STAI) made more errors on the difficult materials and fewer errors on the easy materials, than did low state anxious students.

These findings were extended by O'Neil, Hansen, and Spielberger (1969). The results of this study confirmed the findings in the earlier study in that: (1) state anxiety scores increased while students worked on difficult materials and decreased when they responded to easy materials; and (2) high state anxious students made significantly more errors on the difficult materials than did low state anxious students. Although there was no relation between trait anxiety and performance, high trait anxious students responded throughout the learning task with higher levels of state anxiety than did low trait anxious students.

These results also provide support for Trait-State Anxiety Theory by demonstrating the need to: (1) distinguish between trait anxiety and state anxiety; and (2) obtain measures of state anxiety in the experimental situation.

On the assumption that the CAI situation involved some threat to self-esteem, Trait-State Anxiety Theory would predict that the magnitude of state anxiety would be greater for high trait anxious students than for low trait anxious students, but this expectation was not confirmed in the study in which students were selected on the basis of trait anxiety scores. A possible explanation is that while the CAI task was stressful because it was difficult, it was not necessarily more stressful for high trait anxious students than for low trait anxious students, because it did not evaluate the adequacy of the student's performance relative to others. If explicit negative evaluations concerning performance were given by the computer, then high trait anxious students might be expected to perceive the CAI situation as more threatening than did low trait anxious subjects, and to respond with higher levels of state anxiety.

O'Neil (1969) investigated this interpretation of the effects of negative evaluations on state anxiety and on performance. Female introductory psychology college students who differed in anxiety proneness were used as subjects. High trait anxious students in the stress condition showed a significantly greater initial increase in state anxiety from pretask levels than did the low trait anxious students. During the learning task, high trait anxious students in the stress condition showed a marked decline in state anxiety whereas level of state anxiety remained relatively constant for low trait anxious students. In the nonstress condition, both groups showed almost the same increase in state anxiety from pretask levels and approximately parallel changes in the level of state anxiety during the CAI learning task.

Students with high levels of state anxiety made more errors than did low state anxious students throughout the learning task. The differences in the performance of high state anxious and low state anxious subjects were significant on the easier sections of the CAI task, but not for the most difficult part of the task. These relationships between state anxiety and errors differed from previous research (O'Neil, Hansen, & Spielberger, 1969).

All of these anxiety and CAI studies have highlighted the need to distinguish between trait anxiety and state anxiety, and to obtain measures of state anxiety in the CAI situation. However, these conclusions have been generated by using a single set of CAI mathematical learning materials. Thus, another series of studies was done to test the generality of these conclusions by using verbal and graphical learning materials rather than mathematical ones, and, in addition, to discover how different response modes might affect anxiety.

In the first of this series of studies, Leherissey, O'Neil, and Hansen (1971a) investigated the effects of trait and state anxiety levels (Low, Medium, High) and response modes (Reading, Covert, Modified Multiple Choice, Constructed Response) on posttest achievement for familiar and technical materials dealing with heart disease. The four versions of the learning materials, presented to 148 students via an IBM 1500 computer-assisted instruction system, were as follows: (a) Reading version, in which the students were not required to make overt responses, but merely to read successive frames of material. Response blanks were filled in, and frames asking a question were presented in declarative form. (b) Covert version, which contained response blanks and interrogative frames. No overt responses were required. Students were instructed to merely "think" their answer to themselves and then signal to obtain the correct answer. (c) Modified Multiple Choice version, which required overt responses in the form of a typed word where response blanks appeared on the familiar and technical verbal portions of the material. On the technical pictorial portion of the learning material, which contained EKG drawings and tracings, students were required to read each frame and choose one of three or four multiple choice answers before being shown the correct response. (d) Constructed Response version, which was identical to the Modified Multiple Choice version on the familiar and technical verbal frames, but in which students were required to respond to technical pictorial frames by constructing graphical representations of EKG tracings using various keyboard dictionary characters in succession prior to receiving the correct answer.

High trait anxiety was associated with high state anxiety for all groups. Although prior programmed instruction research using the same materials indicated that the Constructed Response mode would lead to superior performance on a posttest covering technical materials when compared with a Reading mode, Leherissey, et al. (1971a) found no differences between Constructed Response and Reading groups presented the same technical materials via CAI. This failure to replicate the programmed instruction findings may have been due to the fact that the students in the Constructed Response groups had significantly higher state anxiety during learning of the technical materials and the posttest than the the Reading group. The Constructed Response groups also took over twice as long as the Reading group to complete the CAI materials. Negative debriefing comments by the Constructed Response group also indicated that they may have been more hostile than the Reading group toward the task. Thus, the amount of time required to complete the materials, high state anxiety, and hostility toward the task may have served to depress the posttest performance of the Constructed Response group.

In the second study of the series, Leherissey, O'Neil, Heinrich, and Hansen (1971) sought to replicate the findings of the preceding study, as well as to reduce state anxiety and improve performance by shortening the amount of time spent on the instructional materials. One hundred twenty-eight students were randomly assigned to one of four treatment groups consisting of Long and Short versions of the Reading and Constructed Response forms of the verbal and graphical materials. Hostility was also measured to explicate and extend the previous findings.

As in the previous study, high trait anxious students had higher levels of state anxiety throughout the experimental tasks than either medium or low trait anxious students, thus supporting predictions from Trait-State Anxiety Theory. The state anxiety analyses for both studies indicated that state anxiety scores decreased for both the Reading and Constructed Response groups from a pre-task measure to a second measure presented immediately upon completion of the familiar material. Whereas state anxiety scores remained at a lower level for the Reading group on a third measure taken upon completion of the technical materials, state anxiety scores for the Constructed Response groups increased during the technical materials. Students in the Constructed Response groups were also found to have higher levels of state anxiety during the posttest than students in the Reading groups in both studies.

Regarding the replicated performance results, neither levels of trait anxiety nor levels of state anxiety affected student performance on the pretest. Analysis of posttest performance for both studies indicated that students in the Reading groups performed better than students in the Constructed Response groups on portions of the posttest pertaining to familiar materials. Students in the Constructed Response groups generally took approximately twice as long as their Reading group counterparts to learn the instructional materials.

The hypothesis that shortening program length would lead to reductions in level of state anxiety, particularly for those students in the Constructed Response Short version, was not supported.

It was also hypothesized that shortening program length would improve the posttest performance of students in the Short Constructed Response group relative to the performance of students in the Short Reading group. Relevant to this hypothesis was the finding that students in the shortened program versions performed significantly better than in the longer version on the familiar posttest. Analysis also revealed a significant interaction between response modes and program length on familiar posttest, which indicated that whereas there was little difference in the performance of students in the Long and Short Reading groups, students in the Short Constructed Response version performed significantly better than students in the Long Constructed Response versions.

In addition, there was a significant interaction between level of state anxiety, response mode, and program length on the familiar portion of the posttest which indicated that level of state anxiety was less debilitating to the performance of students in the Long Constructed Response version. That is, students with Medium or High levels of state anxiety in the Short Constructed Response group performed at approximately the same level as students in the Reading versions, whereas for students in the Long Constructed Response group, Medium or High levels of state anxiety were particularly debilitating to performance. This interaction thus provides indirect evidence of the differential effects of state anxiety for students in the Long and Short program versions.

Analysis of the performance of students on the portion of the posttest pertaining to initial technical materials failed to support the hypothesis that shortening instructional time would improve performance.

As predicted, students in the Constructed Response groups had higher hostility scores than students in the Reading group. Contrary to predictions, however, students in the Long and Short program groups did not differ in mean hostility engendered by the learning task.

With respect to performance results of the two studies, several findings failed to replicate. Interactions involving level of trait anxiety and response mode on the familiar posttest were observed in both studies, but were in opposite directions. In addition, low state anxious students in the second study were found to perform significantly better than high state anxious students on the familiar posttest, while there was no main effect of state anxiety in the first study.

In summary, the findings of these two studies indicated that the impact of the Constructed Response mode variable is paramount, in that students in this response mode condition had high levels of anxiety and hostility, and poorer performance on the total technical posttest than students in the Reading groups. The major findings of both studies generally supported Trait-State Anxiety Theory and replicated the effects of response modes on state anxiety and performance in the CAI task. However, the instructional treatment of shortening time spent on the CAI task was not effective in reducing state anxiety. Shortening program length did prove effective in improving the performance of the Constructed Response group on the posttest for familiar and initial technical materials, an effect which may have been due to a decreased memory load for this group.

The findings in these two recent studies seem to indicate that it is not instructional time per se that is the critical variable for reducing state anxiety and improving performance. The intrinsic differences in the nature of the CAI learning task for the Reading and Constructed Response groups, including their differential affective and cognitive effects, imply a need for research efforts in the area of anxiety reduction which also take into account relevant task variables. It seems desirable that a variety of anxiety reduction techniques, from instructional design considerations to clinical treatments, should be investigated in a range of computer-based situations.

Anxiety Reduction

A problem of concern to educators is to determine the effects of anxiety on the learning of students, and in particular, to discover appropriate means to reduce the disruptive effects of anxiety on learning. Although it has been recognized that anxiety can interfere with the learning process (Sarason, 1960; Spielberger, 1966), relatively little research has been concerned with reducing anxiety in the learning situation. A major reason for the scarcity of research in the area of anxiety reduction may be the theoretical and methodological confusion regarding the construct of anxiety and how it should be measured. However, with the recent formulation of the Trait-State Anxiety Theory (Spielberger, 1972), it has been possible to differentiate conceptually between anxiety as a transitory state and as a relatively permanent trait.

With respect to a methodological solution to the confusion, several recent CAI studies have examined anxiety in the situation and have supported the contention that periodic A-State measures are needed to understand the relationship between anxiety and performance (Leherissey, O'Neil, & Hansen, 1971b; O'Neil, Spielberger, & Hansen, 1969; O'Neil, Hansen, & Spielberger, 1969; O'Neil, 1969). These CAI studies have shown that it is possible methodologically not only to measure performance as a function of anxiety, but also to measure changes in state anxiety as a function of experimental treatment.

The implication of these conceptual and methodological distinctions for research in anxiety reduction is primarily that one can actually measure whether anxiety has been in fact reduced, rather than inferring this reduction on the basis of improved performance. The major studies which have particular relevance to the problem of anxiety reduction will be reviewed briefly.

A series of studies have looked at performance in a test situation as a function of experimental treatment and trait or test anxiety. An early study by Smith and Rockett (1958) found that if students were allowed to comment on ambiguous or misleading items while taking a multiple choice exam, the performance of high anxious students was improved. The authors inferred from this result that commenting allowed high anxious students to discharge their tensions over the exam. Sinclair (1969) found that students in a low ego-involving situation in which they were told their performance was to be used for the purpose of improving the testing instrument did better than students in an ego-involving situation in which they were told they would be evaluated.

Other techniques which have improved the performance of high anxious students include: (1) the provision of feedback in programmed instruction (Campeau, 1968); (2) the opportunity to observe models first perform a task (Sarason, Pederson, & Nyman, 1968); (3) reassurance instructions to students that their performance will not be used in evaluation (Sarason, 1958); (4) relaxation instructions prior to a paired-associate learning task (Strauglin & Duford, 1969); and (5) the provision of memory support on problem solving and concept formation (Sieber & Kameya, 1967; Paulson, 1969; Sieber, Kameya & Paulson, 1970).

Most of the research studies which have been concerned with experimental treatments that reduce the disruptive effects of anxiety on performance have not measured state anxiety. In addition, they have used a behavioral or performance index from which anxiety reduction was inferred. Many of the treatments which have been employed have been shown to improve the performance of high anxious students and are thus suggestive of appropriate anxiety reduction techniques.

Our progress to date with respect to various anxiety reduction techniques will now be reviewed. The first section will deal with "The Effects of Stimulating State Epistemic Curiosity on State Anxiety and Performance in a Computer-Assisted Learning Task, and the second with "Effects of Various Anxiety Reduction Techniques within Computer-Based Intelligence Testing."

STUDY I

THE EFFECTS OF STIMULATING STATE EPISTEMIC CURIOSITY ON STATE ANXIETY AND PERFORMANCE IN A COMPUTER-ASSISTED LEARNING TASK

The purpose of this research was (a) to investigate the hypothesis that stimulating state epistemic curiosity within a complex computer-assisted instruction (CAI) task would reduce state anxiety and improve performance for students differing in levels of trait curiosity, trait anxiety, and response mode conditions, and (b) to further validate a theoretically-derived measure, the State Epistemic Curiosity Scale (SECS) (Leherissey, 1971), within a computer-assisted instruction situation. The importance of extending research in this direction relates to the need to investigate (a) the role of curiosity behaviors in motivation and learning, and (b) the experimental manipulation of curiosity within an individualized instructional program.

The conceptual framework which guided the development of the State Epistemic Curiosity Scale was derived in part from Berlyne's (1960) distinction between diversive and specific curiosity. Diversive curiosity refers to stimulus-seeking exploration induced by a state of boredom, specific curiosity to information-seeking exploration induced by a state of subjective uncertainty due to incomplete information about a particular stimulus. Berlyne (1960) further distinguishes between types of specific curiosity--perceptual and epistemic. Epistemic curiosity is aimed not only at acquiring sensory information, but also at acquiring knowledge in order to reduce conceptual conflict. As such, epistemic curiosity is related to thinking and problem-solving, and is assumed to lead to permanent storage of information (Berlyne, 1971).

Recent research suggests that curiosity enhances the acquisition of knowledge (e.g., Berlyne, 1971; Day, 1969). The particular type of curiosity most relevant to the learning process is epistemic or knowledge-seeking curiosity. A factor which has been identified as detrimental to both the arousal of curiosity behaviors and optimal performance within a learning task is anxiety (Day, 1967, 1969; Lester, 1968). In addition, research evidence indicates that an inverse relationship exists between curiosity and anxiety. That is, when levels of curiosity are high, levels of anxiety are relatively low (Day, 1969; Lester, 1968; Penny, 1965). These findings suggest that the precise measurement and experimental manipulation of curiosity during a learning task could lead to the discovery of optimal feeling states (i.e., curiosity) which could form the basis for optimal performance. The present research, therefore, was aimed at (a) explicating the effects of manipulating state epistemic curiosity on state anxiety and performance in a CAI task, and (b) establishing the feasibility and validity of computer-based measurement of state epistemic curiosity.

Researchers in the area of personality processes and learning have recognized the necessity of distinguishing between personality traits and states (e.g., Day, 1969; Spielberger, 1966). Whereas traits imply relatively stable personality predispositions, states imply transitory emotional conditions which fluctuate over time as a function of situational factors. As previously mentioned, CAI research has demonstrated that state variables (anxiety) are more closely related to behavior in a particular learning situation than trait variables (e.g., Leherissey, O'Neil, & Hansen, 1971a; O'Neil, Spielberger, & Hansen, 1969).

In particular, two recent CAI studies which use the same learning materials as the present study found differential effects of state anxiety for students in Reading (R) and Constructed Response (CR) modes (Leherissey, O'Neil, & Hansen, 1971 Leherissey, O'Neil, Heinrich, & Hansen, 1971b). That is, students in Constructed Response groups were found to have higher levels of state anxiety during the technical learning materials than students in the Reading groups in both studies. Furthermore, whereas superior performance was expected for the Constructed Response groups relative to the Reading groups on the basis of previous programmed instruction research with these materials (Tobias, 1969), no difference in the total technical posttest performance for the Constructed Response and Reading groups was found in either of the prior studies. These findings, plus the interactive relationship which has been found between curiosity and anxiety (Day, 1969; Lester, 1968) suggests that one possible means for reducing the disruptive effects of state anxiety on learning and performance may be to stimulate state epistemic curiosity in this CAI task.

This suggestion implies the need to take into account individual differences in both trait curiosity and trait anxiety, and the need for empirically sound measures of both trait and state curiosity and trait and state anxiety. A reliable and valid measure of trait and state anxiety, the State-Trait Anxiety Inventory (STAI), has been developed by Spielberger, Gorsuch and Lushene (1970). In addition, progress has been made in the development of a measure of trait curiosity, the Ontario Test of Intrinsic Motivation (OTIM) by Day (1969).

To date, however, no measures of state epistemic curiosity have appeared in the expermmental literature. In order to investigate the relationships between state curiosity, state anxiety, and performance, a measure of state epistemic curiosity, the State Epistemic Curiosity Scale, was developed by Leherissey (1971). The criteria which guided the development of the State Epistemic Curiosity Scale were related to the student's desire to: (1) know more about a learning task; (2) approach a novel or unfamiliar learning task; (3) approach a complex or ambiguous learning task; and (4) persist in information-seeking behavior in a learning task. Preliminary reliability and validity data collected in two studies indicated that the State Epistemic Curiosity Scale had high internal consistency, as well as substantial concurrent and construct validity (Leherissey, 1971).

The state of epistemic curiosity was conceptualized as a transitory motivational condition of the student, the arousal level of which was expected to vary across time, both with the nature of the specific learning task and the student's personality characteristics or predispositions. Thus, dependent upon the student's level of trait epistemic curiosity (i.e., relatively stable tendency to engage in specific knowledge-seeking behaviors under conditions of conceptual conflict) and past experiences with specific types of learning tasks, he would be expected to exhibit differential levels of state epistemic curiosity across time.

Thus, one of the purposes of this study was to further test the validity of the State Epistemic Curiosity Scale. A major purpose of the study was to investigate the effect of stimulating state epistemic curiosity on reduction of state anxiety in a CAI task.

Major predictions for the study were as follows: (a) high state curious students would have lower levels of state anxiety throughout the experimental task than low state curious students; (b) high state curious students would make more correct responses on the achievement measure than low state curious students; (c) students in the Curiosity Stimulating Instruction condition would have higher levels of state curiosity than students in the No Instruction condition; (d) high trait curious students would have higher levels of state curiosity than low trait curious students; and (e) levels of state curiosity would change during the task. In addition, based on the findings of prior studies, it was expected that students in the Constructed Response-No Instruction condition would have higher levels of state anxiety than students in the Reading-No Instruction condition during the experimental session.

Method

Subjects

One hundred and fifty-two female undergraduates were randomly assigned to the Reading-Curiosity Stimulating Instruction, Reading-No Instruction, Constructed Response-Curiosity Stimulating Instruction, and Constructed Response-No Instruction groups based on their levels of trait curiosity (low, high) and trait anxiety (low, high). These students were selected from a group of 222 students on the basis of their extreme trait curiosity and trait anxiety scores obtained in an initial group testing session. The cut-off scores were in the upper and lower 30% of the normative trait anxiety distribution and students were matched for extreme scores on the OTIM.

Apparatus

An IBM 1500 system presented learning materials, state epistemic curiosity and state anxiety scales. System terminals consisted on cathode-ray tube, light pen, and typewriter keyboard, and were located in a sound-deadened, air-conditioned room. The CAI system recorded all student responses to the learning materials and state affective measures during the CAI portion of the task.

Learning Program and Achievement Measures

The Reading and Constructed Response instructional programs covering technical materials on the diagnosis of myocardial infarction described by Leherissey, O'Neil, and Hansen (1971a) were presented via CAI. The learning material and posttest were divided into initial technical and remaining technical sections. Both verbal and graphical (e.g., EKG drawings and tracings) frames were included in these technical materials.

The basic learning program was divided into two versions, each containing the same subject matter and frame structure: (a) Reading, and (b) Constructed Response. In the Reading versions of the instructional program, students were not required to make any overt responses, but merely to read each frame successively. In the Constructed Response versions, overt responses were required in the form of a typed word for response blanks on the technical verbal materials. On the technical graphical materials containing EKG drawings, students "constructed" EKG trainings by special program coding (Leherissey et al., 1971a).

The Reading and Constructed Response versions were modified for the curiosity stimulating condition by the insertion of special curiosity stimulating instructions (Curiosity Stimulating Instruction condition) following a brief introduction to the learning task. The Curiosity Stimulating Instruction materials were presented in three instructional frames which stated:

FRAME 1

DID YOU KNOW THAT--

HEART DAMAGE CAUSES MORE THAN HALF OF ALL DEATHS IN THIS COUNTRY?

MAJOR TYPES OF HEART DAMAGE CAN BE IDENTIFIED BY ELECTROCARDIOGRAM TRACINGS?

THE STAGES OF RECOVERY FROM HEART DAMAGE CAN BE TRACED BY AN ELECTROCARDIOGRAM?

ALTHOUGH YOU MAY KNOW THE GENERAL FACTS ASSOCIATED WITH THE ABOVE STATEMENTS, THE PRECISE MEDICAL KNOWLEDGE CONCERNING HEART DAMAGE AND ITS DIAGNOSIS IS PROBABLY NEW TO YOU.

FRAME 2

FOR EXAMPLE, DO YOU KNOW-

1) THE MEDICAL NAME FOR THE HEART MUSCLE?
2) THE MEDICAL NAMES FOR THE THREE MAJOR TYPES OF HEART DAMAGE?
3) HOW AN ELECTROCARDIOGRAM TRACING IS OBTAINED?
4) HOW HEART DAMAGE IS DIAGNOSED BY AN ELECTROCARDIOGRAM TRACING?
5) HOW LONG IT TAKES TO RECOVER FROM MAJOR HEART DAMAGE?

FRAME 3
THE ANSWERS TO THOSE QUESTIONS AND MANY OTHERS ARE GIVEN IN THE INSTRUCTIONAL MATERIALS YOU ARE ABOUT TO LEARN. FOR EXAMPLE, YOU WILL LEARN THE MEDICAL TERMS FOR HEART DAMAGE, HOW ELECTROCARDIOGRAM TRACINGS ARE RECORDED, HOW TO DIFFERENTIATE BETWEEN ELECTROCARDIOGRAM TRACINGS, AND THE STAGES IN THE HEALING PROCESS.

In the No Instruction condition, students were told to take a one-minute break, which was the average time required to read the curiosity stimulating instructions.

Affective Measures

The State Epistemic Curiosity Scale (Leherissey, 1971) was used to assess state epistemic curiosity periodically throughout the learning task. The Ontario Test of Intrinsic Motivation (Day, 1969) was used to match subjects on trait curiosity levels. The State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970) was used to measure both state and trait anxiety.

The trait curiosity and trait anxiety scales were administered during the initial group testing session with standard instructions, i e., "indicate how you generally feel." The long form (20-item) state epistemic curiosity and long form (20-item) state anxiety scales were also administered during the initial group testing session with instructions, "indicate how you think you would feel while learning new materials." The short form (5-item) state anxiety scales and short form (7-item) state curiosity scales were administered a total of six times during the CAI task: immediately after the introduction to the learning materials, after the Curiosity Stimulating Instruction or No Instruction conditions, after the first and second halves of the initial technical learning materials, and after the first and second halves of the remaining technical learning materials. The state curiosity and state anxiety scales followed each other in random order between presentations, and individual items on each scale were also randomly ordered within presentations. In addition, the short form State-Trait Anxiety Inventory state anxiety scale and 20-item State Epistemic Curiosity Scale were administered after the achievement posttest via paper and pencil.

The state curiosity and state anxiety scales presented after the introduction to the learning materials and after the Curiosity Stimulating Instruction or No Instruction conditions were presented with standard instructions, i.e., "indicate how you feel right now"; the remaining state curiosity and state anxiety scales presented during the CAI task were administered with retrospective state instructions, i.e., "indicate how you felt during the task you have just finished." The short form State-Trait Anxiety Inventory state anxiety scale administered immediately after the posttest instructed students to "indicate how you felt while take the posttest," and the 20-item State Epistemic Curiosity Scale administered at the end of the experimental session instructed students to "indicate how you felt while you were learning the instructional materials."

Procedure

The experimental session was divided into three periods: (1) a pretask period, during which students were assigned to response mode and instruction conditions, and read instructions on the operation of the CAI terminal; (2) a performance period, during which students received differential instructions (Curiosity Stimulating Instruction or No Instruction conditions), learned the technical CAI materials, and took six combined short form (7-item) state curiosity and short form (5-item) state anxiety scales at six equal intervals within the learning materials; (3) a posttask period, during which students were administered the achievement posttest, short form state anxiety scale, and long form (20-item) State Epistemic Curiosity Scale. Each of these periods is further described below.

1. Pretask Period. Upon arrival at the CAI Center, students were assigned to one of four experimental conditions based upon their level of trait curiosity and level of trait anxiety to insure an equal number of students in each group. These four conditions were (a) Reading with Curiosity Stimulating Instructions, (b) Reading without Curiosity Stimulating Instructions, (c) Constructed Response with Curiosity Stimulating Instructions, and (d) Constructed Response without Curiosity Stimulating Instructions. Following assignment to experimental treatments, students were asked to read written instructions on the operation of the CAI terminal. The students were run in small groups of 8 to 13 in a total of 13 experimental sessions, each of which was proctored by two to four experimenters.

2. Performance Period. All students were seated at CAI terminals and after "signing on," were presented short introductory materials on the general nature of the experimental task. The students were then presented the first short form state curiosity and state anxiety scales. Depending upon whether students were in the Curiosity Stimulating Instruction or No Instruction conditions, they received differential instructions, followed by the second combined state curiosity and state anxiety scales. The students were then presented with differential instructions as to how they should proceed through the learning task, depending upon whether they were in the Reading or Constructed Response mode groups. All students were instructed to proceed through the materials at their own rate; specific instructions given to each response mode group are reported in detail by Leherissey et al. (1971a). The Constructed Response group was given practice in the operation of the keyboard and on the "enter" and "erase" functions. On the technical graphical materials, the Constructed Response group was given a handout of 10 possible EKG tracing segments and instructed to type in the combination of numbers from 0 to 9 which would complete the appropriate tracing (Leherissey et al., 1971a). During this performance period, all students were presented the short form state curiosity and state anxiety scales with retrospective state instructions at four points in the instructional program: (a) following the first half of the initial technical materials; (b) following the second half of the initial technical materials; (c) following the first half of the remaining technical materials; and (d) following the second half of the remaining technical materials.

3. Posttask Period. After each student had completed the instructional program and final state curiosity and state anxiety scale, he was taken to another room and given a posttest package. Included in the posttest package were the technical portion of the posttest and a short form state anxiety scale with retrospective state instructions. In addition, all students were given a handout of the 10 possible EKG tracing segments and instructed to use this handout when they were required to "draw" EKG tracings on the posttest. The students, therefore, chose the appropriate sequence of numbers to construct particular tracings, rather than actually drawing these tracings. After completion of the posttest package, students were asked to respond to the 20-item State Epistemic Curiosity Scale. The students were then informed that the task was quite difficult and reassured that their performance was satisfactory. They were also given some additional information concerning the nature of the experiment and cautioned not to discuss the experiment with their classmates.

Results

State Curiosity Data

In order to investigate the hypotheses that (a) high trait curious students would have higher levels of state curiosity during the CAI task than low trait curious students, (b) levels of state curiosity would change over time, and (c) students in the Curiosity Stimulating Instruction conditions would have higher levels of state curiosity than students in the No Instruction conditions, a 2 x 2 x 2 x 6 ANOVA with repeated measures on the last factor was calculated. Independent variables in this analysis were response mode conditions (Reading, Constructed Response), instruction conditions (Curiosity Stimulating Instruction, No Instruction), levels of trait curiosity (low, high), and measurement periods (six in-task periods). The dependent variable was mean state curiosity during the task. The means and standard deviations of low and high trait curious students in response mode and instruction conditions on the six in-task State Epistemic Curiosity Scale measures are reported in Table 1.

Results of this analysis revealed two significant interactions: (a) response mode by levels of trait curiosity by measurement periods ($F = 2.60$, $df = 5/720$, $p < .05$); and (b) response modes by measurement periods ($F = 5.30$, $df = 5/720$, $p < .001$). The triple interaction, shown in Figure 1, indicated that low trait curious students in both the Reading and Constructed Response groups had lower levels of state curiosity throughout the task than high trait curious students; however, the sharpest decreases in state curiosity scores across time were noted for the Constructed Response groups relative to the Reading groups. In addition, whereas high trait curious students in the Constructed Response groups decreased in state curiosity scores to a level comparable to low trait curious students in the Reading group by the end of the CAI task, high trait curious students in the Reading groups retained a relatively high level of state curiosity throughout the CAI task.

Table 1

Mean State Curiosity Scores on the Six In-Task SCS Measures for Low and High Trait Curious Students in Response Mode and Instruction Conditions

Groups	Pre Instructions	Post Instructions	Measurement Periods Post T_I1	Post T_I2	Post T_R1	Post T_R2
All groups (n=152)						
Mean	24.21	24.20	24.41	22.31	22.49	20.30
SD	3.14	3.45	3.77	4.97	5.04	5.84
Reading - CSI						
LC (n=19)						
Mean	22.74	23.63	25.16	20.74	22.37	19.21
SD	2.75	3.62	3.24	5.33	5.36	6.28
HC (n=19)						
Mean	25.74	25.58	24.90	23.68	25.00	23.47
SD	2.02	2.46	3.45	3.73	3.62	3.92
Reading - NI						
LC (n=19)						
Mean	23.95	24.11	24.11	21.16	22.53	20.11
SD	2.97	3.11	3.41	5.54	4.78	5.32
HC (n=19)						
Mean	24.00	23.68	25.68	23.47	23.42	23.47
SD	3.13	3.53	3.06	3.86	3.27	4.27
Constructed Response - CSI						
LC (n=19)						
Mean	23.00	24.42	22.84	22.11	21.21	19.32
SD	3.74	3.98	3.98	4.84	5.42	5.06
HC (n=19)						
Mean	25.58	24.48	23.95	22.58	21.74	18.58
SD	2.67	3.81	4.34	4.11	4.41	6.62
Constructed Response - NI						
LC (n=19)						
Mean	22.84	22.58	23.37	20.11	19.37	17.58
SD	3.75	3.85	4.72	6.64	6.39	6.11
HC (n=19)						
Mean	25.84	26.05	25.26	24.63	24.32	20.68
SD	2.04	1.84	3.45	4.07	4.92	6.58

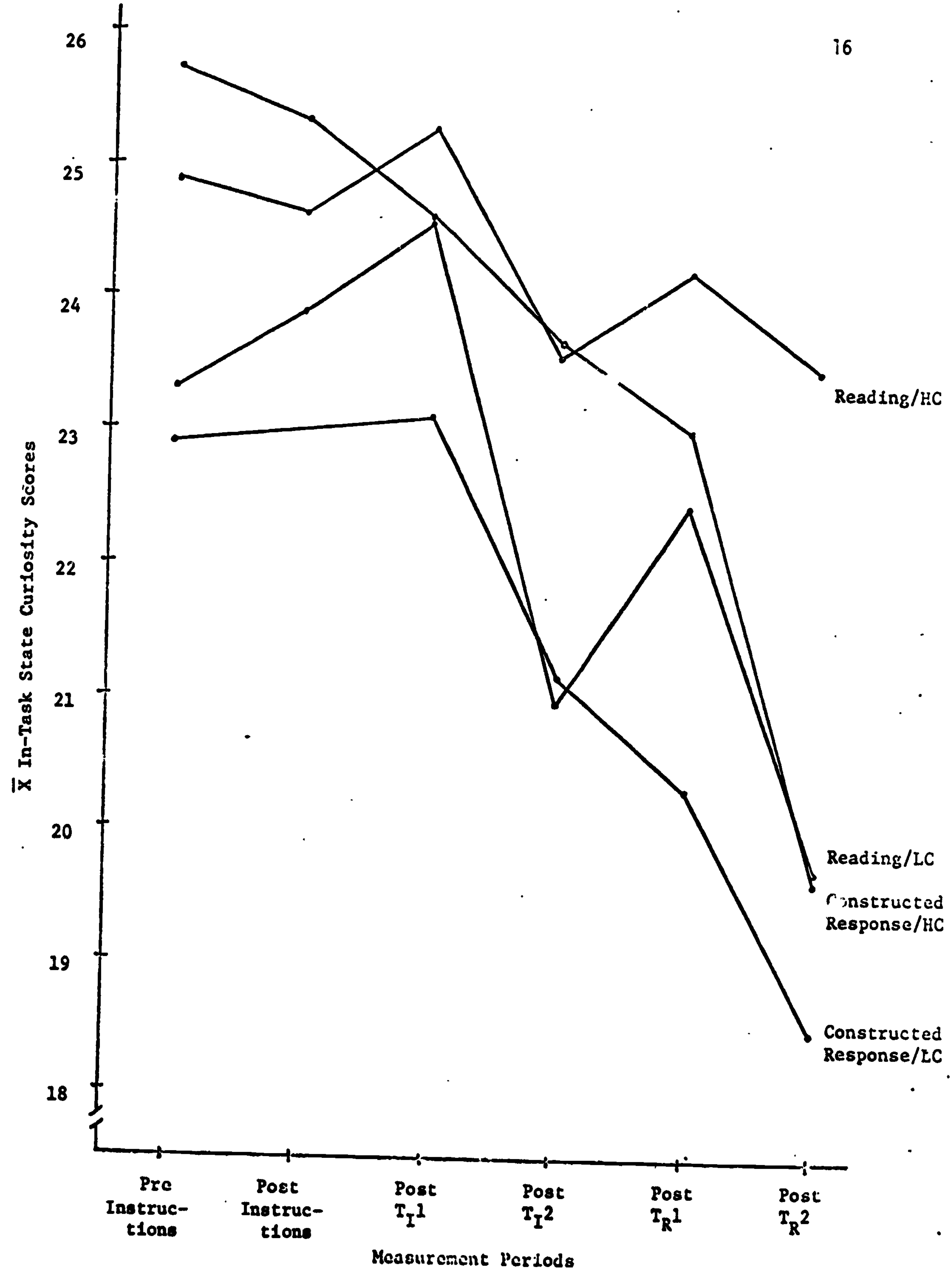

Figure 1.--Response modes by levels of trait curiosity by measurement periods interaction on state curiosity scores.

The response modes by measurement periods interaction, which is shown in Figure 2, indicated that the Constructed Response groups had the sharpest decline in state curiosity during the CAI task relative to the Reading groups. All groups were also found to have decreases in state curiosity across the six measurement periods ($F = 44.48$, $df = 5/720$, $p < .001$) and high trait curious students had higher state curiosity scores than low trait curious students ($F = 13.50$, $df = 1/144$, $p < .001$). The main effect of instruction conditions was not found to be significant.

State Anxiety Data

In order to determine if (a) high trait anxious students would have higher levels of state anxiety than low trait anxious students throughout the CAI task, and (b) levels of state anxiety would change over time, a 2 x 2 x 2 x 6 ANOVA with repeated measures on the last factor was calculated. The independent variables in this analysis were response modes (Reading, Constructed Response), instruction conditions (Curiosity Stimulating Instruction, No Instruction), levels of trait anxiety (low, high) and measurement periods (six in-task state anxiety measures). The dependent variable was mean state anxiety during the task. The means and standard deviations of low and high trait anxious students in response mode and instruction conditions on the six state anxiety measures are reported in Table 2.

Results of this analysis revealed two significant interactions: (a) response modes by measurement periods by levels of trait anxiety ($F = 2.48$, $df = 5/720$, $p < .05$), and (b) response modes by measurement periods ($F = 11.37$, $df = 5/720$, $p < .001$). The triple interaction is shown in Figure 3 and indicates that whereas high and low trait anxious students in the Constructed Response groups had relatively the same pattern of increases and decreases in state anxiety during the CAI task, high trait anxious students in the Reading groups had steady decreases in state anxiety across measurement periods, while low trait anxious students in the Reading groups had decreases in state anxiety during the initial technical materials and slight increases in state anxiety during the remaining technical materials.

The response modes by measurement periods interaction shown in Figure 4 indicated that state anxiety scores for students in the Reading groups decreased during the CAI task, whereas state anxiety scores for students in the Constructed Response groups tended to increase across time. In addition, students in the Constructed Response groups were found to have higher state anxiety scores than students in the Reading groups ($F = 4.53$, $df = 1/144$, $p < .05$), high trait anxious students had higher state anxiety scores than low trait anxious students ($F = 38.28$, $df = 1/144$, $p < .001$), and state anxiety scores significantly changed across the six measurement periods ($F = 14.50$, $df = 5/720$, $p < .001$). That is, all groups had highest levels of state anxiety initially, lowest levels of state anxiety following the Curiosity Stimulating Instruction and No Instruction conditions, and increases in state anxiety following the initial technical material. A subsidiary ANOVA indicated that the decrease in state anxiety after the Curiosity Stimulating Instruction and No Instruction conditions was significant ($F = 81.63$, $df = 1/144$, $p < .001$); however, no differential effects of instruction conditions were found.

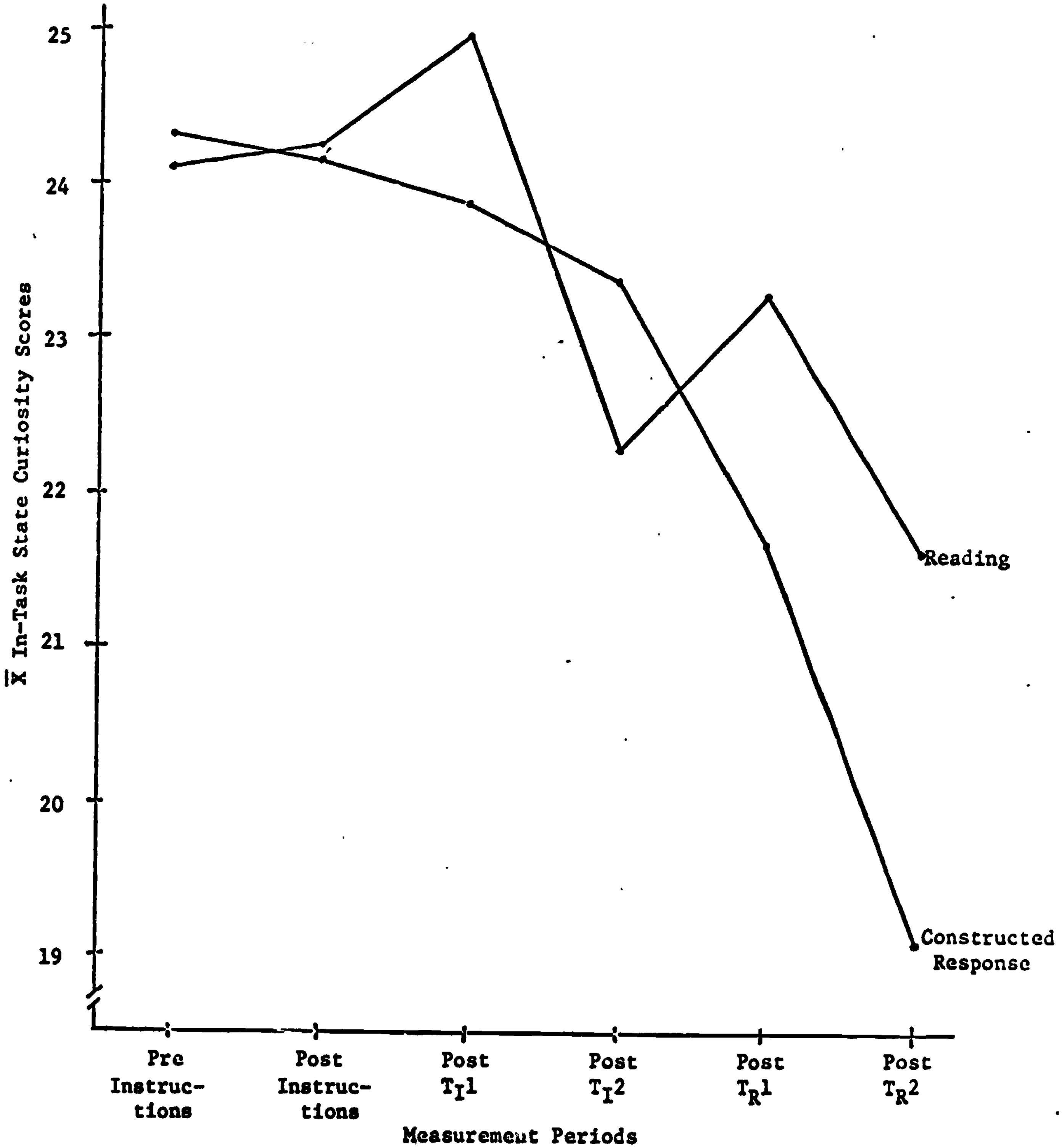

Figure 2.--Response modes by measurement periods interaction on state curiosity scores.

TABLE 2

Mean A-State Scores on Six In-Task STAI A-State Measures for Low and High A-Trait Students in Response Mode and Instruction Conditions

	Groups	Pre-Instructions	Post-Instructions	Measurement Periods Post T_I1	Post T_I2	Post T_R1	Post T_R2
	All groups (n=152)						
	Mean	9.70	7.95	8.90	8.16	8.38	8.17
	SD	3.39	2.83	3.23	3.27	3.36	3.36
Reading - CSI	LA (n=19)						
	Mean	8.47	7.05	6.68	6.16	6.37	6.79
	SD	3.39	2.35	2.08	1.86	1.86	2.66
	HA (n=19)						
	Mean	11.90	10.26	9.84	8.74	8.37	8.00
	SD	3.64	2.86	3.64	3.45	3.29	2.62
Reading - NI	LA (n=19)						
	Mean	7.47	6.63	6.90	6.37	6.21	6.47
	SD	2.72	2.14	2.31	1.74	1.58	2.26
	HA (n=19)						
	Mean	11.63	9.74	9.21	9.26	8.84	7.84
	SD	2.85	3.00	2.89	3.58	3.29	2.71
Constructed Response - CSI	LA (n=19)						
	Mean	8.68	6.53	9.16	7.37	8.16	7.58
	SD	2.89	1.58	3.32	2.77	2.57	2.55
	HA (n=19)						
	Mean	10.16	7.79	9.90	9.63	9.63	10.00
	SD	2.85	2.37	3.23	3.29	3.55	4.16
Constructed Response - NI	LA (n=19)						
	Mean	8.37	6.58	8.74	8.00	8.32	8.05
	SD	3.13	2.59	3.57	3.50	3.28	3.64
	HA (n=19)						
	Mean	10.90	9.05	10.74	9.74	11.11	10.32
	SD	3.07	2.80	2.58	3.62	4.28	4.23

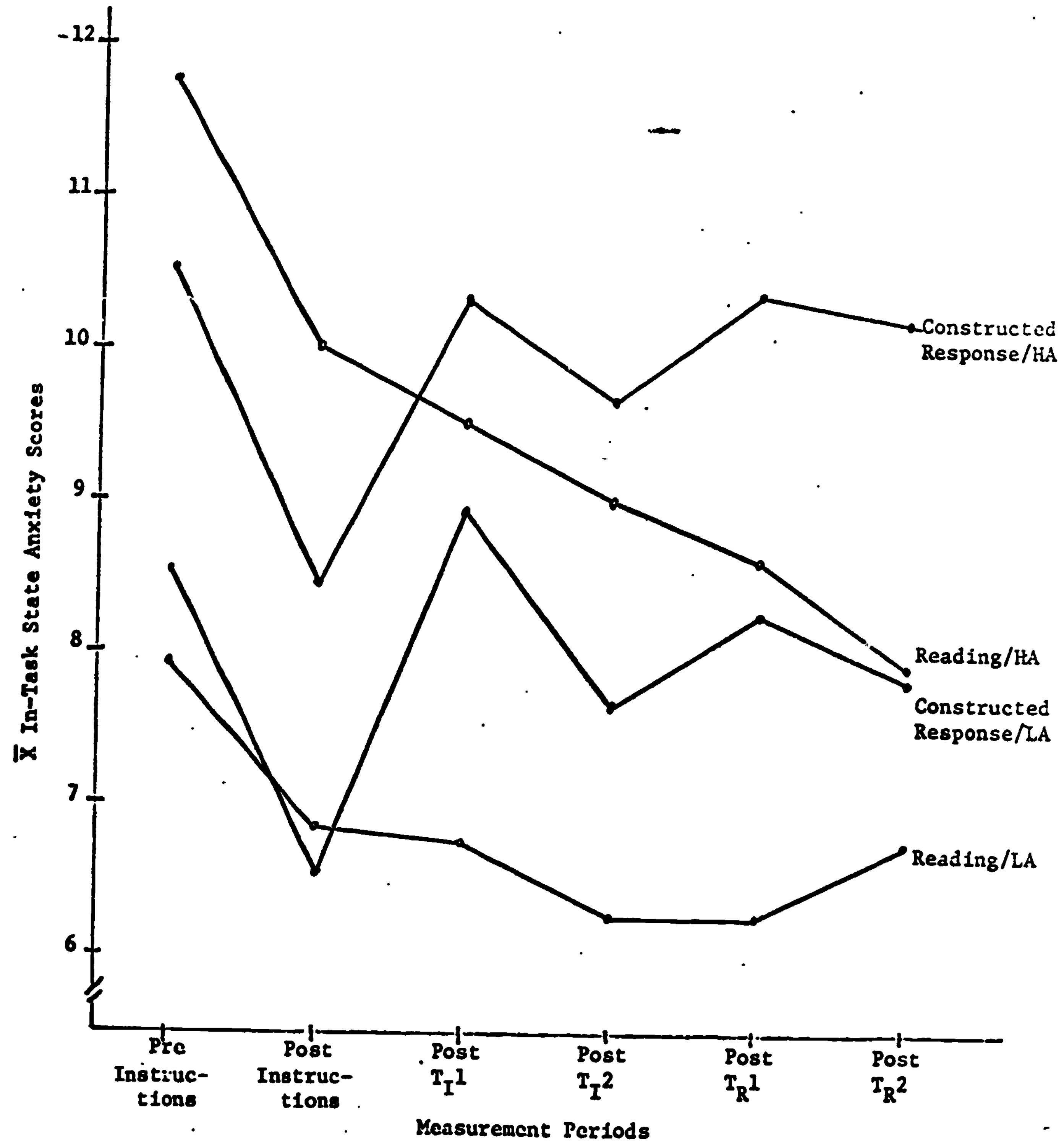

Figure 3.--Response modes by levels of trait anxiety by measurement periods interaction on in-task state anxiety scores.

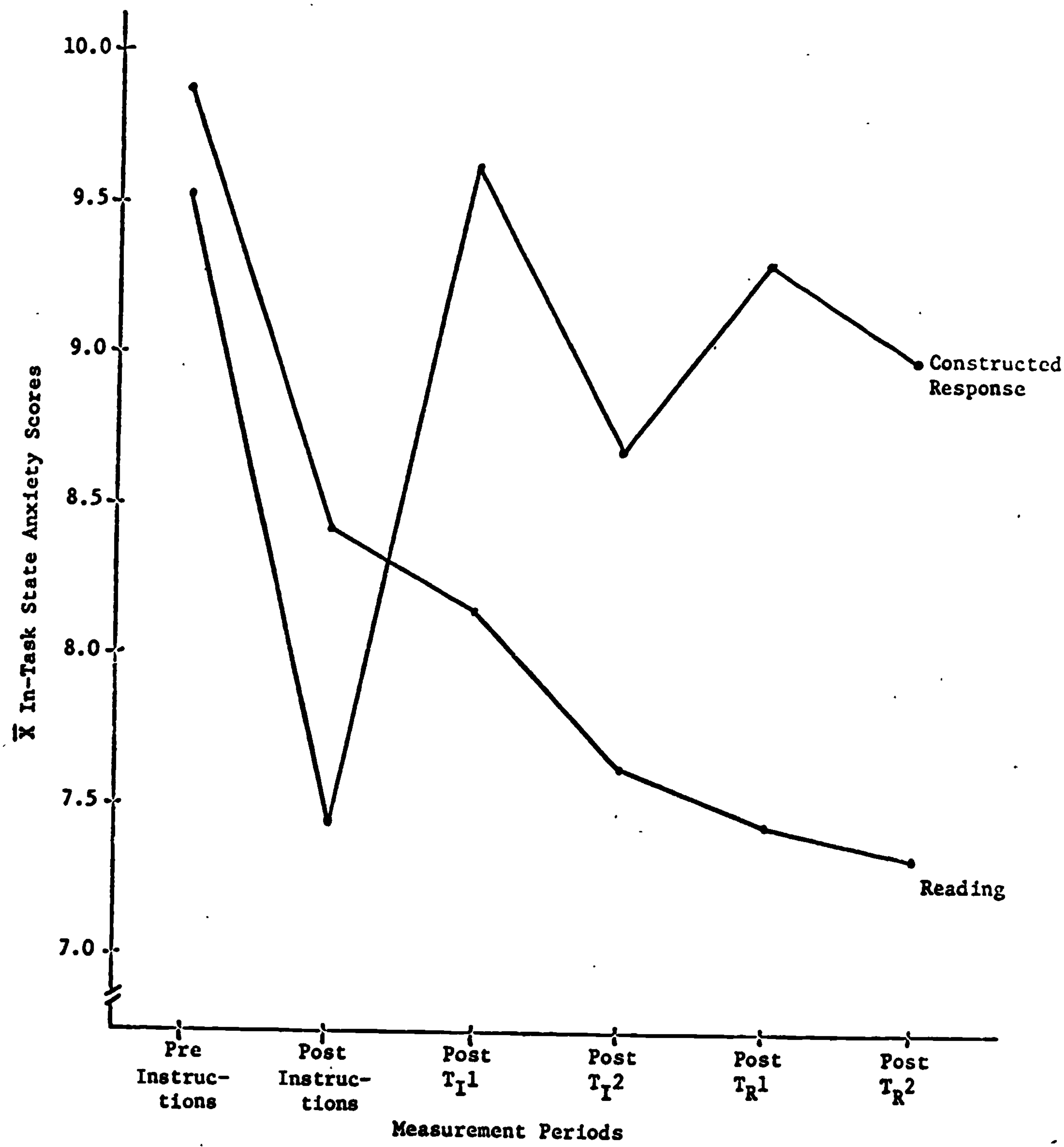

Figure 4.--Response modes by measurement periods interaction on in-task state anxiety scores.

Performance Data

Several ANOVAS were calculated to determine the effects of trait curiosity and state curiosity on the posttest achievement measure. An analysis of initial technical posttest scores for students differing in trait curiosity, response modes, and instruction conditions revealed a significant interaction between trait curiosity levels and instruction conditions ($F = 4.34$, $df = 1/144$, $p < .05$) which is shown in Figure 5. The interaction indicated that low trait curious students in the Curiosity Stimulating Instruction conditions performed better than low trait curious students in the No Instruction conditions, whereas there was relatively little difference in the performance of high trait curious students in the Curiosity Stimulating Instruction and No Instruction conditions. In addition, students in the Constructed Response groups performed better than students in the Reading groups ($F = 27.62$, $df = 1/144$, $p < .001$) on the initial technical portion of the posttest. A similar analysis on the remaining technical posttest revealed no significant main effect or interactions. The means and standard deviations for low and high trait curious students in response mode and instruction conditions on the initial technical and remaining technical posttest are presented in Tables 3 and 4 respectively.

Results of the second set of ANOVAS as a function of posttask state curiosity levels on initial technical posttest performance also revealed that high state curious students made more correct responses than medium or low state curious students ($F = 7.83$, $df = 2/130$, $p < .001$). Results of the second set of analyses on remaining technical posttest performance revealed two significant interactions: (a) response mode by instruction conditions ($F = 5.52$, $df = 1/140$, $p < .05$), and (b) instruction conditions by posttask state curiosity levels ($F = 3.85$, $df = 2/140$, $p < .05$).

The response mode by instruction conditions interaction shown in Figure 6 indicated that whereas students in the Constructed Response-Curiosity Stimulating Instruction condition performed better than students in the Constructed Response-No Instruction group on the remaining technical posttest, the reverse was true for students in the Reading-Curiosity Stimulating Instruction and Reading-No Instruction groups. The instruction conditions by state curiosity level interaction shown in Figure 7 indicated that whereas there was relatively little difference in the performance of low, medium, and high state curious students in the Curiosity Stimulating Instruction conditions, for students in the No Instruction conditions, high state curious students performed better than medium or low state curious students on the remaining technical posttest. In addition, high posttask state curious students were found to make more correct responses than medium or low posttask state curious students ($F = 12.65$, $df = 2/140$, $p < .001$). The means and standard deviations of low, medium, and high posttask curious students in response mode and instruction conditions on the initial technical and remaining technical posttest are reported in Tables 5 and 6 respectively. Students were divided into low, medium, and high curious groups on the basis of their scores on the 20-item State Epistemic Curiosity Scale measure given at the end of the experimental session. The range of low posttask state curiosity scores was 26-53; medium posttask state curiosity scores ranged from 54-67; the range of high posttask curiosity scores was 68-80.

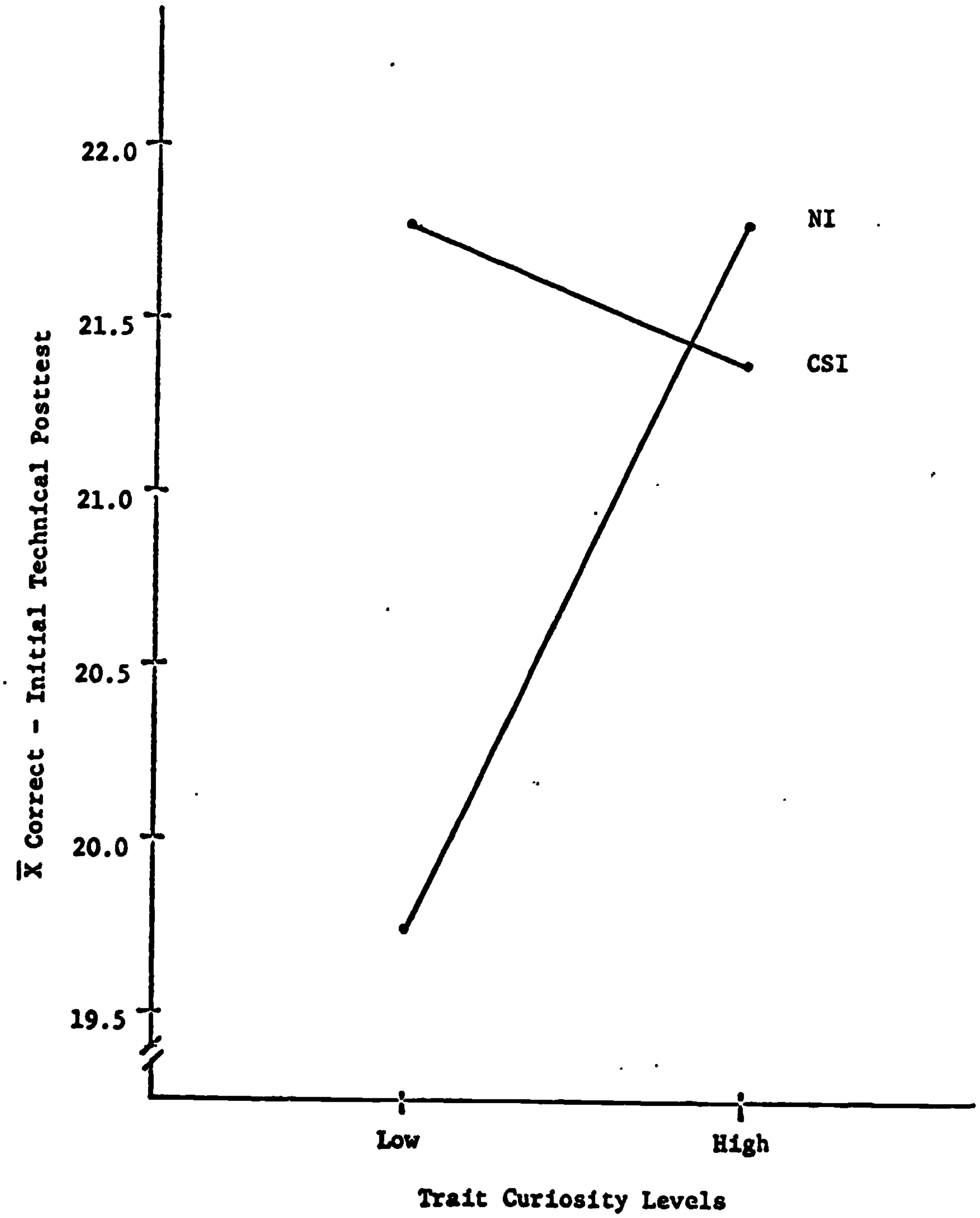

Figure 5.--Instruction conditions by trait curiosity levels interaction on initial technical posttest scores.

TABLE 3

Mean Correct Responses on the Initial Technical Posttest for Low and High Trait Curious Students in Response Mode and Instruction Conditions

Groups	Trait Curiosity Levels	
	Low	High
Reading - CSI (n=38)		
Mean	19.42	20.58
SD	3.53	5.00
Reading - NI (n=38)		
Mean	18.59	20.00
SD	3.12	3.04
Constructed Response - CSI (n=38)		
Mean	24.11	22.16
SD	2.54	4.03
Constructed Response - NI (n=38)		
Mean	20.90	23.53
SD	3.78	2.65

TABLE 4

Mean Correct Responses on the Remaining Technical Posttest for Low and High Trait Curious Students in Response Mode and Instruction Conditions

Groups	Trait Curiosity Levels	
	Low	High
Reading - CSI (n=38)		
Mean	32.95	35.05
SD	17.40	16.97
Reading - NI (n=38)		
Mean	35.32	38.42
SD	14.36	18.69
Constructed Response - CSI (n=38)		
Mean	48.84	39.11
SD	13.66	18.42
Constructed Response - NI (n=38)		
Mean	34.42	40.05
SD	21.79	14.18

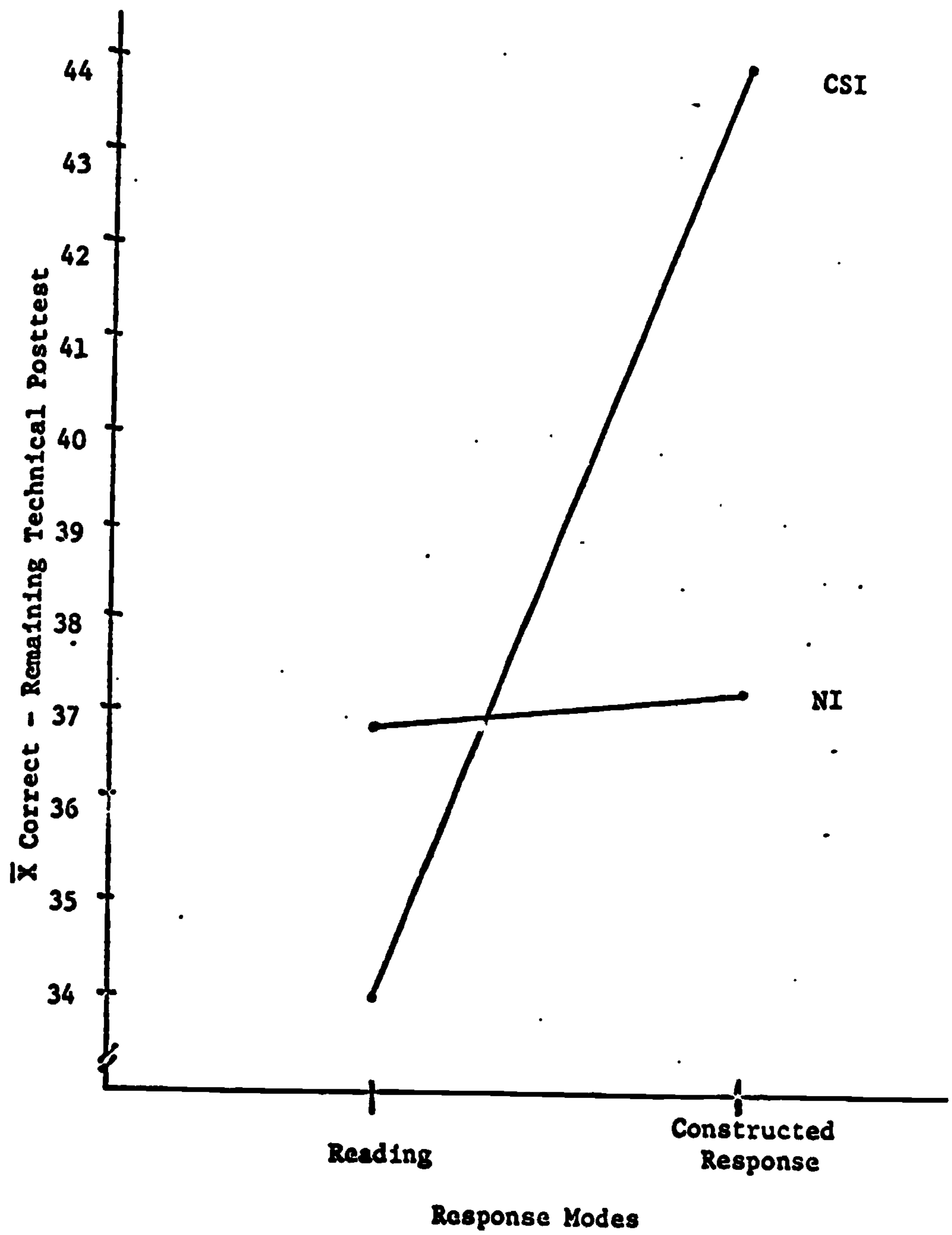

Figure 6.--Instruction conditions by response modes interaction on remaining technical posttest scores.

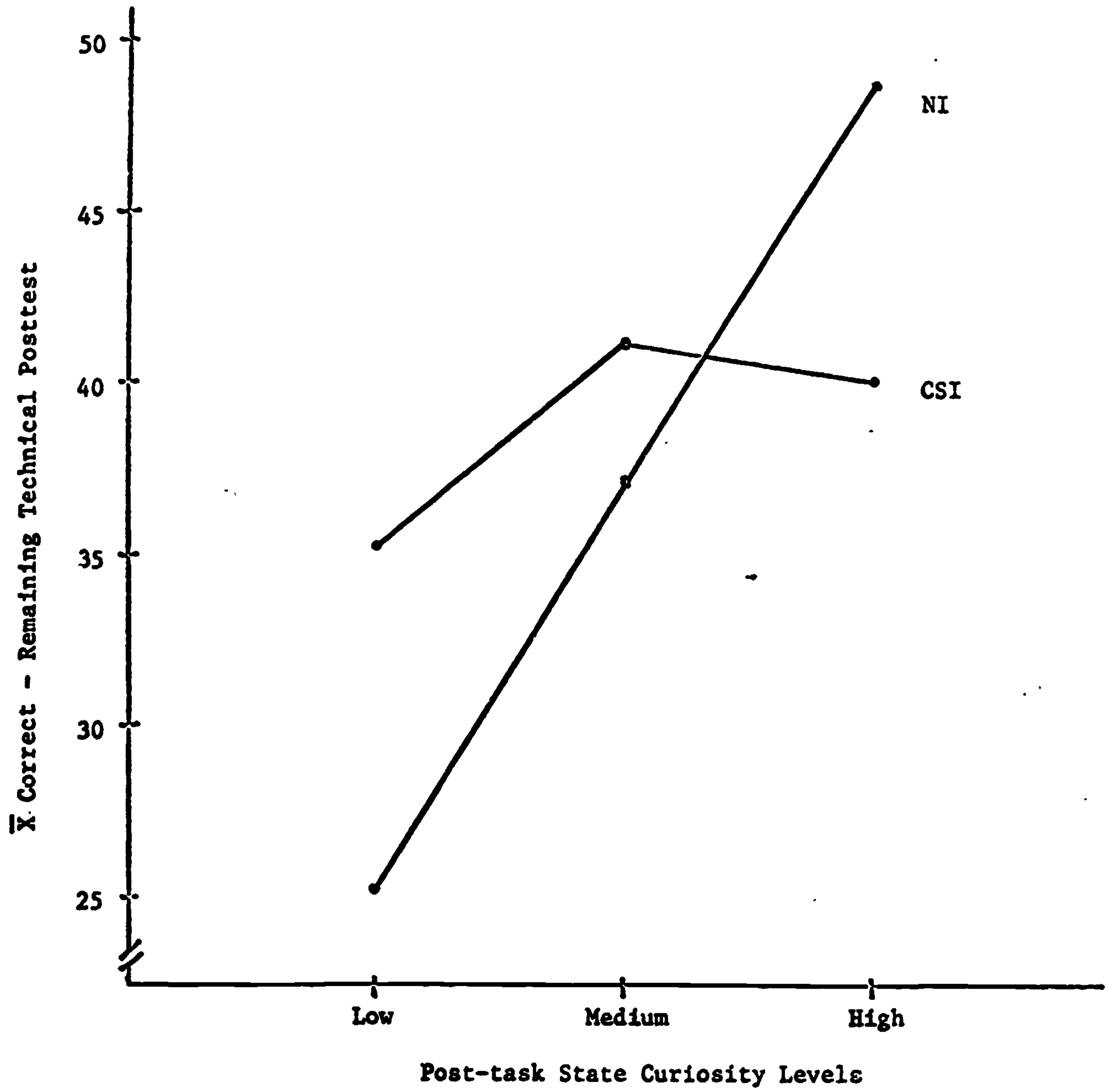

Figure 7.--Instruction conditions by levels of post-task state curiosity levels on remaining technical posttest scores.

TABLE 5

Mean Correct Responses on the Initial Technical Posttest for Low, Medium, and High Post-Task State Curious Students in Response Mode and Instruction Conditions

Groups	Post-Task State Curiosity Level Low	Medium	High
Reading - CSI (n=38)			
Mean	19.80	20.13	20.00
SD	1.99	4.52	5.61
Reading - NI (n=38)			
Mean	17.09	19.44	21.67
SD	3.48	2.20	2.55
Constructed Response - CSI (n=38)			
Mean	22.25	23.31	24.44
SD	3.34	4.37	1.59
Constructed Response - NI (n=38)			
Mean	19.79	22.75	24.06
SD	3.75	2.38	2.44

TABLE 6

Mean Correct Responses on the Remaining Technical Posttest for Low, Medium, and High Post-Task State Curious Students in Response Mode and Instruction Conditions

Groups	Post-Task State Curiosity Level Low	Medium	High
Reading - CSI (n=38)			
Mean	31.40	36.67	32.92
SD	19.78	15.99	16.80
Reading - NI (n=38)			
Mean	23.64	35.56	55.67
SD	10.45	14.11	7.75
Constructed Response - CSI (n=38)			
Mean	37.75	47.00	50.67
SD	18.71	14.38	13.65
Constructed Response - NI (n=38)			
Mean	26.71	40.75	44.69
SD	17.01	12.31	18.46

Also of interest in the present study was an examination of the effects of state anxiety levels on posttest performance. Results of an ANOVA on initial technical posttest performance as a function of posttest state anxiety scores revealed a significant interaction between instruction conditions and levels of state anxiety ($F = 3.24$, $df = 2/140$, $p < .05$), shown in Figure 8. This interaction indicated that whereas there was relatively little difference in the performance of low state anxious students in the Curiosity Stimulating Instruction and No Instruction groups, high posttest state anxiety students in the Curiosity Stimulating Instruction group performed better than high posttest state anxiety students in the No Instruction groups. In contrast, medium state anxiety students in the No Instruction groups were found to perform better than medium state anxiety students in the Curiosity Stimulating Instruction groups. The means and standard deviations for low, medium, and high posttest state anxiety students in response mode and instruction conditions on the initial technical and remaining technical posttest are reported in Tables 7 and 8 respectively.

In addition to those results concerning the relationship between curiosity, anxiety, and performance and the effect of curiosity stimulating instructions and differing response modes, additional data were collected on the reliability and validity of the State Epistemic Curiosity Scale. This was in accordance with the second purpose of this study, that of establishing the feasibility and validity of the computer-based measurement of state epistemic curiosity.

Reliability Data

The means, standard deviations, and alpha reliability coefficients for the 20-item pretask State Epistemic Curiosity Scale, six short form in-task State Epistemic Curiosity Scale, and 20-item posttask State Epistemic Curiosity Scale measures are reported in Table 9.

As Table 9 indicates, the alpha reliabilities of the State Epistemic Curiosity Scale ranged from a low of .81 to a high of .94, indicating high internal consistencies on both the short and long forms of the State Epistemic Curiosity Scale. In addition, the State Epistemic Curiosity Scale showed generally increased internal consistencies from the first in-task to the sixth in-task measurement periods.

Item-remainder correlations for the individual items on each State Epistemic Curiosity Scale were also calculated. Table 10 gives the means, standard deviations, and item-remainder correlations for individual items of the pretask and posttask State Epistemic Curiosity Scale. The means, standard deviations, and item-remainder correlations for the individual items on the six in-task State Epistemic Curiosity Scales are reported in Table 11.

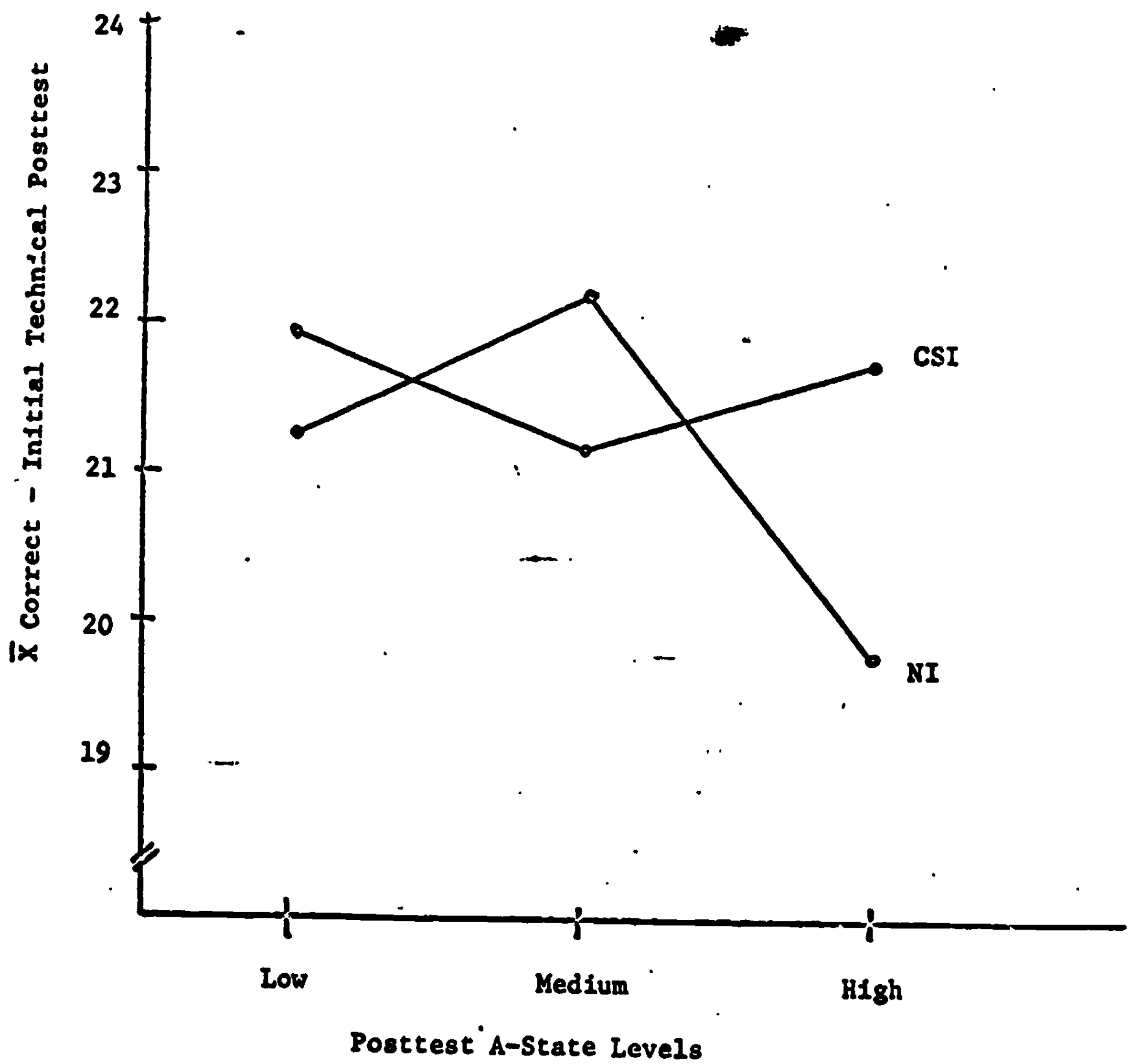

Figure 8.--Instruction conditions by posttest A-State levels interaction on initial technical posttest scores.

TABLE 7

Mean Correct Responses on the Initial Technical Posttest for Low, Medium, and High Posttest A-State Students in Response Mode and Instruction Conditions

Groups	Posttest A-State Level		
	Low	Medium	High
Reading - CSI (n=38)			
Mean	19.92	19.36	20.83
SD	5.45	4.01	3.69
Reading - NI (n=38)			
Mean	19.14	20.27	18.62
SD	3.90	2.41	2.69
Constructed Response - CSI (n=38)			
Mean	24.40	22.61	22.80
SD	1.58	4.27	3.16
Constructed Response - NI (n=38)			
Mean	22.83	24.44	20.59
SD	3.66	1.51	3.45

TABLE 8

Mean Correct Responses on the Remaining Technical Posttest for Low, Medium, and High Posttest A-State Students in Response Mode and Instruction Conditions

Groups	Posttest A-State Level		
	Low	Medium	High
Reading - CSI (n=38)			
Mean	31.17	34.36	36.42
SD	17.07	16.96	17.99
Reading - NI (n=38)			
Mean	34.57	40.00	36.69
SD	18.52	17.16	14.46
Constructed Response - CSI (n=38)			
Mean	51.00	43.00	38.70
SD	11.79	16.57	20.16
Constructed Response - NI (n=38)			
Mean	40.67	42.67	31.94
SD	23.83	8.41	17.27

TABLE 9

Means, Standard Deviations, and Alpha Reliabilities of the Eight State Epistemic Curiosity Scales Administered During Experiment (N=152)

Scale	Scale Range	Mean	SD	Alpha
Pre-task SECS	20-80	63.06	8.07	.88
First In-task SECS	7-28	24.21	3.13	.81
Second In-task SECS	7-28	24.20	3.60	.88
Third In-task SECS	7-28	24.41	3.76	.86
Fourth In-task SECS	7-28	22.31	4.96	.91
Fifth In-task SECS	7-28	22.49	5.02	.91
Sixth In-task SECS	7-28	20.30	5.82	.93
Total In-task SECS	42-168	138.39	21.24	.96
Post-task SECS	20-80	59.00	12.56	.94

TABLE 10

Means, Standard Deviations, and Item-Remainder Correlations for the Pretask and Posttask State Epistemic Curiosity Scales (N=152)

	Pretask SECS			Posttask SECS		
Item	Mean	SD	Item Remainder	Mean	SD	Item Remainder
1	3.37	.63	.58	3.15	.82	.80
2	3.30	.71	.67	3.10	.85	.84
3	3.49	.60	.42	3.25	.82	.69
4	3.04	.76	.62	2.48	1.03	.71
5	3.06	.79	.28	3.21	.90	.62
6	3.41	.70	.62	3.02	.98	.87
7	3.18	.71	.47	2.71	1.08	.57
8	3.29	.72	.44	3.22	.80	.60
9	2.67	.70	.32	2.76	.93	.60
10	3.41	.64	.60	3.04	.83	.85
11	3.02		.42	3.18	.83	.80
12	3.35	.69	.52	3.50	.75	.42
13	3.50	.69	.46	2.99	.99	.67
14	3.15	.83	.41	2.93	1.04	.32
15	3.00	.77	.41	2.99	.96	.62
16	3.19	.73	.56	2.34	.98	.64
17	2.73	.86	.50	2.41	1.01	.50
18	3.28	.71	.68	3.34	.79	.62
19	3.38	.67	.71	2.85	.99	.78
20	2.24	.92	.17	2.55	1.11	.49

TABLE 11

Means, Standard Deviations and Item-Remainder Correlations for the Six In-Task State Epistemic Curiosity Scales (N=152)

Scale	Item	Mean	SD	Item Remainder
First In-task SECS	1	3.60	.59	.43
	2	3.46	.62	.67
	3	3.32	.70	.57
	4	3.74	.52	.47
	5	3.42	.76	.65
	6	3.49	.66	.66
	7	3.19	.72	.37
Second In-task SECS	1	3.53	.64	.49
	2	3.53	.66	.68
	3	3.43	.66	.75
	4	3.68	.64	.65
	5	3.50	.69	.73
	6	3.17	.79	.58
	7	3.20	.88	.68
Third In-task SECS	1	3.41	.82	.30
	2	3.45	.72	.82
	3	3.33	.79	.78
	4	3.72	.54	.57
	5	3.37	.71	.79
	6	3.44	.76	.79
	7	3.68	.72	.43
Fourth In-task SECS	1	3.18	.87	.61
	2	3.18	.83	.82
	3	3.00	.92	.84
	4	3.39	.80	.69
	5	3.17	.86	.78
	6	3.19	.86	.81
	7	3.19	.96	.61
Fifth In-task SECS	1	3.31	.89	.64
	2	3.16	.95	.82
	3	3.14	.87	.87
	4	3.44	.78	.71
	5	3.14	1.00	.85
	6	3.13	.91	.82
	7	3.31	.88	.46
Sixth In-task SECS	1	2.84	.96	.59
	2	2.89	.97	.85
	3	2.82	.99	.86
	4	3.05	.98	.80
	5	2.89	.99	.86
	6	2.87	.98	.86
	7	2.94	1.06	.65

For the data reported in Table 10, it should be noted that, with the exception of items 12, 13, and 18, the item-remainder correlations increased or remained the same from the pretask to posttask State Epistemic Curiosity Scale measures. Item-remainder correlations on the in-task State Epistemic Curiosity Scale measures ranged from .30 to .87, as can be noted in Table 11. These in-task State Epistemic Curiosity Scale correlations fluctuated depending on the measurement period in which the State Epistemic Curiosity Scales were given.

Concurrent Validity Data

As evidence of the concurrent validity of the two 20-item State Epistemic Curiosity Scales and the six short form State Epistemic Curiosity Scales, these scales were correlated with the Ontario Test of Intrinsic Motivation total scale and Ontario Test of Intrinsic Motivation subscales. Since the Ontario Test of Intrinsic Motivation was considered to be a trait measure of specific curiosity and the State Epistemic Curiosity Scale was assumed to be a state measure of specific epistemic curiosity, moderately high positive correlations between these measures were expected. The correlations between the pretask and posttask 20-item State Epistemic Curiosity Scale, Ontario Test of Intrinsic Motivation total scale, and Ontario Test of Intrinsic Motivation subscales are reported in Table 12. The correlations between the six in-task short form State Epistemic Curiosity Scales, Ontario Test of Intrinsic Motivation total scale, and Ontario Test of Intrinsic Motivation subscales can be found in Table 13.

As Tables 12 and 13 show, significant positive correlations were found between both the 20-item and short form State Epistemic Curiosity Scales and the Ontario Test of Intrinsic Motivation total scales. In addition, the State Epistemic Curiosity Scales were found to correlate significantly with a majority of the Ontario Test of Intrinsic Motivation subscales. It should be noted that these moderately high positive correlations are within the range of the correlations found between trait and state anxiety, as measured by the State-Trait Anxiety Inventory (Spielberger et al., 1970).

Construct Validity Data

Evidence which can be considered to bear on the construct validity of the State Epistemic Curiosity Scale is provided by the correlations of the various State Epistemic Curiosity Scales with the State-Trait Anxiety Inventory A-State and A-Trait scales. Since evidence exists that curiosity and anxiety are inversely related, this inverse relationship should lead to moderately high negative correlations between these state measures. In contrast, since trait anxiety implies relatively stable personality predispositions, relatively low negative correlations between the State-Trait Anxiety Inventory A-Trait scale and the State Epistemic Curiosity Scale measures would be expected.

TABLE 12

Correlations of Pre-Task and Post-Task SECS Scales with OTIM Total Scale and OTIM Subscales (N=152)

Ontario Test of Intrinsic Motivation Scales	Correlations Pre-Task SECS	Post-Task SECS
Total Scale	.52**	.37**
Ambiguity	.03	-.13
Complexity	.02	-.11
Novelty	.20*	.08
Ambiguity-Thinking	.44**	.31**
Ambiguity-Consultation	.34**	.19
Ambiguity-Observation	.47**	.32**
Complexity-Thinking	.43**	.35**
Complexity-Consultation	.45**	.22*
Complexity-Observation	.44**	.32**
Novelty-Thinking	.40**	.36**
Novelty-Consultation	.47**	.32**
Novelty-Observation	.44**	.27**
Diversive Curiosity	.04	.07
Scientific Interest	.52**	.42**
Social Desirability	.25**	.14

* $p < .05$
** $p < .01$

TABLE 13

Correlations of Six In-Task Short Form SECS Scales with OTIM Total Scale and OTIM Subscales (N=152)

Ontario Test of Intrinsic Motivation Scales	Correlations - SECS Scales 1	2	3	4	5	6
Total Scale	.34**	.29**	.16*	.29**	.19*	.25**
Ambiguity	-.04	-.04	-.07	-.06	-.08	-.16
Complexity	.01	-.03	-.09	-.07	-.08	-.14
Novelty	.14	.11	.04	.10	.07	.01
Ambiguity-thinking	.30**	.24**	.14*	.22*	.16*	.18*
Ambiguity-consultation	.21*	.20*	.06	.18*	.05	.08
Ambiguity-observation	.29**	.26**	.11	.33**	.20	.25**
Novelty-thinking	.31**	.21*	.19*	.26**	.22*	.30**
Novelty-consultation	.38**	.30**	.10	.25**	.19*	.21*
Novelty-observation	.25**	.23**	.10	.22*	.12	.16*
Diversive Curiosity	-.09	-.08	.00	.07	-.02	.01
Scientific Interest	.37**	.32**	.23**	.26**	.19*	.27**
Social Desirability	.29**	.24**	.16*	.19*	.23**	.18*

* $p < .05$
** $p < .01$

The correlations between the eight State Epistemic Curiosity Scales and eight State-Trait Anxiety Inventory A-State scales given during the experimental session are reported in Table 14. This table also shows the intercorrelations of the various State Epistemic Curiosity Scales and the A-State scales. Table 15 gives the correlations between the eight State Epistemic Curiosity Scales and the State-Trait Anxiety Inventory A-Trait scale.

For the data reported in Table 14, it can be noted that the majority of the State Epistemic Curiosity and A-State scales were significantly correlated in a negative direction, particularly for the scales given close in time during the experimental session. In addition, both the State Epistemic Curiosity Scales and the State-Trait Anxiety Inventory A-State scales were found to correlate highly among themselves. For the data reported in Table 15, it is instructive to note that, with the exception of the posttask State Epistemic Curiosity Scale measure, all correlations of the State Epistemic Curiosity Scales with trait anxiety were low negative, and only the correlation between A-Trait and the first in-task State Epistemic Curiosity Scale reached significance. Furthermore, the correlational analyses revealed: (a) a correlation of -.01 between the Ontario Test of Intrinsic Motivation and State-Trait Anxiety Inventory A-Trait scale; and (b) correlations ranging from -.10 and -.15 between the Ontario Test of Intrinsic Motivation and State-Trait Anxiety Inventory A-State scales, which were not significant.

Additional evidence bearing on the construct validity of the State Epistemic Curiosity Scale is provided by the correlations between the various State Epistemic Curiosity Scales and portions of the achievement posttest. State epistemic curiosity was assumed to facilitate performance, particularly during the learning task, and thus it would be expected that those students scoring high on the State Epistemic Curiosity Scale measures would make more correct responses on the achievement posttest. That is, moderately high positive correlations would be expected between the State Epistemic Curiosity Scales and posttest sections. These correlations are reported in Table 16.

As Table 16 indicates, a majority of the State Epistemic Curiosity Scales given during the experimental session were found to correlate significantly with the posttest achievement measures. Only the short form State Epistemic Curiosity Scales which were given at the end of the introductory instructions and after the first half of the initial technical learning materials (i.e., short form State Epistemic Curiosity Scales 1 and 2) were not found to correlate significantly with posttest performance. In addition, the highest positive correlations were found between the posttask State Epistemic Curiosity Scale measure, which asked students to reflect on how they felt while learning the materials, and posttest performance. The predictability of the State Epistemic Curiosity Scale as a state measure of curiosity is further supported by the fact that correlations between trait curiosity, as measured by the Ontario Test of Intrinsic Motivation, and the achievement posttest were negligible ($r = .07$ and $r = .05$ for the initial and remaining technical portions of the posttest, respectively).

TABLE 14

Correlation Matrix of the Eight SECS Scales
and Eight STAI A-State Scales (N=152)

SCALES	Pre Task	State Anxiety 1	2	3	4	5	6	Post Test	Pre Task	State Curiosity 1	2	3	4	5	6	Post Task
Pre-Task A-State	.91	.41*	.43*	.48*	.41*	.39*	.34*	.41*	-.30*	-.27*	-.26*	-.07	-.12	-.10	-.11	-.10
A-State 1		.88	.73*	.45*	.39*	.28*	.37*	.45*	-.08	-.22	-.22	-.14	-.09	-.01	-.05	.00
A-State 2			.86	.55*	.52*	.42*	.48*	.50*	-.11	-.32*	-.30*	-.18	-.14	-.10	-.06	-.06
A-State 3				.89	.75*	.65*	.60*	.62*	-.13*	-.24	-.25*	-.30*	-.22	-.22	-.25*	-.14
A-State 4					.91	.75*	.68*	.63*	-.17	-.35*	-.27*	-.21	-.29*	-.30*	.30*	-.20
A-State 5						.86	.70*	.58*	-.19	-.36*	-.30*	-.26*	-.26*	-.31*	.27*	-.21
A-State 6							.86	.73*	-.12	-.27*	-.19	-.21	-.23	-.19	-.31*	-.19
Posttest A-State								.92	-.18	-.28*	-.22	-.20	-.20	-.12	-.21	-.20
Pre-Task State Curiosity									.88	.57*	.55*	.20	.31*	.26*	.26*	.54*
State Curiosity 1										.81	.77*	.41*	.50*	.57*	.45*	.52*
State Curiosity 2											.88	.53*	.58*	.54*	.45*	.57*
State Curiosity 3												.86	.51*	.58*	.48*	.57*
State Curiosity 4													.91	.65*	.64*	.63*
State Curiosity 5														.91	.79*	.68*
State Curiosity 6															.93	.72*
Post-Task State Curiosity																.94

Correlations underlined are significant at the $p < .05$ level; correlations followed by an asterisk are significant at the $p < .01$ level. The alpha reliabilities for the respective scales are given on the diagonals.

TABLE 15

Correlations of Eight State Epistemic Curiosity Measures with Posttest Achievement Measures (N=152)

Measures	Initial Technical	Posttest Sections Remaining Technical	Total Technical
Pre-Task SECS	.20*	.18*	.19*
Short Form SECS 1	.16*	.17*	.18*
Short Form SECS 2	.15	.11	.13
Short Form SECS 3	.06	-.02	-.01
Short Form SECS 4	.28**	.26**	.29**
Short Form SECS 5	.19*	.18*	.20*
Short Form SECS 6	.20*	.30**	.30**
Post-Task SECS	.25**	.31**	.32*

* p < .05
** p < .01

TABLE 16

Correlations of the 20-Item and Short Form State Epistemic Curiosity Scales with the STAI A-Trait Scale (N=152)

State Epistemic Curiosity Scale	Correlations A-Trait Scale
Pre-Task SECS	-.11
Short Form SECS-1	-.19*
Short Form SECS-2	-.14
Short Form SECS-3	-.08
Short Form SECS-4	-.05
Short Form SECS-5	-.14
Short Form SECS-6	-.05
Posttask SECS	.02

* p < .05

Evidence of the construct validity of the State Epistemic Curiosity Scales is also provided by the correlations of these scales with the Diversive Curiosity subscale of the Ontario Test of Intrinsic Motivation. Since the State Epistemic Curiosity Scale is assumed to be a measure of specific epistemic curiosity, low positive correlations between the State Epistemic Curiosity Scale measures and a measured of Diversive Curiosity would be expected. Tables 14 and 15 indicate that the State Epistemic Curiosity Scales did not correlate significantly with the Diversive Curiosity subscale of the Ontario Test of Intrinsic Motivation.

In addition, the results of several of the ANOVAS discussed earlier provide additional validity data. More specifically, as discussed earlier, high trait curious ($\bar{X}$ = 24.00) students had higher levels of state curiosity than low trait ($\bar{X}$ = 21.98) students ($F = 13.50$, $df = 1/144$, $p < .001$) and state curiosity changed across the six in-task measurement periods ($F = 44.48$, $df = 5/720$, $p < .001$). In addition, high in-task state curious ($\bar{X}$ = 7.88) students had lower levels of state anxiety ($F = 5.76$, $df = 2/140$, $p < .01$) than medium ($\bar{X}$ = 8.25) or low ($\bar{X}$ = 9.49) state curious students. High state curious ($\bar{X}$ = 44.68) also performed better on the posttest ($F = 12.65$, $df = 2/140$, $p < .01$) than medium ($\bar{X}$ = 40.33) or low ($\bar{X}$ = 30.43) state curious students. These data, then, serve to demonstrate the fact that the measures obtained using this scale are congruent with theoretical assumptions concerning state epistemic curiosity.

Discussion

The major findings of the present study may be summarized as follows. The hypothesis that high state epistemic curious students would have lower levels of state anxiety during the CAI task than low state epistemic curious students was supported with the present data. This finding is consistent with previous research indicating an inverse relationship between curiosity and anxiety (Day, 1969; Lester, 1968; Penney, 1965). In addition, the prediction of facilitated performance for students scoring high in state epistemic curiosity relative to those scoring low was substantiated, in that high state curious students made more correct responses on the initial technical and remaining technical portions of the posttest than low state curious students.

The hypothesis that students in the Curiosity Stimulating Instruction condition would have higher levels of state epistemic curiosity than students in the No Instruction condition was not supported. In contrast, regardless of whether students were in the Curiosity Stimulating Instruction or No Instruction conditions, levels of state curiosity remained relatively high prior to the CAI task. To interpret this finding, however, several other factors should be taken into consideration. First, the high levels of state epistemic curiosity for all groups prior to the CAI task may be attributable to the novelty of the CAI experience for the majority of students, confounding the effect of experimental manipulation of curiosity through special instructions with that of curiosity aroused by the instructional mode (i.e., a CAI learning experience). Furthermore, the initial manipulation of curiosity prior to adaptation to a novel learning situation may not have been sufficient to maintain levels of curiosity throughout the learning task. Alternately, the possibility that the present Curiosity Stimulating Instruction condition was not efficacious for stimulating or maintaining curiosity must be considered. To more adequately assess the effects of stimulating state epistemic curiosity, therefore, additional research is needed on (a) a variety of Curiosity Stimulating Instruction formats, and (b) more frequent manipulations of curiosity through insertion of curiosity stimulating instructions periodically during the learning task.

Although no main effects of the Curiosity Stimulating Instruction condition were found, there were several interesting interactions with respect to the posttest performance data. For example, both trait curiosity and post-task state curiosity interacted with instruction conditions, indicating a facilitating effect of the Curiosity Stimulating Instruction condition for students with low levels of either trait or state curiosity. In addition, instruction conditions interacted with response modes on the remaining technical portion, which indicated that whereas the Curiosity Stimulating Instruction condition facilitated performance relative to the No Instruction condition for students in the Constructed Response groups, the reverse was found for students in the Reading groups. A further interaction of interest was found between instruction conditions and levels of posttest state anxiety, indicating that the Curiosity Stimulating Instruction condition facilitated the performance of high state anxious students on the initial technical post-test relative to high state anxious students in the No Instruction condition to a level comparable to low state anxious students in the Curiosity Stimulating Instruction and No Instruction conditions.

In general, therefore, whereas there was no main effect of the Curiosity Stimulating Instruction condition on increasing state epistemic curiosity, reducing state anxiety, or improving performance, this condition appears to have the cumulative effect of improving performance for low curious and high anxious students. In addition, the effect of the Curiosity Stimulating Instruction condition was more pronounced for students in the Constructed Response groups, suggesting that the provision of such instructions was particularly helpful when students were required to overtly respond to the learning materials, rather than merely required to read the materials. Assuming that the nature of the task was more complex for students in the Constructed Response groups relative to students in the Reading groups, the Curiosity Stimulating Instruction condition may have served to reduce task complexity by providing students with the scope and direction of the learning task, which resulted in facilitated performance for students in the Constructed Response groups. In contrast, given the less complex nature of the task for students in the Reading groups, the provision of curiosity stimulating instructions which structured the task may have reduced motivational levels and their facilitating effect.

Additional analyses were performed on the present data in order to investigate the hypothesis that the Constructed Response-No Instruction group would have higher levels of state anxiety during the experimental session than the Reading-No Instruction group. These analyses revealed that students in the Constructed Response-No Instruction group had higher levels of state anxiety during the CAI task, but not on the achievement posttest, than the Reading-No Instruction group. Thus, the above hypothesis was partially supported, and these findings partially replicated those found in prior studies with the same learning materials (Leherissey et al., 1971b; Leherissey, O'Neil, Heinrich, & Hansen, 1971). However, it was also found that state anxiety scores during the CAI task for the Reading and Constructed Response groups in the present study (i.e., $\bar{X} = 8.07$ and $\bar{X} = 9.16$, respectively) were lower relative to the state anxiety scores during the technical CAI task for the Reading and Constructed Response groups in prior studies.

EFFECTS OF VARIOUS ANXIETY REDUCTION TECHNIQUES WITHIN

COMPUTER-BASED INTELLIGENCE TESTING

STUDY II

ADDITIONAL RELIABILITY DATA ON A COMPUTER-BASED

SLOSSON INTELLIGENCE TEST

In a previous research report, Hedl (1971) investigated the feasibility and validity of a computer-based administration and scoring program for the Slosson Intelligence Test (SIT) (Slosson, 1963). Using a Latin square design, the computer-administered Slosson Intelligence Test was found to correlate .75 with a Slosson Intelligence Test administered via examiners. Furthermore, equivalent concurrent validity indices between these two scores and the WAIS were found. The computer-administered Slosson Intelligence Test correlated .54 with the WAIS, while the examiner-administered test correlated .52 for this college sample.

Of further interest in the above validation study was the impact of the computer and examiner testing procedures on stress and state anxiety.[1] It was reasoned that evaluative stress to the degree that it depended on a dyadic interpersonal relationship should be mitigated or eliminated in a computer test situation. However, the computer test was found to elicit higher levels of stress and consequently higher state anxiety indices in comparison to either of the two examiner testing situations, irrespective of testing order. Thus, the control and elimination of the human examiner did not necessarily result in a concomitant reduction of stress within the testing situation. In fact, the results of this study suggested the opposite. These affective results were interpreted as being a function of certain procedural considerations involved in the administration of the Slosson Intelligence Test via computer (Hedl, et al., in press).

The purpose of the present study was to replicate the affective relationships within computer testing and to provide additional reliability information on a computer administration of the Slosson Intelligence Test. This dual focus was investigated within the context of a test-retest design.

[1]Portions of these data are reported in Hedl, J. J., Jr., O'Neil, H. F., Jr., & Hansen, D. N. Affective reactions toward computer-based intelligence testing. Journal of Consulting and Clinical Psychology, in press.

Method

Subjects

A total of thirty undergraduate female students participated in the experiment. The sample was drawn from undergraduate psychology and educational psychology courses; volunteers received course credit for their participation. These students were selected on the basis of their medium value trait anxiety scores. Students with either upper or lower quartile scores based on the normative sample of the STAI for college students were assigned to participate in Study III.

Apparatus

The IBM 1500 system, as described in Study I, was also used for this experiment. See Study I for a further description.

Computer-Based Intelligence Test

The Slosson Intelligence Test (Slosson, 1963) was designed to be an individual screening instrument for the assessment of intelligence for both children and adults. The test has adapted many items from the Stanford-Binet, Form L-M (Terman & Merrill, 1960). The items are grouped according to difficulty, mixing content randomly. The major validity study using Stanford-Binet, Form L-M scores and Slosson Intelligence Test (SIT) scores showed concurrent validity coefficients of .90 to .98 with students (N = 701) ranging in age from four years to 18 and above. In another study (Slosson, 1963), a correlation of .92 was found between Binet and SIT scores for students ranging in age from four years to 19 (N = 141).

The computer-based Slosson Intelligence Test (CB-SIT) has been designed to operate with an IBM 1500 Instructional System (IBM, 1967). The test items are presented individually on a cathode ray tube. Students then enter their answers for immediate computer evaluation. If the student is judged as totally correct, or incorrect, the program then continues to the next item. If he is judged as partially correct, the student is instructed to more fully explain his response.

Affective Measures

The State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970) was used to measure both state and trait anxiety. The short form (five-item) state anxiety scale was administered both before and after each of the testing situations. In addition, at the beginning of the first testing situation the students completed the trait anxiety scale of the STAI.

An adjective scale modeled after the Semantic Differential (Osgood, Suci, & Tannenbaum, 1957) was used to measure attitudes toward the concept of "computer testing" and "examiner testing." This scale was given in the future tense form for the pretest of each testing session and the past tense form for each posttest session. Mathis, Smith, and Hansen (1969) have reported test-retest reliabilities ranging from .88 to .96. More importantly, these scale scores have been shown to be related to performance in a computer-assisted instruction mode (Mathis, et al., 1969). Students who committed many errors on the CAI program rated the chapter as more dull, bad, worthless, weak, unpleasant, boring, and depressing than did students who made fewer errors. This measure was selected in order to study the relationship between student test attitudes and the test administration method.

In addition, the hostility scale of the Multiple Affect Adjective Check List (Zuckerman & Lubin, 1965) was administered at the completion of each testing session.

Procedures

For each of the testing situations, the procedures were administered in the following sequence: (1) a pre-task period during which students responded to the short form (5-item) of the state anxiety scale of the State-Trait Anxiety Inventory (STAI) (Spielberger, et al., 1970) and an attitude scale dealing with feelings toward computer testing in general; (2) an evaluative period during which the Slosson Intelligence Test was administered via computer; and (3) a posttest period during which students responded to the short form (5-item) of the STAI state anxiety scale, the attitude scale, and the hostility scale of the Multiple Affect Adjective Check List (Zuckerman & Lubin, 1965). Each of these periods is further described below. The students were tested twice with a one-week interval between test administration sessions. In addition, at the beginning of the first testing session, students completed the trait anxiety scale of the STAI.

1. Pretask Period. The students were first given a detailed description of terminal operations for the administration of the Slosson Intelligence Test on the computer terminal. A brief attitude scale dealing with feelings toward computer testing was then administered to the students via paper and pencil procedures. After "signing on" to the terminal, students completed the five-item state anxiety scale with instructions to indicate how they felt at present.

2. Evaluation Period. The computer-administered Slosson Intelligence Test was then taken by each student. Test items were individually presented on the computer terminals and students responded to each question by typing in their responses. The computer program immediately evaluated the adequacy of answers and when the scoring of particular items was indeterminate, students were asked to amplify their answers (Hedl, 1971).

The computer program began by presenting item 12-0 to all students. Following this, the program proceeded in a linear fashion to present all items until students completed the entire test. As with the previous study (Hedl, 1971), no adaptive test administration procedures were used.

3. Posttask Period. When the test had been completed, the five-item STAI state anxiety scale was re-administered on terminal, with instructions to respond in terms of how they felt during the intelligence test. Following this, the attitude scale and the hostility scale of the MAACL were administered in paper and pencil form. The students were then debriefed and dismissed.

Results

The results section is organized into four major sections. In the first section, the test-retest reliability analysis is presented. In the second section, the STAI state anxiety scores as a function of the experimental procedures, are evaluated. The third section deals with the evaluation of the attitude results. Finally, the hostility data is evaluated as a function of procedures. procedures.

Test-Retest Reliability

To obtain estimates of IQ, the student test item responses obtained from detailed computer performance records were given to a skilled examiner for scoring. This examiner then derived IQ scores from his individual evaluation of each student's performance data. The two sets of IQ scores were then correlated using the Pearson product-moment coefficient technique.

Results of this analysis showed a test-retest correlation of .82, indicating a high positive relationship between the two administrations of the computer test. The mean IQ score obtained from the initial administration of the test was 118.2 (sd = 12.4), whereas the mean IQ score was 125.4 (sd = 9.8) from the second test administration. A t test for correlated sample indicated that there was a significant difference between these two mean IQ scores ($t = 5.6$, $df = 28$, $p < .001$). In other words, there was a significant increase in performance or test-retest effect over the one-week inter-test interval, a finding to be expected considering the test-retest time interval.

Evaluation of Anxiety Levels

To investigate state anxiety levels as a function of testing sessions and pre-, posttest periods, a 2 x 2 repeated measures ANOVA was run on the STAI state anxiety data. The independent variables were week of test administration and pre-, posttest periods. The means and standard deviations of these STAI state anxiety scores by test administration and measurement periods are presented in Table 17.

The results of the 2 x 2 ANOVA with repeated measures are reported in Table 18. As the reader may note, STAI state anxiety scores were higher during the first administration of the test ($\bar{X} = 12.6$) in comparison to the second test administration ($\bar{X} = 9.6$) as indicated by the significant main effect for test administration. Furthermore, the state anxiety levels appeared to remain relatively constant within each of the testing sessions. This finding was substantiated

Table 17
Means and Standard Deviations of STAI State Anxiety Scores by Experimental Conditions (N = 30)

Test Administration	Pretest	Posttest
First Test Administration		
Mean	12.83	12.27
SD	3.70	3.60
Second Test Administration		
Mean	9.47	9.63
SD	3.14	3.65

Table 18

Analysis of Variances of STAI State Anxiety Scores by Order of Test Administration and Measurement Periods

Source	df	MS	F
Between Subjects	29	29.94	
Within Subjects			
Test Administration (A)	1	270.0	20.74**
A x Subjects Within Groups	29	18.0	
Pretest-Posttest Periods (P)	1	1.20	< 1
P x Subjects Within Groups	29	3.87	
AP	1	4.03	
AP x Subjects Within Groups	29	2.98	1.35

**p < .01

via the non-significant F value for the comparison of pretest and posttest anxiety scores. Differential effects for the two testing conditions were also non-significant.

A further analysis was conducted to determine the reliability characteristics of the STAI scale for the present sample. In this regard, Kuder-Richardson 20 reliability estimates were derived via analysis of variance techniques. Table 19 presents the KR-20 estimates for the four administrations of the five-item STAI state anxiety scale.

As the reader may note, the range of the estimates was from .87 to .92 for this sample of college students. Thus, the five-item scale is impressive with regard to its internal consistency of anxiety measurement. These results are consistent with the normative data reported on the full 20-item scale (Spielberger, 1972).

Evaluation of Attitude

The impact of the testing procedures was further evaluated along a student attitudinal dimension. A 2 x 2 ANOVA with repeated measures was used to examine the attitude scale scores with respect to the two test administrations and before-after measurement periods. This statistical analysis was based on a total summated score for the 15-item Likert rating scale.

The means and standard deviations of the attitude scale scores are reported in Table 20. The results of the ANOVA are presented in Table 21 in which the most salient finding was the significant interaction between the different test administrations and the before-after dimension. The main effects for test administration and measurement periods were found to be non-significant.

This interaction is graphically presented in Figure 9. As the reader may note, the students had a moderately high attitude toward the computer testing situation prior to the first test administration. However, this attitude was lowered following their initial experience with the terminal testing. Student attitudes appeared to have remained at this lower level during their second encounter with the computer test.

It should be noted that although attitude levels decreased as a function of computer testing exposure, the resultant mean attitude scores were still above the theoretically neutral score of 60 on the scale. Thus, student attitudes appeared to remain on the positive domain of the attitude rating scale. This general trend of a decrease in attitude scores was also found in the previous study using the same computer testing situation (Hedl, 1971).

As with the previously cited data on the STAI scale, Kuder-Richardson 20 estimates were derived for the four administrations of the attitude scale and are presented in Table 22.

With the exception of the pretest measure in the first test administration, the resulting KR-20 coefficients compare favorably to previously reported data on the scale (Hedl, 1971; Mathis, Smith, & Hansen, 1969).

Table 19

Kuder-Richardson 20 Reliability Estimates for the

Four Administrations of the STAI State Anxiety Scale ($\underline{N}$ = 30)

Test Administration	Pretest	Posttest
First Test Administration	.87	.89
Second Test Administration	.90	.92

Table 20

Means and Standard Deviations of Total Attitude Scale Scores by Experimental Conditions

Test Administration	Pretest	Posttest
First Test Administration		
Mean	72.20	68.90
SD	6.73	9.36
Second Test Administration		
Mean	68.87	69.57
SD	9.86	10.66

Table 21

Analysis of Variance of Attitude Scale Scores by Experimental Conditions

Source	df	MS	F
Between Subjects	29	250.96	
Within Subjects			
Test Administration (A)	1	53.33	1.35
A x Subjects Within Groups	29	39.45	
Pretest-Posttest Periods (P)	1	50.70	2.06
P x Subjects Within Groups	29	24.65	
AP	1	120.00	4.20*
AP x Subjects Within Groups	29	28.60	

*$p < .05$

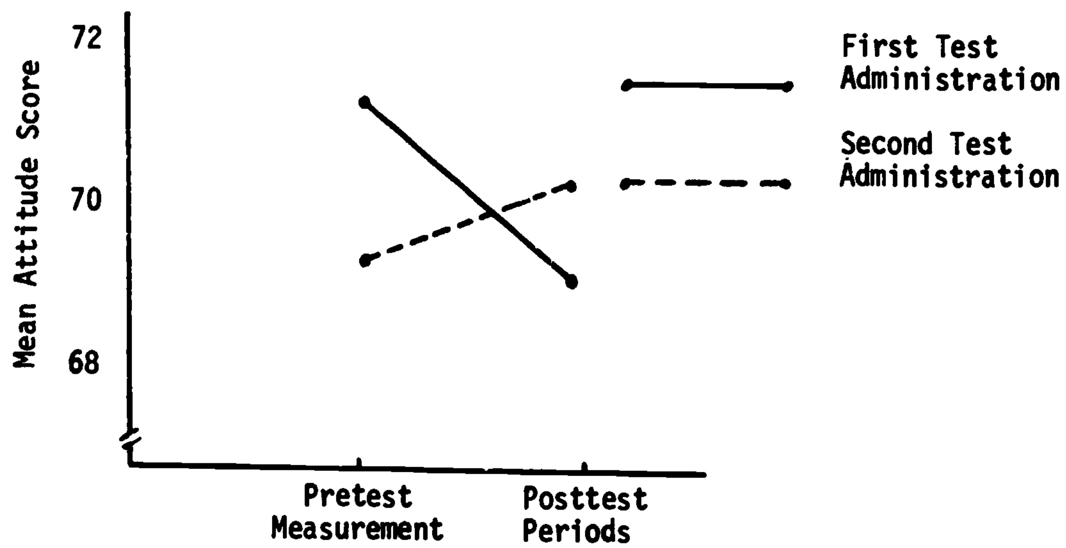

Figure 9. Interaction between test administrations and measurement periods for the attitude scale scores

Table 22

Kuder-Richardson 20 Reliability Estimates for the Four Administrations of the Attitude Scale Scores (N = 30)

Test Administration	Pretest	Posttest
First Test Administration	.52	.75
Second Test Administration	.81	.85

Evaluation of MAACL Hostility Scores

An additional measure in this study was the inclusion of the Hostility scale of the MAACL consisting of 28 hostility items and 12 buffer items. The statistical analysis of the hostility scale was based on the total hostility score for the scale and consisted of a *t* test for correlated samples.

The means and standard deviations of the Hostility scale scores are presented in Table 23 along with Kuder-Richardson 20 estimates. The observed minor decrease in Hostility scores was not significant (t = .03, df = 28), indicating relatively consistent hostility reactions across the experimental situations. Thus, whereas both state anxiety levels and attitude scores decreased over the experimental situation, hostility scores did not show the same relationship over the course of the experiment.

Discussion

The primary purposes of this study were to provide test-retest reliability for the computer-based Slosson Intelligence Test and to replicate the affective data obtained on the program. With regard to the first objective, the test-retest correlation of .82 indicated the temporal stability of the IQ scores considering the homogeneous nature of the college students' abilities, thus establishing the reliability of the test scores.

Additional concurrent validity data on the computer-based Slosson Intelligence Test was obtained by relating the IQ scores to college grade point averages (GPA). The correlations were .37 and .52 for the first and second administrations with GPA, respectively. Both of these correlations were significant at the .01 level. Thus, it appears that the computer test produces reasonably stable and valid scores in reference to GPA averages.

Of perhaps greater interest is the affective data. It was expected that the first administration of the computer test would yield affective indices equivalent to the previous research. In addition, it was expected that the affective data from the second administration would yield somewhat reduced anxiety levels and increased attitudinal reactions. The data from the present study generally conformed to these expectations. The mean state anxiety scores obtained from the first computer test closely replicated the previous data, again indicating elevated state anxiety levels in comparison to traditional testing conditions. The attitude data also replicated the previously reported findings in which the attitude scores were found to be lower after the students had completed the testing procedures.

The data from the second testing session also conformed to expectations. Performance was found to increase presumably due to practice effects. In addition, anxiety levels were found to decrease and compare favorably with the human examiners' data, although one cannot determine the anxiety-performance causative

Table 23

Means, Standard Deviations, and Kuder-Richardson 20 Estimates for the Hostility Scale of the MAACL (N = 30)

Test Administration	Mean	SD	KR-20
First Test Administration	9.33	3.51	.78
Second Test Administration	9.13	3.70	.78

factors in this study. It is conceivable that their familiarity with the procedures in the second testing situation caused students to feel more at ease with the computer test and evaluation procedures, and therefore, to be less anxious.

The test administration procedures used in the present study were not adaptive, i.e., not based on student performance. A straightforward presentation strategy was employed. In this respect, the computer program lacked the provision to cease testing when the student had reached ceiling--defined as 10 consecutive incorrect responses. Therefore, students received the entire array of test items regardless of performance. If a student's performance indicated that he reached ceiling, he still would have received the remaining test items. Thus, a student could conceivably terminate the computer test with a massive failure experience, and this may well have contributed to the observed high levels of state anxiety.

Previously, it was found that this computer presentation procedure appeared to have adverse anxiety effects within the testing situation (Hedl, 1971). Comparison of state anxiety scores for those students who did not reach ceiling on the computer test with those students who received the remaining test items after their performance indicated ceiling revealed higher state anxiety scores for the latter group. Thus, failure to terminate the test according to performance led to an increase in the level of stress in the situation as measured by the state anxiety scale of the STAI. These findings would appear to indicate that a revision of the test administration procedures, allowing the student to terminate once his performance reached ceiling, would act to decrease stress in the testing procedure. Study III was concerned with an evaluation of this possibility.

STUDY III

EFFECTS OF PROGRAM ADMINISTRATIVE REVISIONS ON STATE ANXIETY WITHIN COMPUTER-BASED INTELLIGENCE TESTING

The findings of the previous study replicated the affective results of the validation study (Hedl, 1971) and also provided promising test-retest reliability for the computer administration of the Slosson Intelligence Test. State anxiety indices were found to be higher than the previous examiner data and to be most similar to the prior computer administration findings. In addition, suggestive evidence was provided to support the interpretation that both specific test administration procedures in terms of ceiling considerations and familiarity with the terminal testing procedures were important causative stress factors within the computer test, although the relative importance of terminal familiarity on state anxiety reactions may have been confounded with performance increments. These data also suggest the hypothesis that computers are not anxiety producing in and of themselves. The more causal factors would seem to arise from the nature of the computer task (Hansen & O'Neil, 1970).

This study sought to evaluate a number of procedural revisions in the computer administration of the Slosson Intelligence Test and their effect on subsequent anxiety levels within the computer test. In this respect, a further clarification of the relative importance of within-test presentation methods as well as familiarity with the terminal operations could be determined. The revision of the test administration procedures was based on both empirical and heuristic considerations in terms of the sequence of item presentation, the increase in permissible computer time for responding, and the provision for a more performance-sensitive item presentation in terms of the inclusion of basal age and ceiling age branching options. These test revisions were then evaluated within a test-retest design on test-retest reliability, anxiety reduction effectiveness, and reduction of testing time. An additional dimension within the present study was the inclusion of students differing in A-Trait levels to investigate the possible interactive effects of computer presentation strategies and anxiety-proneness of the students.

Method

Subjects

A total of 30 female students participated in this experiment. The sample was drawn from undergraduate psychology and educational psychology courses with volunteers receiving course credit for their participation. These students were selected on the basis of their upper or lower quartile scores based on the normative STAI college sample. Eighteen low trait-anxious and 12 high trait-anxious were included in the sample.

Apparatus

The IBM 1500 system as used in Studies I and II was also used in this experiment. See Study I for a further description.

Computer-Based Intelligence Test

As in Study II, the Slosson Intelligence Test was used as the intelligence measure; however, the test was revised in an effort to reduce anxiety and testing time while maintaining reliability and validity. The initial version of the computer-based Slosson Intelligence Test used a straightforward item presentation strategy. Item 12-0 was presented first to all students followed by a linear progression of test items until the student completed the entire range of items. Thus, the program was insensitive to a number of presentation strategies used by the typical human examiner.

The first major revision focused on this problem. It was decided to begin the test at item 21-3 and present items in a sequentially reverse fashion until basal age defined as 10 consecutive item successes were reached. When basal age had been determined, the computer program branched to item 21-6 and continued to present the more difficult items until either the student reached ceiling defined as 10 consecutive item failures or completed the entire array of test items.

The revisions of this nature were based on prior data indicating the highest basal age observed for any student and similar considerations concerning students' reaching ceiling. In other words, no prior student had received a basal age equal to or higher than 21-3 or shown a ceiling at this item. In addition, this presentation strategy more optimally simulates the nature of a human presentation of the same test in which the examiner attempts to determine basal age initially and then proceeds through the more difficult items in a sequential fashion to determine the ceiling or possible ceiling of the student.

The second major program revision was concerned with the time allowed for item responses. Previous student comments substantiated by analysis of "time outs" indicated that certain items, especially the mathematical questions had inappropriate time limits. When a time out was reached, students were asked by the computer program to "Please answer the question or type pass." Apparently this program reminder had adverse affective reactions on the students, and possibly performance effects. In order to reduce the effects of this variable, the time limits for the mathematical and verbal items in question were increased twofold.

The research design consisted of testing high and low trait-anxious students with the revised computer-based Slosson Intelligence Test twice with a one-week interval between the test administrations. In this respect, additional reliability information could be gained on the computer test in addition to the effects of program revisions on state anxiety levels, attitude reactions, time to complete the test, and hostility scores.

Affective Measures

As in the previous study, the State-Trait Anxiety Inventory, the Hostility scale of the Multiple Affect Adjective Check List, and an attitude scale dealing with feelings toward computer testing were used as affective measures. (See Study II for further description.)

Procedure

The experimental procedures in Study III were exactly the same as those used in Study II. During the pretask period, students responded to the state anxiety scale of the STAI and an attitude scale toward computer testing. The evaluation period consisted of the revised computer administration of the Slosson Intelligence Test. As with Study II, the five-item STAI state anxiety scale, attitude scale, and Multiple Affect Adjective Check List Hostility scale were completed during the posttest period. These procedures were repeated for both testing sessions. For a more detailed description of these procedures see Study II.

Results

The results section is organized into five major sections. In the first section, results of the test-retest reliability analysis are presented. In the second section, the STAI state anxiety scores are evaluated as a function of trait anxiety and experimental conditions. The third and fourth sections deal with analyses of the attitude and hostility scores, respectively. The final section deals with an evaluation of total testing time.

Test-Retest Reliability

The IQ scores from the computer-based Slosson Intelligence Test were obtained in a similar manner as in Study II. To obtain a test-retest reliability estimate, the scores from the two administrations of the computer test were correlated via the Pearson product-moment method.

The results of this analysis showed a test-retest correlation of .72 indicating a moderately high positive relationship between the two sets of IQ scores. The mean IQ score from the first computer test administration was 120.5 ($SD = 7.83$), with a mean of 123.7 ($SD = 9.52$) from the second test administration. A t test for correlated samples indicated that this difference of 3.2 points was statistically significant ($t = 2.70$, $p < .02$). Thus, there was a significant increase in test scores over the one-week interval.

Evaluation of Anxiety Levels

To study state anxiety levels as a function of trait anxiety, testing session, and pre-post measurement periods, a 2 x 2 x 2 analysis of variance with repeated measures on the last two factors was run. The means and standard deviations of these STAI state anxiety scores are presented in Table 24. Six low anxious students were deleted on a random basis to achieve sample size comparability in the data analysis.

The results of this 2 x 2 x 2 ANOVA are presented in Table 25. As the reader may note, high anxious students were shown to exhibit higher levels of state anxiety throughout the two testing sessions. This finding is indicated by the significant main effect for trait anxiety. State anxiety levels were also shown to decrease across the testing conditions as shown by the main effect for testing sessions. The overall state anxiety mean for the first test was 11.0, whereas a mean of 8.1 was shown for the second test.

State anxiety levels appeared to remain relatively constant within the testing sessions as substantiated by the nonsignificant F value for the pretest-posttest state anxiety overall comparison. In addition, none of the second or third order interactions were found to be significant.

A further analysis was run to determine the internal consistency reliability of the STAI state anxiety scale for the present sample. Table 26 presents the KR-20 estimates for the four administrations of the five-item scale. These results are consistent with the estimates obtained in Study II and with the normative data on the scale.

Evaluation of Student Attitudes

Parallel data analyses were conducted on the attitude scale scores as a function of trait anxiety, testing session, and pre-post measurement periods. The means and standard deviations of these scores are presented in Table 27. As with the state anxiety analyses, the same six students were deleted from the original sample of 30 to achieve equal sample sizes for the statistical analysis.

The results of the 2 x 2 x 2 ANOVA with repeated measures on the last two factors are presented in Table 28. The most salient finding of this analysis was that low anxious students were found to show higher overall attitude scores in comparison to the high anxious students. This result can be seen from the significant main effect for trait anxiety level. The overall mean attitude score of low anxious students was 74.52, whereas high anxious students showed a mean score of 64.1. The two additional main effects for testing session and pre-post measurement periods were not found to be statistically significant. Furthermore, none of the second or third order interactions reached significance.

Table 24

Means and Standard Deviations of STAI State Anxiety Scores by Trait Anxiety, Testing Sessions, and Pre-Post Measurement Periods ($N = 24$)

Trait Anxiety	First Test Administration		Second Test Administration	
	Pretest	Posttest	Pretest	Posttest
High Anxious				
Mean	12.67	12.42	11.17	8.67
SD	4.10	4.48	3.97	3.73
Low Anxious				
Mean	9.33	9.75	6.58	6.08
SD	3.63	3.17	1.83	2.02

Table 25

Analysis of Variance of STAI State Anxiety Scores by Trait Anxiety, Testing Sessions, and Pre-Post Measurement Periods

Source	*df*	*MS*	*F*
Between Subjects			
Trait Anxiety (A)	1	260.00	9.79*
Subjects within groups	22	26.56	
Within Subjects			
Testing Sessions (T)	1	204.17	23.79*
AT	1	2.04	< 1
T X subjects within groups	22	8.58	
Pre-Post Periods (P)	1	12.04	1.94
AP	1	10.67	1.72
P X subjects within groups	22	6.20	
TP	1	15.05	2.08
ATP	1	2.67	< 1
TP X subjects within groups	22	7.24	

* $p < .01$.

Table 26

Kuder-Richardson 20 Reliability Estimates for the Four Administrations of the STAI State Anxiety Scale (N = 24)

Test Administration	Pretest	Posttest
First Test Administration	.83	.93
Second Test Administration	.85	.93

Table 27

Means and Standard Deviations of Attitude Scores by Trait Anxiety, Testing Session, and Pre-Post Measurement Periods (N = 24)

Trait Anxiety	First Test Administration		Second Test Administration	
	Pretest	Posttest	Pretest	Posttest
High Anxious				
Mean	68.92	62.67	62.50	62.17
SD	11.02	13.07	13.71	12.90
Low Anxious				
Mean	74.75	73.58	75.33	74.42
SD	12.56	7.24	8.47	11.34

Table 28

Analysis of Variance of Attitude Scores by Trait Anxiety, Testing Session, and Pre-Post Measurement Periods

Source	df	MS	F
Between Subjects			
Trait Anxiety (A)	1	2625.00	6.36*
Subjects within groups	22	417.65	
Within Subjects			
Testing Sessions (T)	1	112.67	2.90
AT	1	30.38	< 1
T X subjects within groups	22	38.80	
Pre-Post Periods (P)		45.38	1.25
AP		104.17	2.86
P X subjects within groups		36.41	
TPP		57.04	1.40
ATP		48.17	1.18
TP X subjects within groups		40.65	

*$p < .05$.

Subsequent analyses were conducted to determine internal consistency reliability estimates for the 15 item attitude scale. Table 29 presents the results of these analyses. These estimates compare favorably to those found in Study II and to previously reported data on the scale.

Evaluation of Multiple Affect Adjective Check List Hostility Scores

The hostility scores obtained from the Multiple Affect Adjective Check List (28 hostility items and 12 buffer items) were subjected to similar ANOVA analysis procedures. The means and standard deviations of the Multiple Affect Adjective Check List scores are presented in Table 30. These data were evaluated by a 2 x 2 ANOVA with repeated measures on the final factor.

Results of this analysis (Table 31) revealed that the high anxious students showed higher levels of hostility than did their low anxious counterparts. This finding is substantiated via the significant main effect for trait anxiety. The overall mean for the high anxious students was 11.13, while the low anxious students had a mean hostility score of 7.29.

The hostility scores were also shown to be relatively consistent across administration within trait anxiety groups via the nonsignificant main effect comparing the two test administrations. In addition, the trait anxiety by test administration interaction was also nonsignificant.

Internal consistency reliability estimates (KR-20's) were found to be .86 and .88 for the two administrations of the Multiple Affect Adjective Check List hostility scale. These results compare favorably to the estimates found in Study II.

Evaluation of Testing Time

A major concern in the revision of the computer test was the reduction of testing time via providing for a more response-sensitive administration procedure. In this regard, provision was made to provide for basal and ceiling age cutoff strategies in the test administration. This procedure did serve to reduce the number of test items taken by an individual. The first computer test presented 70 items regardless of student performance. This figure was reduced to a mean number of 60.0 items in the first administration of the revised computer program. The second administration of the revised program yielded a mean of 55.23 presented items. However, this reduction in the number of items did not lead to a concomitant decrease in testing time.

Table 32 presents the mean "time on system" for the two administrations of the computer test in Study II and for Study III. "Time on system" was defined as the time elapsed from when the student "signed on" to the computer terminal to

Table 29

Kuder-Richardson 20 Reliability Estimates

for the Four Administrations of the

Attitude Measure (N = 24)

Test Administration	Pretest	Posttest
First Test Administration	.82	.87
Second Test Administration	.86	.89

Table 30

Means and Standard Deviations of the MAACL Hostility Scores by Trait Anxiety and Test Administration

Trait Anxiety	First Administration	Second Administration
<u>High Anxious</u>		
Mean	12.08	10.17
<u>SD</u>	3.75	5.11
<u>Low Anxious</u>		
Mean	6.75	7.83
<u>SD</u>	3.62	4.00

Table 31

Analysis of Variance of MAACL Hostility Scores by Trait Anxiety and Test Administrations

Source	*df*	*MS*	*F*
Between Subjects			
Trait Anxiety (A)	1	176.33	8.14*
Subjects within groups	22	21.66	
Within Subjects			
Test Administrations (T)	1	2.08	< 1
AT	1	27.00	2.08
T X subjects within groups	22	12.99	

*p < .01.

Table 32

Mean Testing Time for the Initial

and Revised Program*

Computer Test	1st Administration	2nd Administration
Initial Version (Study II)	51.0	40.5
Revised Version (Study III)	54.2	40.8

*In minutes based on 30 students in each study.

when he had completed the test and was automatically "signed off" by the computer program. Within Study III, the amount of testing time was compared for both high anxious and low anxious students and found to be the same. Thus, these data were pooled together for the present analyses.

As the reader may note, the total time on system was fairly equivalent in both studies across the two testing conditions. Thus, the revised presentation strategy did not lead to a decrease in testing time.

A further analysis was conducted on the latencies for 14 of the math items with the increased allowable time limits. Comparisons were made between those items for both Study II and Study III (Table 33). Latency was defined as the time elapsed from the presentation on the cathode ray tube of the test item to the time when the student entered his response into the computer program.

The reader may note that students took more time answering questions on the revised computer program. This was true for each of the 14 items in question. Thus, there appeared to be a strong tendency for students to use the additional time if available on the items. Even though the program revisions succeeded in reducing the total amount of items presented, the possible reduction of testing time due to this decrease was eliminated by the students spending additional time on the items with the increased allowable time limits.

Discussion

The primary purposes of the present study were to evaluate a number of program administration revisions on stress and state anxiety within a computer testing situation. Of primary concern was the effect of these test administrative procedures on the subsequent state anxiety levels observed within the computer test. An indirect method to assess these effects would be to compare these anxiety scores with the previous data on the first program version. Figure 10 presents a graphical comparison of the pretest and posttest anxiety scores from Study II in which the initial version of the computer test was employed and the present study in which the revised program was used. In addition, anxiety data from a previous study using human test examiners are presented (Hedl, 1971).

Inspection of Figure 10 indicates that the revised program strategies were effective in anxiety reduction during the computer test. The pretest and posttest means from the initial computer program administration were 12.83 and 12.27, respectively, based on 30 female students. The same means were 11.0 and 11.1 for the revised program based on 24 female students. Thus, these data would suggest the effectiveness of the administration revisions for anxiety reduction purposes.

However, the revised program still produced higher anxiety levels in comparison to previous data obtained from an examiner-administration of the Slosson Intelligence Test. These latter means were based on a sample of 16 college students

Table 33

Mean Response Latencies in Seconds

for the 14 Math Items in

Study II and Study III

Item	Study II (N = 30)	Study III*
14-0	15.81	23.98 (N = 24)
14-6	51.18	56.53 (N = 24)
14-10	32.39	36.13 (N = 26)
15-6	22.73	24.98 (N = 28)
16-0	34.64	37.76 (N = 30)
16-9	36.54	46.24 (N = 30)
19-0	33.06	52.18 (N = 30)
19-9	48.39	77.09 (N = 30)
20-3	46.71	71.44 (N = 30)
20-9	53.39	67.39 (N = 30)
21-9	44.08	54.91 (N = 28)
22-3	43.74	62.49 (N = 28)
22-9	42.79	62.88 (N = 28)
23-3	49.12	66.53 (N = 27)
24-0	45.81	63.41 (N = 26)
24-3	41.67	69.44 (N = 26)

*N for item is dependent on the number of students to which the item was presented.

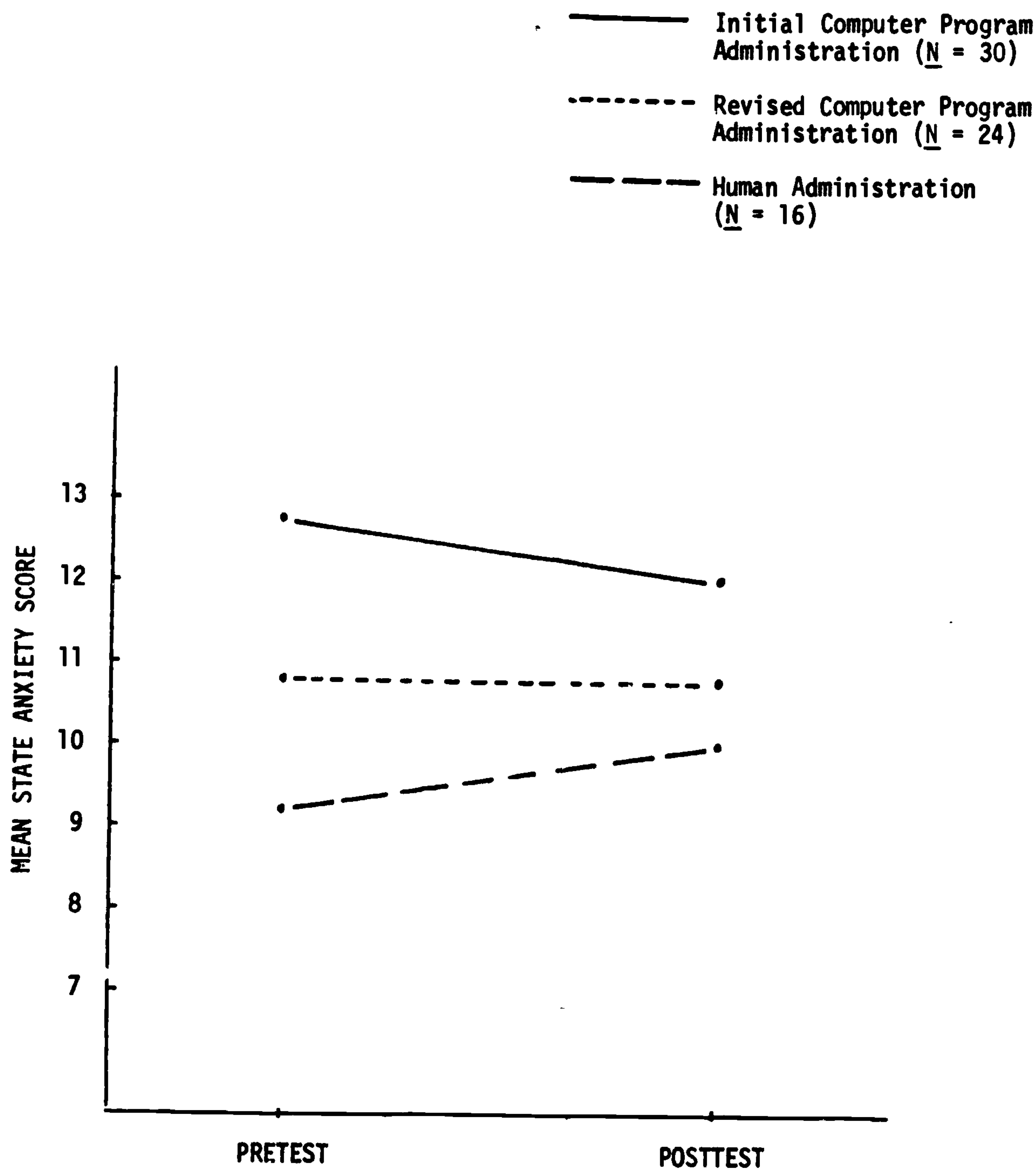

Figure 10.--Comparison of anxiety score as a function of different administrative procedures.

(8 males, 8 females) obtained during the validation of the programmed approach (Hedl, 1971). The pretest and posttest means were 9.44 and 10.00, respectively. Thus, although the revisions of the program were effective in comparison to initial version of the program, the revised computer test compared less favorably to the anxiety levels obtained from an examiner-administration of the same intelligence test. Given the possible differences in the experimental samples, these data should be considered as being suggestive.

Another evaluation index for the effect of the revised administrations on the reduction of negative affective variables would be the attitude scores. The primary consideration in this respect would be the observed attitude scale levels from the different computer test administration procedures. Figure 11 presents a graphical description of the pretest and posttest attitude scores from Study II and Study III. Furthermore, attitude data obtained from examiner-administered tests are also included.

Inspection of the attitude scores in Figure 11 would seem to indicate that the program administrative revisions were not effective in altering the students' reactions toward computer testing in general in comparison to the initial computer administration procedures. The means for the attitude scale from the initial program were 72.2 and 68.9 for the pretest and posttest measurements. In comparison, the same means for the revised program were 71.8 and 68.9 indicating only minor differences in attitude scores as a function of test administration procedures.

These results also indicate that both computer tests lead to lower attitude scores in relation to an examiner-administration of the same test. The latter pretest and posttest means were 73.1 and 70.5. Thus, while there do not appear to be differences among these groups in terms of their pretest attitude scores, there appears to be a difference in the posttest attitude scores which asked the students specifically about their reactions to the testing procedures and conditions. In this respect, although attitude scores decreased in all cases, the students in the examiner test administration group tended to have more favorable opinions about the testing experience.

Of additional interest in the present study was the effect of program revisions on time to complete the test. It was reasoned that a program that was more sensitive to responses in terms of administration would reduce the total number of items presented and thus lead to a reduction in total testing time. The benefits to be derived from this procedure were hoped to be related to anxiety reduction within the task. The results, as previously indicated, indicated that the total number of items was reduced in the present study; however, there was no concomitant reduction in total testing time. This effect was apparently mitigated in that students also spent more time on certain of the items which had increased time limits for responding. In this sense, students tended to use the allowable time limits per question.

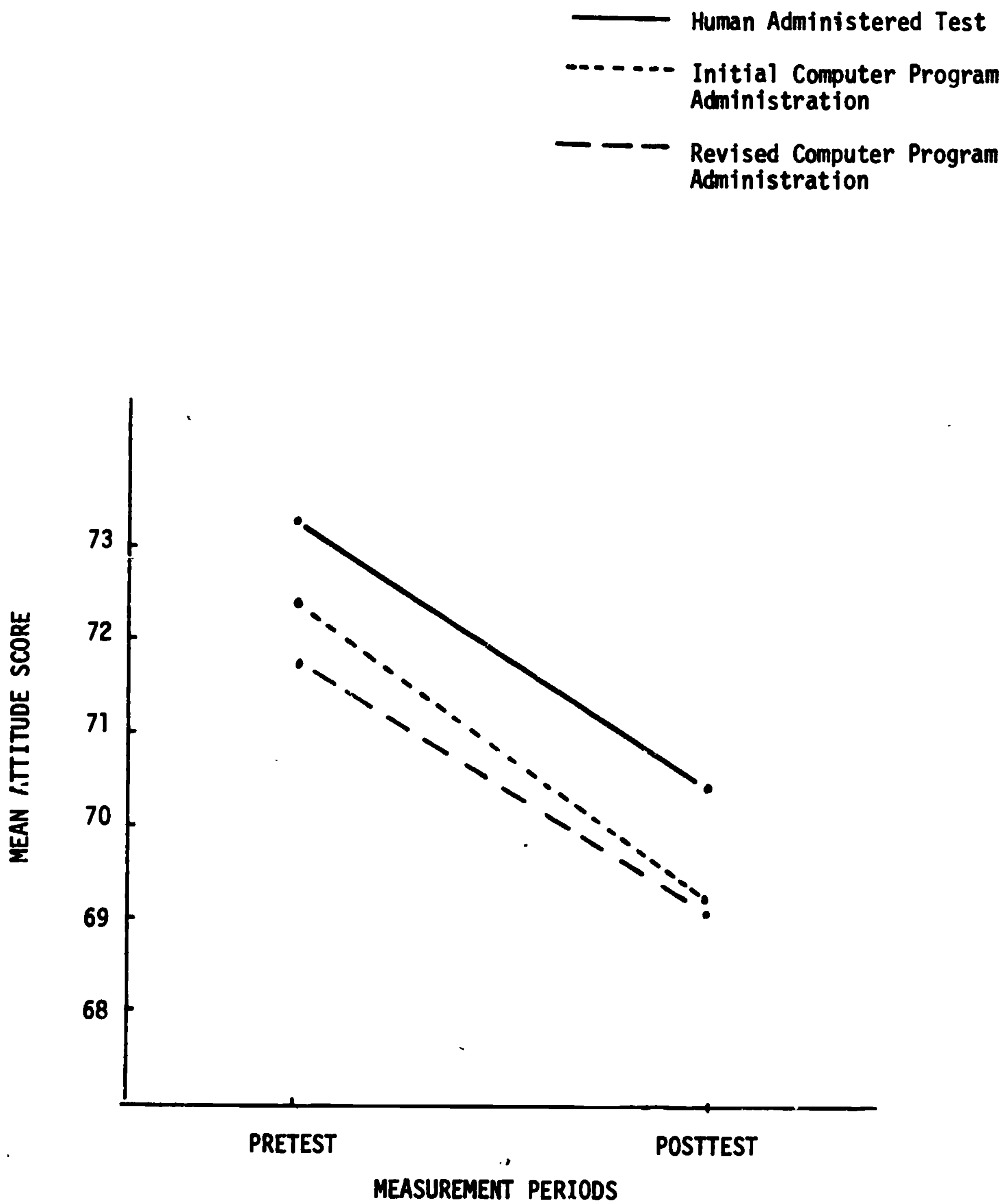

Figure 11.--Comparison of attitude scores as a function of different intelligence test administration procedures.

Examination of the correlations between total testing time and state anxiety scores from the present study indicated that this relationship was not as important as first hypothesized. Correlations of the post state anxiety scores which asked the students to respond in terms of their anxiety during the test and total time testing were nonsignificant for both administrations of the computer test. Thus, it would appear that the amount of time spent in the testing situation is not a contributing factor in the elicitation of state anxiety. It would seem that variables such as lack of familiarity with the novel terminal procedures and the nature of the evaluative interaction with the computer test would be more important determiners of anxiety than merely length of task period. Thus, the duration of the task does not appear nearly as important as the actual events during the task period.

STUDY IV

EFFECTS OF PRIOR TERMINAL EXPERIENCE ON AFFECTIVE INDICES IN COMPUTER-BASED INTELLIGENCE TESTING

The results of Studies II and III indicated that the initial administration of computer tests led to affective reactions which were similar to the first version of the computer-based intelligence test. The reliability and concurrent validity indices showed promising psychometric support for the utility of computer intelligence testing; however, the affective indices indicated that additional factors were operative in the computer testing situation in comparison to the human-administered tests.

The affective results from the second administration in both Studies II and III compared favorably with data from human-administrations of the Slosson Intelligence Test. However, it was impossible to determine whether these affective results were a function of the observed increase in performance, therefore making the situation less threatening in nature, or the familiarity with the unique testing procedures resulting in reduced anxiety levels and therefore higher performance, or practice effects related to the inter-test interval of one week. It is likely that the observed results from the second administration of the computer test were a function of both test practice effects and familiarity with testing procedures.

In order to discriminate between these factors, the fourth study was concerned with the specific effects of prior terminal experience on subsequent affective indices within the computer testing situation. In this regard, a more precise delimitation of the anxiety-causative effects could be determined. If familiarity with terminal procedures is the more important variable within a presumed stressful testing situation, the observed affective indices in the testing situation should be differentially related to prior exposure and practice on the computer terminals. Study IV was designed to explore this hypothesis in more detail.

The primary purpose of this study was to investigate the effects of prior terminal experience in the form of a computer game on subsequent anxiety indices within an evaluative computer testing situation. In this respect, the efficacy of computer gaming as an anxiety reduction technique could be explored.

Method

Subjects

A total of 60 students participated in this experiment (25 males, 35 females). The sample was drawn from an undergraduate educational psychology course. Course credit was given by their instructors for the volunteers' participation in the experiment.

Computer-Based Intelligence Test and Computer Game

As in the two prior studies the Slosson Intelligence Test, with the revisions made in Study III, was used as the intelligence measure in this experiment. For a further description of this test, see Study II.

In addition, a computer game was used to provide the students with prior terminal experience in an effort to reduce anxiety on the intelligence test. The game consisted of a simulated horse race in which artificial odds on six horses were given and an imaginary budget of $10,000 allotted for "betting." Students were then asked to indicate on which horses they wanted to bet; the amount of the bet; and whether they wished to bet, win, place, or show. Students could bet on as many of the horses as they desired. After the bets had been placed, the actual race was viewed on the cathode ray terminal.

Six horses were included in each race, represented on the screen by elongated "m's." During the race, the computer program simulated the race by allowing the various "horses" to flash across the screen at a predetermined, randomly-selected pace. Races were unique to each individual student. In other words, if there were five students playing the game, there would be five different races occurring at any point in time. Winning or losing at the game was randomly determined.

At the conclusion of any one race, students were informed on the screen of the present status of their imaginary bank account based upon an update using the previous race results. The upcoming race odds were then presented to the student for betting selection, followed by the simulated running of the race. This procedure was identical for each race during the 20-25 minute game period.

Affective Measures

As in Studies II and III, both the state anxiety and trait anxiety scales of the State-Trait Anxiety Inventory and the attitude scale dealing with feelings toward computer testing were used as affective measures in this study. In addition, the Test Anxiety Scale (Sarason, 1958) was administered as a pretask affective measure. (See Study II for a further description of these measures.)

Procedure

The procedures were administered in the following sequence for all students: (1) a pretask period, during which students responded to the Test Anxiety Scale (Sarason, 1958) and the trait anxiety scale of the State-Trait Anxiety Inventory (STAI) (Spielberger, Gorsuch, & Lushene, 1970); (2) a non-evaluative game period during which students played a game on the cathode ray terminal and then completed the five-item state anxiety scale; (3) an evaluation period during which the students first responded to the five-item STAI state anxiety scale and the attitude scale concerning computer testing, then were administered the Slosson Intelligence Test; and (4) a posttask period during which students responded to the five-item STAI state anxiety scale. Each of these periods is further described below.

1. Pretask Period. During this period, the students first completed a paper and pencil form of the Test Anxiety Scale and the trait anxiety scale of the STAI. The students then received written instructions informing them how to operate the computer terminal and then responded to a paper and pencil administration of the brief five-item state anxiety scale, with instructions to respond in terms of "How do you feel right now." Practice in the operation of the terminal keyboard was presented and students "signed on" to the computer game.

2. Nonevaluative Game Period. During the game, students worked individually at computer terminals. The game period was defined by a time interval of 20-25 minutes and not by a total number of races completed. Following completion of the game period, the five-item state anxiety scale was re-administered with instructions to indicate "How did you feel during the game you just played?"

3. Evaluative Period. The experimental procedures used for the computer administration of the Slosson Intelligence Test were identical to those used in the previous study. Prior to the administration of the test, students responded to the attitude scale toward computer testing and the five-item state anxiety scale. The revised computer test was then taken by each student individually at the computer terminals.

4. Posttask Period. As with the previous studies, the students then completed the five-item state anxiety scale on-line, and the paper and pencil attitude scale concerning feelings toward computer testing. The students were then debriefed and dismissed.

Results

The results section is divided into three major sections. The first section provides descriptive performance data from the computer test. State anxiety data in relation to experimental conditions is presented in the second section. The final section deals with the evaluation of student attitudes.

Descriptive Performance Data

As with the previous studies, the IQ scores were obtained by a skilled examiner reviewing the detailed item performance records of each student. The mean IQ score for the sample of students was found to be 118.5 with a standard deviation of 9.92. As the reader may note, this performance data is quite consistent with the previous data on college students for the same computer test.

Evaluation of Anxiety Levels

To evaluate the differential impact of stress on state anxiety in the nonevaluative game period and evaluative computer test session, a 2 x 2 ANOVA within subjects repeated measures analysis was conducted (Winer, 1971). The independent variables were game versus testing situation and pre-post anxiety measures. Table 34 presents the means and standard deviations of the state anxiety scores.

Table 34

Means and Standard Deviations of STAI State Anxiety Scores by Experimental Sessions and Pre-Post Periods (N = 60)

Experimental Sessions	Pretest	Posttest
Game		
Mean	8.07	7.47
SD	2.83	2.90
Computer Test		
Mean	9.63	10.40
SD	3.02	3.72

The results of the analysis indicated that the game led to lower levels of state anxiety in comparison to the computer testing experience. This overall main effect was statistically significant as indicated in Table 35. In addition, a significant conditions by pre-post interaction was found.

This interaction is graphically depicted in Figure 12. As the reader may note, state anxiety scores decreased slightly over the nonevaluative game period, whereas the state anxiety scores increased for the same students during the computer test. Thus, during a nonevaluative session students responded with decreasing levels of state anxiety while during the computer test, their state anxiety scores increased during the test.

A subsidiary analysis was conducted to determine if sex of the student interacted with state anxiety, and thus modify the general trend of the data. This analysis, however, revealed no evidence of such an interaction.

Attitude Scores

The means for the pretest and posttest administrations of the attitude scale were 73.4 (sd = 12.0) and 70.2 (sd = 11.5), respectively. These results are comparable to those attitude scores found in Studies II and III, with the exception of the low anxiety subjects in Study III, in that once more a slight decrease in attitude scores from the pretest to the posttest measure is found; however, these attitude scores are somewhat higher than those found in Studies II and III. These findings will be examined further in the discussion section.

Discussion

Game

The primary purpose of this study was to evaluate the effectiveness of prior terminal experience in the form of a nonevaluative computer game on the reduction of state anxiety. The results of the study clearly indicated that there was an overall difference in the amount of state anxiety evoked within the two experimental conditions with the game experience leading to lower levels of state anxiety in comparison to the computer testing situation.

However, the question of most importance to the purposes of this study was the efficacy of terminal game experience on the reduction of subsequent anxiety reactions within the computer test. Suggestive evidence is provided for this if a comparison of the obtained state anxiety data in this study with similar data from previous studies is made.

Figure 13 presents a graphical comparison of the pretest and posttest state anxiety scores from this study, the same scores obtained from the revised

Table 35

Analysis of Variance of STAI State Anxiety Scores by Experimental Conditions and Pre-Post Measurement Periods

Source	df	MS	F
Between Subjects	59	23.86	
Within Subjects			
Conditions (C)	1	303.75	51.4**
Subjects x C	59	5.91	
Pre-Post Periods (P)	1	.42	< 1
Subjects x P	59	6.00	
C x P	1	28.02	6.59*
Subjects x CP	59	4.25	

*$p < .05$
**$p < .01$

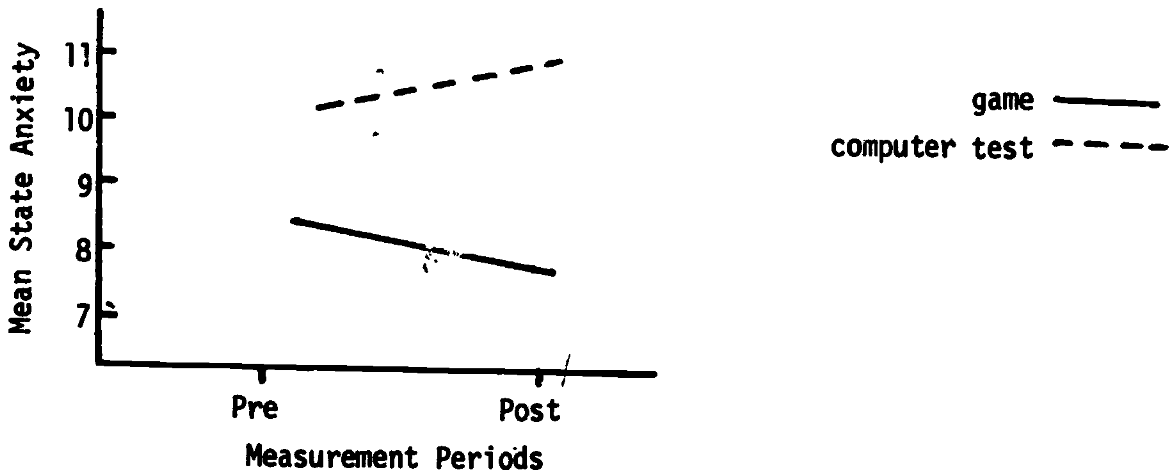

Figure 12. Interaction between experimental conditions and pre-post measurement periods.

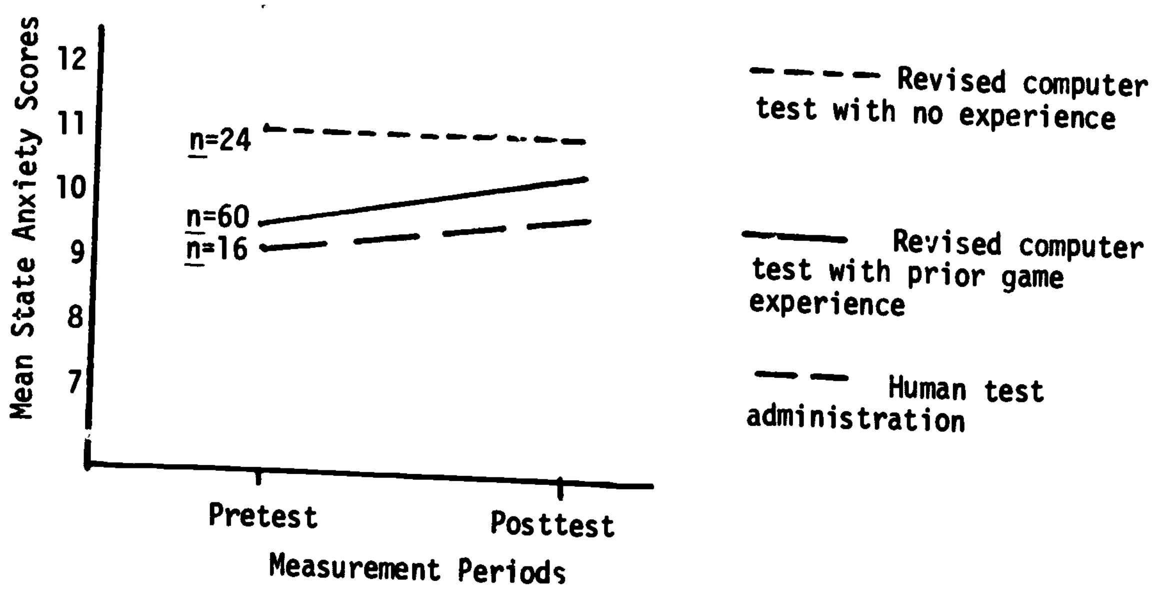

Figure 13. Comparison of state anxiety scores on the computer test as a function of prior terminal experience.

program used in Study III, and the data previously obtained from examiner administrations of the Slosson Intelligence Test. Inspection of Figure 13 indicates that the students who received terminal experience in the form of a game showed lower state anxiety in the computer testing situation than those students in Study III who took the revised computer test but had no prior terminal experience. The mean state anxiety scores of the students with prior game experience for the pretest and posttest were 9.63 and 10.40, respectively. For the students with no prior terminal experience, the same means were 11.0 and 11.1. Thus, the results suggest that the terminal game experience was somewhat effective as an anxiety reduction technique.

Even though there is evidence concerning the positive effects of terminal experience on subsequent stress level in computer testing as measured by the STAI state anxiety scale, one must consider these results as only suggestive due to the possible differences in experimental samples and the limitations inherent in post hoc analysis.

Further evidence concerning the effects of terminal experience in terms of a game experience can be deduced from a similar comparison of attitude scores from this study with the previous data obtained from examiner testing with the same test instrument. These data are presented in Figure 14. Also included is the data from Study III for students who completed the revised computer test with no prior terminal experience.

As the reader may note, there appear to be different attitude scores as a function of prior terminal experience. Students who had terminal experience prior to completing the revised computer test showed similar attitudes toward computer testing as the students in the validation study had expressed toward examiner testing. Conversely, students who took the revised computer test without the benefit of prior terminal experience reported less favorable reactions in comparison to the examiner testing group.

This interpretation is given further support if one views the consistencies between the state anxiety and attitude scores. Students who had prior experience tended to show lower state anxiety scores than the students who took the revised computer test without previous terminal exposure and practice. In addition, the students with prior experience also showed higher attitude scores than those without terminal experience. Thus, the prior experience of the nonevaluative game not only appeared to reduce state anxiety but also to lead to more favorable attitudes in that the testing situation was probably perceived as less stressful in nature.

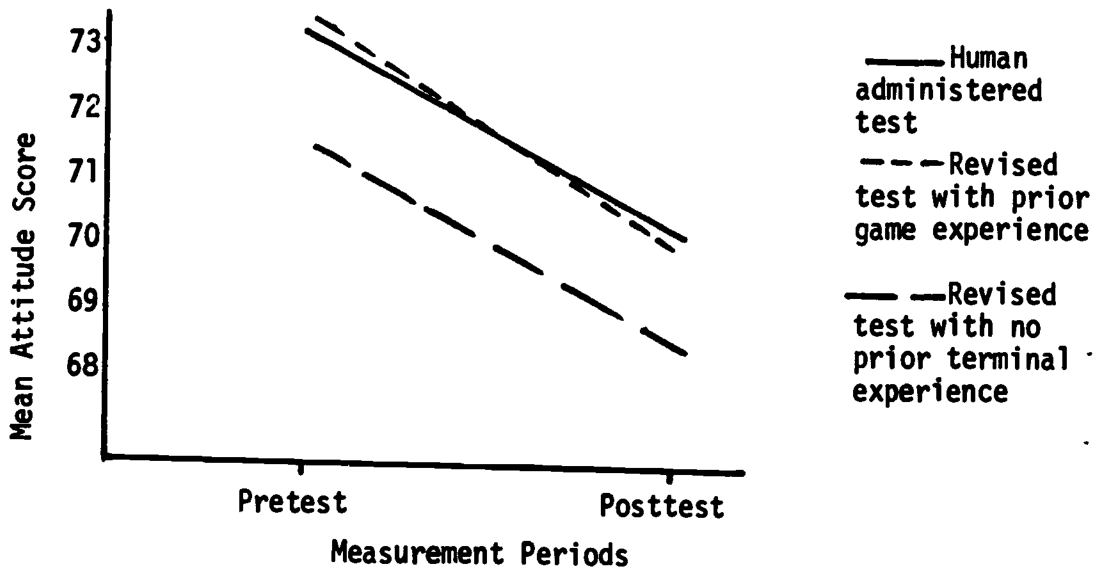

Figure 14. Comparison of attitude scores as a function of different administration procedures.

STUDY V

EFFECTS OF PRIOR TERMINAL EXPERIENCE VIA A LEARNING SITUATION ON SUBSEQUENT AFFECTIVE INDICES IN COMPUTER-BASED INTELLIGENCE TESTING

The results of Study IV suggested that terminal experience in the form of a nonevaluative game situation acted to decrease the initial anxiety indices in computer testing. Thus, it appeared that some form of practice with the operation of the computer would serve to reduce the students' anxiety and thus improve the validity of the intelligence testing.

In conjunction with this hypothesis, previous research in the areas of transfer of training has suggested that the positive transfer effects of practice are dependent upon the similarity of the important task elements within the practice and task conditions. Thus, the more similar the practice period is to the actual task conditions, the greater the positive transfer effects.

In order to relate this rationale to the present research and thus attempt to further reduce anxiety in the testing situation, Study V was designed to provide the students with prior terminal experience, as in Study IV, but to make this prior experience more similar to the actual testing situation. As the computer game used in Study IV involves relatively few cognitive skills and is, by its game-like nature, a relatively non-stressful condition, it differed greatly in substance from the intelligence test itself. In addition, and perhaps even more importantly, the computer game required little use of the terminal keyboard except for the simple typing of numbers, while the intelligence test required a much greater interaction with the keyboard. Thus, while the computer game did serve to reduce anxiety somewhat, it was anticipated that some type of prior experience which was more similar to the testing experience would be even more beneficial.

In order to investigate this hypothesis, Study IV used a CAI instructional program as prior experience for the subjects. This program involved greater use of the keyboard by requiring typed phrase-like responses and, in addition, the cognitive-type skills involved, while somewhat simple, were more similar in kind to the types of behavior involved in an intelligence test. This program will be described in depth in a later section.

Method

Subjects

A total of 30 educational psychology undergraduate students participated in this experiment (10 males, 20 females). Course credit was given by their instructors for voluntary participation in the study.

Apparatus

The IBM 1500 system as used in the previous studies was also used in this experiment. See Study I for a further description.

Computer-Based Intelligence Test and CAI Learning Program

As in the previous studies the Slosson Intelligence Test with the revisions made in Study III was used as the intelligence measure in this experiment. For a further description of this test, see Study II.

In addition, a CAI learning program was given in order to provide the students with some prior terminal experience.

This instructional program consisted of 55 frames dealing with the incidence and prevalence of heart disease and of fatalities resulting from coronary attacks (Tobias, 1969). This program covered, in non-technical language, various risk factors with respect to contracting heart disease such as smoking, cholesterol, tension, and lack of exercise. In conception this program deals in a systematic manner with material on heart disease which is widely available in the public media. This fact was substantiated in that pretest data for a college population approached 33 percent familiarity with the material (Tobias, 1969). Thus, it would appear that these materials are relatively easy for a college population, an interpretation also substantiated by less than a 3 percent error rate for these materials (Tobias, 1969).

Affective Measures

As in Study IV, the State-Trait Anxiety Inventory, the Test Anxiety Scale, and the attitude scale dealing with feelings toward computer testing were used as affective measures in this study. (See Study II for a further description of these measures.)

Procedure

The experimental procedures were administered in the following sequence for all students: (1) a pretask period during which students responded to the Test Anxiety Scale (Sarason, 1958) and the trait anxiety scale of the State Trait Anxiety Inventory (Spielberger, 1972); (2) a prior experience period during which students worked with a CAI learning program on the computer terminal; (3) an evaluation period during which the Slosson Intelligence Test was administered via computer; and (4) a posttask period during which the students completed the short form (five item) STAI state anxiety scale and the attitude scale. The five-item STAI state anxiety scale was administered before and after both the CAI learning program and the intelligence test. An attitude scale was also administered immediately before and after the computer intelligence test. This procedure was identical to that used in Study III except for the nature of the material during the prior experience period. Each of these periods is further described below.

1. Pretask Period. During this period, the students first completed the Test Anxiety Scale and the trait anxiety scale of the STAI. The students then received written instructions informing them how to operate the computer terminal. Practice in the operation of the terminal keyboard was presented and students "signed on" to the CAI learning program. Immediately after "signing on," students responded to the five-item state anxiety scale with instructions to respond in terms of their present anxiety feelings. Students then began working with the CAI instructional program.

2. Prior Experience Period. During the instructional program, students worked individually at computer terminals. They constructed their responses to the individual frames and received feedback or confirmation of their answers immediately when the correct answer was flashed on the cathode ray tube screen to be followed by the next frame.

At the conclusion of the instructional program, students completed a paper and pencil administration of the five-item state anxiety scale with retrospective state instructions to indicate "How did you feel during the material you just studied?" The instructional period was defined by a time interval of 20 - 25 minutes and not by completion of the 55 frame program.

3. Evaluative Period. The experimental procedures used for the administration of the Slosson Intelligence Test were identical to those used in the previous studies with the revised program. Students first responded to the attitude scale toward computer testing and the short form (five-item) state anxiety scale. The revised test was then taken by each student individually at the computer terminals.

4. Posttask Period. Following completion of the test, students again responded to the five-item state anxiety scale and attitude scale. The students were then debriefed and dismissed.

Results

The results section is divided into three major sections. The first section provides descriptive performance data obtained from the computer test. The second section deals with the analysis of the state anxiety data as a function of experimental conditions. Finally, descriptive data is presented concerning student attitudes toward the computer testing situation.

Descriptive Performance Data

In accord with the previous studies, the IQ scores were obtained by a skilled examiner reviewing the detailed item performance records of each student according to the test manual. A mean IQ score of 121.9 (sd = 9.03) was found for this sample of 30 college students. The reader may note that this data is in relative agreement with the data in the previous studies using the same computer test.

Evaluation of Anxiety Levels

To evaluate the different impact of stress on state anxiety in the instructional program and evaluative computer test, a 2 x 2 ANOVA repeated measures analysis was conducted (Winer, 1971). The independent variables were instructional program versus testing situation and pre-post state anxiety measures. Table 36 presents the means and standard deviations of the state anxiety scores.

The results of this analysis are presented in Table 37. The most salient finding in this analysis was the significant interaction between experimental conditions and pre-post periods. As the reader may note, state anxiety scores decreased over the instructional program, remained at this level immediately prior to the computer test, but then increased during the computer test. The main effects for experimental conditions and pre-post periods were not found to be significant.

The significant interaction is graphically depicted in Figure 15. State anxiety scores decreased slightly during the CAI instructional program and increased during the computer test, indicating the operation of stress factors in the test.

Attitude Scores

The mean for the pretest and posttest administrations of the attitude scale were 71.0 (sd = 15.8) and 69.7 (sd = 13.9), respectively. These scores were relatively equivalent with the prior data on both students with prior game experience and the examiner administrations. These comparative results are presented in more detail in the discussion section of this study.

Discussion

The major purpose of the present study was to evaluate the effectiveness of prior terminal experience in the form of a CAI constructed-response instructional program for anxiety-reduction within a subsequent computer intelligence test. The results of the study showed a significant pre-post by conditions interaction indicating the differential impact of stress on state anxiety in the CAI instructional program and testing situation. Clearly, state anxiety scores declined in the learning situation of easy materials and increased in the testing situation. This finding supports predictions from Trait-State Anxiety Theory (Spielberger, 1972).

However, the question of more considerable interest here concerned the effects of prior terminal experience and observed state anxiety within the computer testing situation. Although this effect cannot be directly evaluated in the context of the present study, inferential evidence can be obtained if comparisons of state anxiety scores are made with previous data. Figure 16 presents a description of pre-post state anxiety scores from Study III, Study IV, the validation study, and from the present study.

Table 36

Means and Standard Deviations of STAI State Anxiety Scores by Experimental Sessions and Pre-Post Periods (N = 30)

Experimental Sessions	Pretest	Posttest
CAI Instructional Program		
Mean	9.43	8.93
sd	3.38	3.60
Computer Test		
Mean	8.60	9.80
sd	3.96	3.69

Table 37

Analysis of Variance of STAI State Anxiety Scores by Experimental Conditions and Pre-Post Measurement Periods

Source	df	MS	F
Between Subjects	59	41.06	
Within Subjects			
Conditions (C)	1	.01	< 1
Subjects x C	59	6.28	
Pre-Post Periods (P)	1	3.675	1.25
Subjects x P	59	2.95	
C x P	1	21.675	4.12*
Subjects x CP	59	5.26	

*p .05

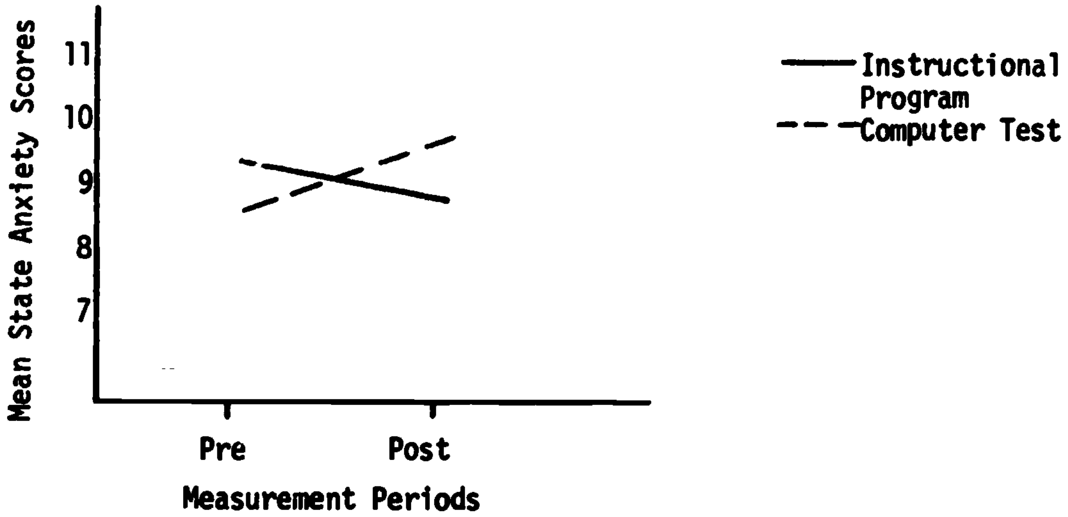

Figure 15. Interaction between experimental conditions and pre-post measurement periods.

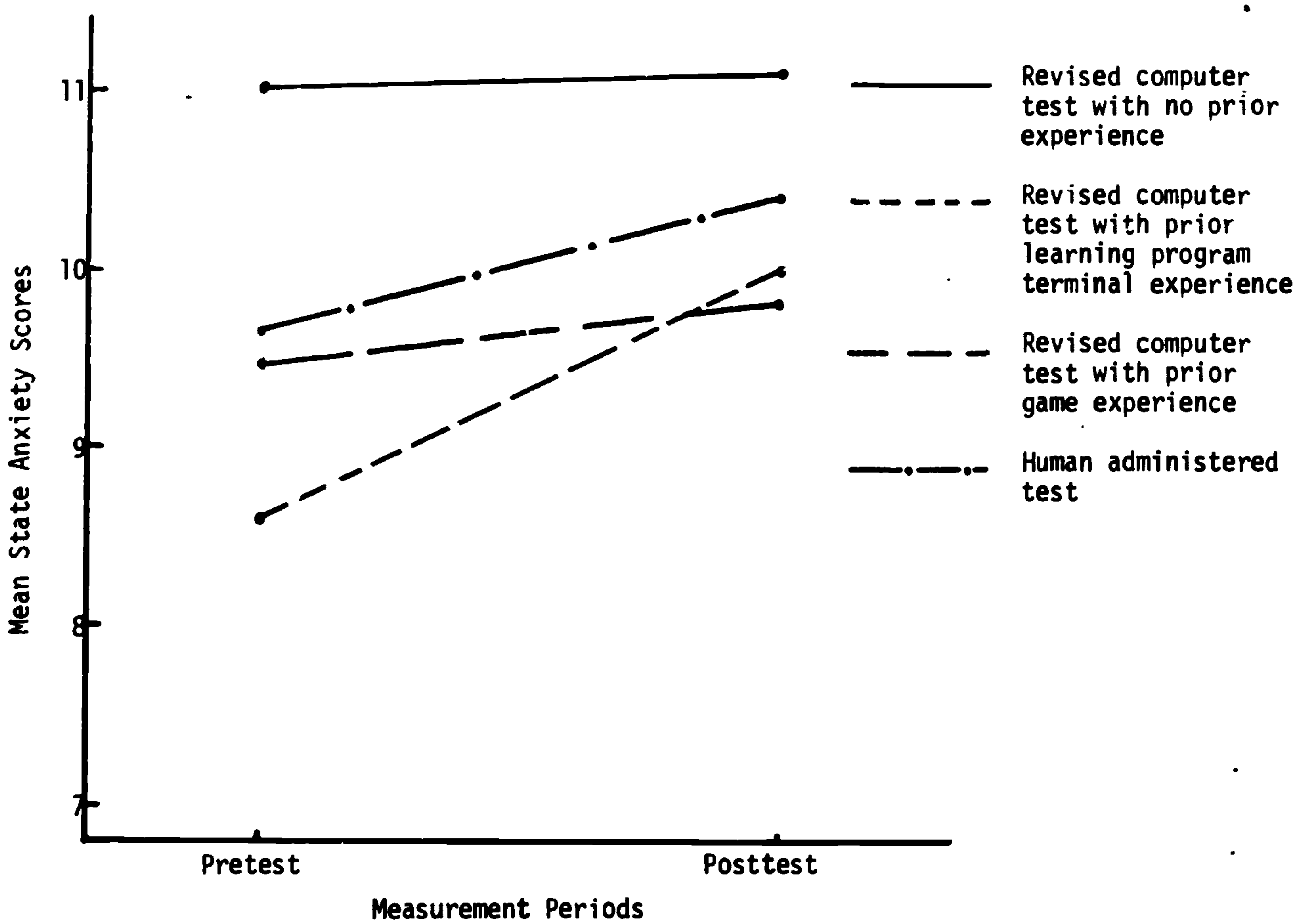

Figure 16. Comparison of state anxiety scores as a function of different administration procedures and prior experience.

The results of this comparison provide some indirect evidence for the effectiveness of prior terminal experience in terms of anxiety-reduction benefits. It may be remembered that the revised administration procedures within the computer test lead to reduced anxiety levels in comparison to the validation study results and the replication (Study II). However, the observed state anxiety scores were still more elevated than when an examiner administered the same test.

In viewing Figure 16, it appears that prior terminal experience either in the form of a computer game or CAI instructional program resulted in reduced anxiety levels compared to the results obtained on the computer test without any prior terminal experience. In addition, these results suggest that the type of terminal experience may be important for anxiety reduction effectiveness. Experience in the form of a computer game led to anxiety results quite similar to the examiner results. However, the experience with a CAI constructed response instructional program seemed to produce even lower state anxiety scores within the computer test than the examiner tests. These data suggest that as previously hypothesized, the effectiveness of prior experience would be partially a function of the degree of task similarity between the prior terminal procedures and the nature of the procedures within the testing situation.

During the computer game, students had minimal practice in the use of the terminal keyboard for entering their responses. Their primary interaction with the terminal was confined to simple typing of numbers. However, more extended practice in typing their responses was required on the CAI instructional program. Students had to construct more phrase-type responses for the learning program, and this type of responding was more similar to the operations required during the computer test in which the students had to define vocabulary words in their own words and answer other questions requiring more extensive verbal answers. This difference in terminal experience seemed to have differential anxiety-reduction benefits in computer testing as indicated by the state anxiety data.

In the previous study, it was found that terminal experience in the form of a computer game appeared to reduce the subsequent stress in the computer intelligence test as indicated by the lower state anxiety scores and more favorable student attitudes. Results of the present study suggested that terminal experience in the form of a constructed response instructional program was more effective as an anxiety reduction technique than was the computer game. This effect also seemed to be somewhat consistent in the attitude scores.

Figure 17 presents a comparison of the attitude scores from Studies III and IV, and the results of the present study. Also included is the data from the previous examiner-administered tests (Hedl, 1971). From these data it would appear that the students who had terminal experience with a CAI learning program showed somewhat lower pretest attitude scores toward computer testing. However, the posttest scores for these students were relatively consistent with the same measure for the students who had the Slosson Intelligence Test administered to them by an examiner and the students who had terminal game experience prior to

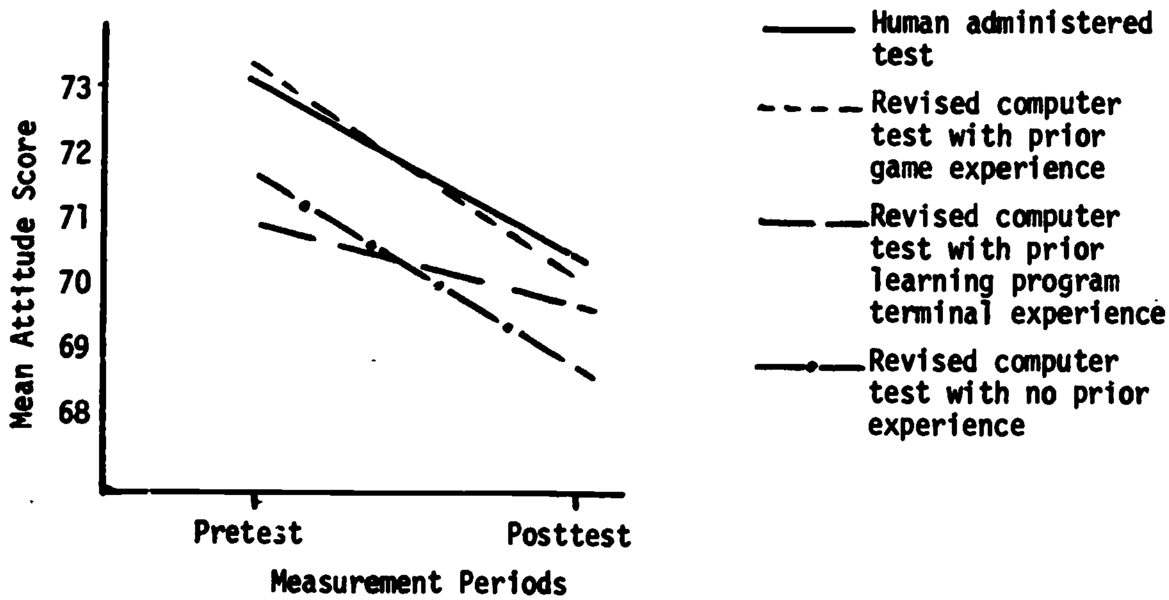

Figure 17. Comparison of attitude scores from Studies III, IV, and V, and from the previous examiner-administered tests.

taking the revised computer administration of the same test. The posttest attitude means for these groups were 69.7, 70.5, and 70.2, respectively. Thus, while the attitude scores of the students given terminal experience with a CAI learning program are lower than the other groups' scores on the pretest measure, the subsequent decrease in scores on the posttest measure is less than that of the other three groups, resulting in [illegible] scores for the two groups with terminal game experience and the human-a[illegible]stered group.

Inspection of these data also suggests that prior terminal experience, regardless of type, resulted in more favorable student attitudes toward computer testing. This can be seen by viewing the attitude results of those students who took the revised computer test without any prior terminal exposure. It may be noted that these attitude scores were lower than either of the two groups with differential terminal experience or the group with human examiners. As previously described, these latter three groups were relatively equivalent in their posttest attitude scores.

In conclusion, the results of this study suggest that some form of prior terminal experience acts to both reduce state anxiety and improve attitudes during a computer-based intelligence test. In addition, the CAI learning program, which is more similar to the intelligence test itself in the type of terminal activity required, was even more effective as an anxiety reduction technique than the computer game, and, in fact, compared favorably with the human-administered test in reducing anxiety.

STUDY VI

EFFECTS OF PRIOR TERMINAL EXPERIENCE VIA A CMI PROGRAM ON SUBSEQUENT AFFECTIVE INDICES IN COMPUTER-BASED INTELLIGENCE TESTING

The results of Study V suggest that, as found in Study IV, some prior terminal experience is beneficial in reducing anxiety and improving attitudinal scores in a computer-based intelligence testing situation. In addition, Study V appears to indicate that the advantages of the prior terminal experience are increased as this experience becomes more similar to the actual testing situation.

Study VI was, therefore, designed to test an extension of these findings. Accordingly, if subjects are given extensive practice on a computer terminal such that they become very familiar with its operation and if this practice involves a situation of some stress, similar to the intelligence testing situation itself, the resulting anxiety and attitudinal scores may be improved even further than in the previous studies.

The authors were given the opportunity to test this hypothesis due to a relatively extensive CMI instructional program being conducted in the undergraduate educational psychology course at The University of Texas at Austin. The program gave the subjects extensive experience in a computer-managed instructional situation under stressful conditions. This program was considered as prior terminal experience analagous to that in the previous studies, making it possible for the authors to test the extension of the results of Studies IV and V.

An additional objective of this study was to obtain some additional information concerning the specific cues in both CMI testing and intelligence testing which elicit anxiety in the student. This was done by conducting structured interviews with each subject following the testing period. It was anticipated that through such interviews additional methods to reduce anxiety might be determined.

Method

Subjects

A total of 42 students (15 males, 27 females) participated in the experiment. The sample was drawn from an undergraduate educational psychology course; volunteers received course credit for their participation.

Apparatus

An IBM 1500 computer-assisted instructional system was used for the presentation of all computer-managed module tests and the five-item state anxiety measure. Test items and anxiety measures were presented to the students and all responses were recorded at the system's terminals.

CMI Module Testing Experience

During the computer-managed instructional portion of introductory educational psychology, students spent about one month studying eight different instructional modules: (1) Computers in Education; (2-7) Psychological Statistics; (8) Cultural Differences; and (9) Classroom Management. Guided by behavioral objectives, each subject studied conventional printed learning resources until he finished the readings for a module; then he came to the CAI Laboratory to take a test, also keyed to the objectives, on a computer terminal. He received feedback at the end of the test consisting of his total score on the module and his percentage score on each objective. If he failed to meet a 75% criterion, he was required to take a retest at a later date. This situation was considered a potentially stressful one for the student, since course grades were affected by module test scores.

As in the previous studies, the Slosson Intelligence Test, with the revisions made in Study III, was used as the intelligence measure.

Affective Measures

As in the previous studies, the short form (five-item) state anxiety scale of the State-Trait Anxiety Inventory was used to measure anxiety both before and after the testing situation. See Study II for a further description of this scale.

In addition, the students were administered the Test Anxiety Scale prior to taking the CMI modules.

Procedures

The experimental procedures were administered in the following sequences: (1) a prior experience period in which students took eight computer-administered module tests, each test followed by the five-item version of the state anxiety scale of the State-Trait Anxiety Inventory (Spielberger, 1972); (2) an evaluation period during which the Slosson Intelligence Test was administered via computer (Hedl, 1971), and the five-item state anxiety scale was administered before and after the intelligence test; and (3) a posttest period during which participants responded to a structured interview concerning the nature of any anxiety they experienced during the intelligence test. Each of these periods is further described below.

1. Prior Experience Period. During this period of about one month the student took each of the eight module tests as described earlier. If he failed to reach criterion he was required to take a retest. At the time of his retest, he was tested only over the previously failed objectives. At the end of each test he responded to the five-item state anxiety scale before receiving his performance score. Since the number of items on retests varied across students, only the data from the first testing was considered in the present study.

2. Evaluative Period. The experimental procedures used for the administration of the Slosson Intelligence Test were modified somewhat from those used in previous studies with the revised program. This modification consisted of the deletion of a number of paragraphs preceding the test which instructed the student in the use of the CRT terminal. Since the participants had had adequate prior experience in the operation of the terminal keyboard, it was felt these instructions were unnecessary. Otherwise the procedures of administering the revised intelligence test itself were identical to those used in previous studies using this program. In addition to taking the test, students responded to the five-item state anxiety scale immediately before and after the test, and their answers were recorded automatically.

3. Posttest Period. In a brief period following all testing proeedures, the students were asked to respond to a 10-minute structured interview concerning any anxiety they might have experienced during the test. Interview questions were arranged so that the nature of anxiety as expressed on the five-item state anxiety measure might be further specified.

Results

The results section is divided into four sections. The first section describes state anxiety data obtained after each of the eight module tests preceding the intelligence test. The second section gives descriptive performance data obtained from the computer-administered intelligence test. In the third section, the state anxiety data collected immediately before and after the intelligence test will be compared with previous studies as a function of prior terminal experience. Finally, qualitative data drawn from interviews following CMI module tests and the computer-based intelligence test will be summarized.

Anxiety Levels During Module Tests

The mean state anxiety levels of subjects in this experiment are listed in Table 38. As may be noted, the mean state anxiety scores of students taking the eight CMI module tests were higher than posttest state anxiety scores in the previous study (Study V), in which students practiced on a CAI instructional program before the computer-administered intelligence test. The mean state anxiety score following practice on the CAI instructional program was 8.93 (sd = 3.60), as compared to a grand mean state anxiety score of 11.91 (sd = 4.04) following the CMI module tests. This elevated state anxiety may be attributed to the real evaluative stress of each CMI module test, which counted toward a course grade, as compared to the practice CAI instructional program, which did not affect a course grade. It was hoped that experience during a CMI treatment of this kind might be more realistic practice for the evaluative stress of the intelligence test, and thus permit a desensitization to state anxiety in a computer terminal situation prior to the intelligence test. However, as will be discussed later, individual differences in responding to the evaluative stress of a course test versus an intelligence test were discovered during the structured interview. Thus practice in course testing did not necessarily better prepare a student for intelligence testing, since some students considered these situations to be two distinct forms of evaluative stress.

Table 38

Means and Standard Deviations of STAI State Anxiety Scores Following Eight CMI Module Tests and a CAI Instructional Program

Module Test	1	2	3	4	5	6	7	8
Mean	11.93	12.00	11.63	11.88	12.76	11.88	12.07	11.13
<u>sd</u>	4.51	3.60	4.11	4.37	3.51	4.44	4.26	3.52
Grand Mean	11.91							
Grand <u>sd</u>	4.04							

CAI Program	
Mean	8.93
<u>sd</u>	3.60

Descriptive Performance Data

In accord with previous studies, the IQ scores were obtained by a skilled examiner reviewing the detailed item performance records of each student according to the test manual. A mean IQ score of 121.48 (sd = 12.24) was found for this sample of 42 college students. These findings are in relative agreement with previous studies using the revised intelligence test.

Evaluation of IQ Test Anxiety Levels

Due to a programming error, one item of the five-item state anxiety scale was not presented after the computer IQ test in Study VI. In order to provide results that could be compared to Studies III, IV, and V, an average score was computed from the four-item posttest state anxiety scale for each of the 42 students, and this average score was added to the four-item total to produce a five-item total for each subject. The validity of this procedure was checked by correlating the four-item posttest scale with the reconstructed five-item scale, which resulted in $r = .97$ ($p < .001$), and by comparing the alpha-reliability (internal consistency) of the four-item scale, the reconstructed five-item scale, and the regular five-item scale, all used as posttest measures of state anxiety after the intelligence test. As can be seen in Table 39, alpha-reliability figures for the three scales are comparable.

Table 40 specifies the means and standard deviations of state anxiety scores for subjects in computer tests who had (1) no prior experience (Study III); (2) prior experience of a computer game (Study IV); (3) prior experience of a CAI instructional program (Study V); and (4) prior experience of eight CMI module tests. Comparisons of the different state anxiety levels are illustrated in Figure 18.

As can be seen from Figure 18, students who received prior experience in the form of CMI module tests showed lower state anxiety scores before the computer-based intelligence test than students in Study V, who received prior experience in the form of a CAI instructional program. This latter group, however, showed lower state anxiety scores than students in Study IV, who received prior experience in the form of a computer game. All students who received one of the three forms of prior experience reported lower state anxiety scores than students in Study III, who received no prior experience.

Unlike previous studies, however, students in the present study who reported the lowest state anxiety scores before the computer-administered intelligence test did not report the lowest state anxiety mean scores after the computer test. Those students who had CMI prior-testing experience reported higher mean state anxiety scores than both student groups with prior experience in the form of a computer-based game and a CAI instructional program, although the CMI prior experience group scores were still lower than the student group with no experience.

Table 39

Kuder-Richardson 20 Reliability Estimates for Three Posttest Administrations of the State Anxiety Scale: Four-Item, Reconstructed Five-Item, and the Regular Five-Item Scale

Scale	N	Pre	Post
Four-Item (Study V)	42	.87	.92
Five-Item (Reconstructed) (Study V)	42	.90	.96
Five-Item (Study II)	30	.87	.89

Table 40

Means and Standard Deviations of STAI State Anxiety Scores by Pre-Post Periods in Four Administrations of the Computer-Based Slosson Intelligence Test

Study	Pretest	Posttest
Study III: No Prior Experience		
Mean	11.0	11.1
Study IV: Game Prior Experience		
Mean	9.63	10.40
sd	3.02	3.72
Study V: CAI Prior Experience		
Mean	8.60	9.80
sd	3.96	3.69
Study VI: CMI Prior Experience		
Mean	7.95	10.74
sd	3.48	4.53

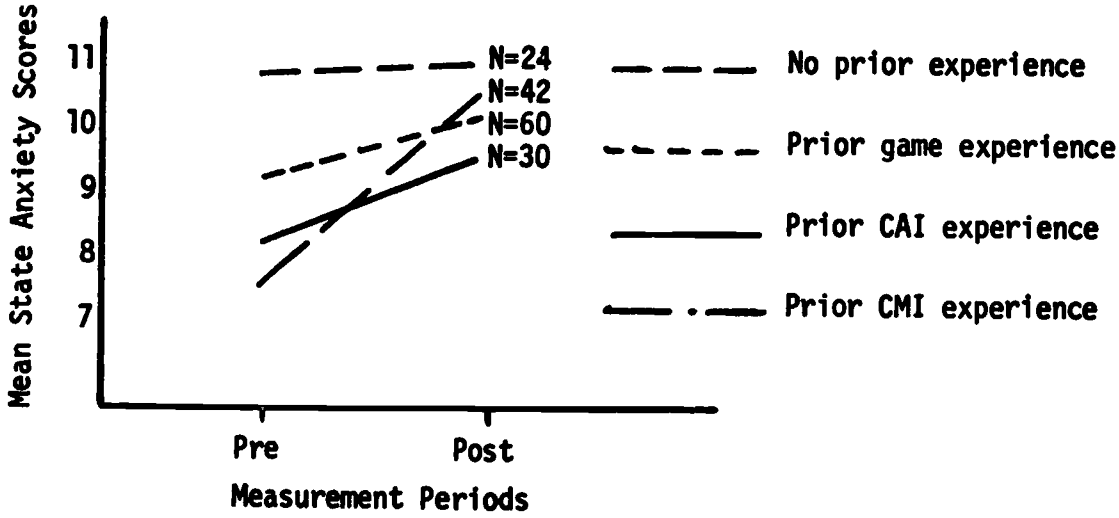

Figure 18. Comparison of state anxiety scores on the computer test as a function of prior terminal experience.

These findings suggest that the CMI-testing prior experience is relatively effective as a reducer of state anxiety immediately before the computer-based intelligence test; however, based on self-reported state anxiety immediately after the intelligence test, the prior CMI testing experience is not as effective in reducing anxiety as prior experience in the form of a computer-based game or a CAI instructional program.

To further analyze these results, the group with prior CMI testing experience was divided into two subsamples, based on categorical data from posttest interviews: (1) those students who reported experiencing more evaluative stress during the intelligence test than during the CMI module tests; and (2) those students who reported experiencing more evaluative stress during the CMI module tests than during the intelligence test. The hypothesis leading to these subsample distinctions was that fear of academic failure (concretely symbolized by a low course grade) is distinct from fear of appearing generally unintelligent, such that the predominance of one behavior over the other would result in differential effects following the same prior experience treatment (CMI module testing). The total sample and subsample group means for Study VI are given in Table 41, and Figure 19 illustrates their relationship.

As may be seen in Figure 19, group pre-post means for the subsample who reported more anxiety during the CMI module tests than during the intelligence test were lower than group pre-post means of the total sample, as well as lower than those of the subsample who reported more anxiety during the intelligence test than during the CMI module tests. The group means of the first subsample are the lowest yet reported, which indicates that prior CMI testing experience can be an effective state anxiety reducer for students who anticipate or report more evaluative stress during a CMI module test than during the intelligence test itself. For students who anticipate or report experiencing more evaluative stress during the intelligence test than during CMI module tests, another form of prior experience, such as the nonevaluative CAI instructional program, would be the preferred treatment.

Interview Data

The need to investigate the specific form of test anxiety that is elicited by situational cues in a computer-administered testing situation has been indicated by Sarason, Kestenbaum, and Smith (1972), who note that:

> . . . self-described high test anxious persons are a particular type of cue-seekers. They search the environment for stimuli that connote or imply evaluation of their achievements . . . persons high in test anxiety differ from others in their attentiveness to certain types of environmental events. Depending upon the stimuli and their reactions to these events, high test anxious individuals may perform either well or poorly. (p. 248)

Table 41

Means and Standard Deviations of STAI State Anxiety Scores by Pre-Post (Computer-Based Slosson Intelligence Test) Periods for Students Reporting Differential Stress from CMI Module Tests and the Computer-Based Slosson Intelligence Test*

Population	Pre	Post
Total (_N_ = 42)	7.95 (_sd_ = 3.48)	10.74 (_sd_ = 4.53)
Reporting IQ Test More Stressful (_N_ = 7)	10.57 (_sd_ = 4.54)	15.18 (_sd_ = 3.78)
Reporting CMI Module Tests More Stressful (_N_ = 21)	7.10 (_sd_ = 2.85)	9.15 (_sd_ = 3.77)

*Total sample statistics also included

Note: Seven subjects of the total sample were not available for interviewing after the computer-based Slosson Intelligence Test; seven other subjects either reported no stress in either situation or could not decide which situation was more stressful.

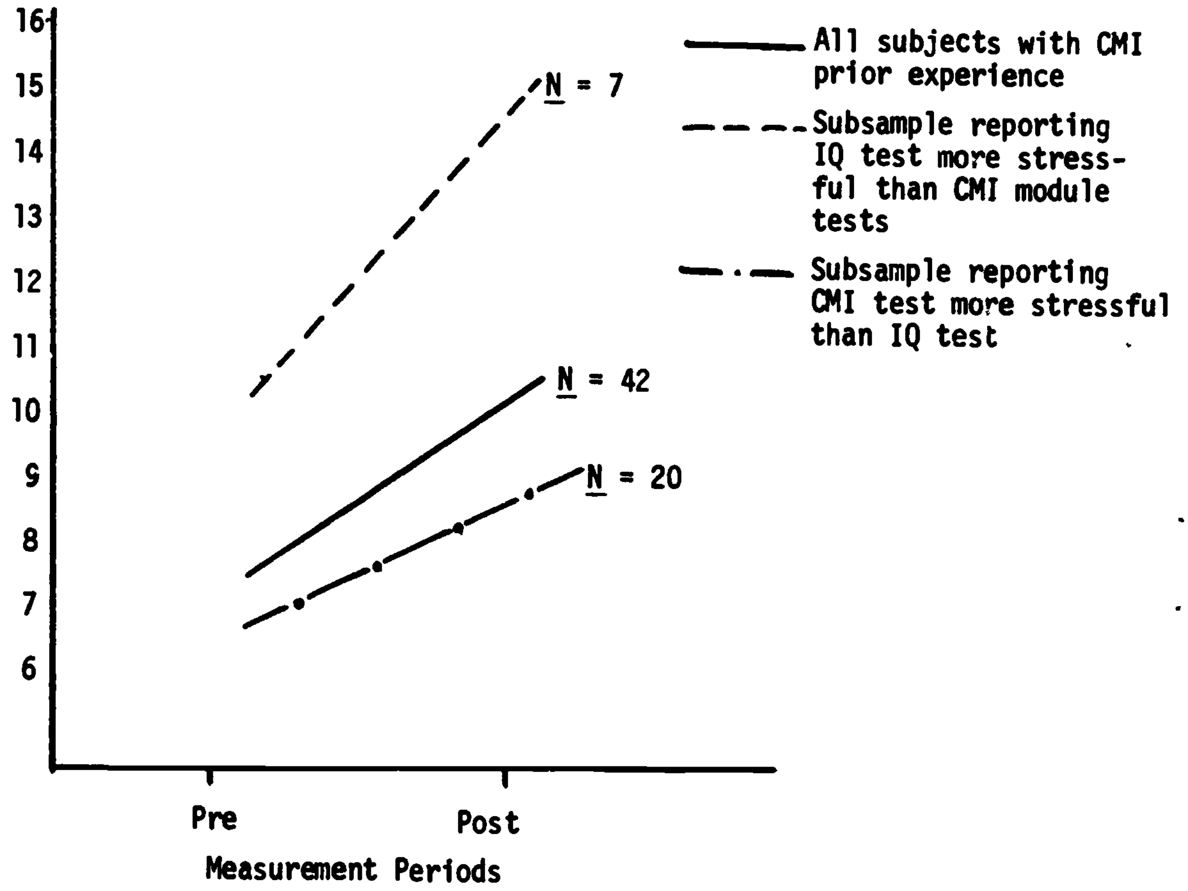

Figure 19. Comparison of state anxiety scores on the computer test between subsamples of the CMI prior experience group (Study VI), based on reported stress in the CMI testing situation and the computer-based Slosson Intelligence Test.

Discussion

Fifty-nine students who took four computer-assisted module tests were interviewed immediately after one of their tests to identify any cues that elicited test anxiety. The structured interview questions focused on the CMI terminal apparatus, overt test-taking behaviors, and internal and external "self-talk" as sources of anxiety cues. A distinction was made between anxiety responses expressed as cognitive concern over performance (worry) and those expressed as autonomic arousal due to the stress of the immediate testing situation (emotionality) (Liebert & Morris, 1967).

Fifteen students reported that random noise in the testing room was distracting, particularly during difficult test items. Ten students noted that the flickering or blurriness of the letters on the CRT screen made them anxious, while six other students reported that the varying time interval between their response entry and the presentation of the next frame made them feel uneasy. Actually, this varying system response time is due to the time-sharing aspects of the IBM 1500 computer, but the six students interpreted this ambiguous cue as evaluative of the correctness of their responses. This finding lends empirical support to Sarason et al.'s (1972) description of the high test-anxious individual as a cue-seeker who searches the environment for stimuli that connote evaluation.

Few overt behaviors were reported to be associated with high test anxiety. Three students reported becoming quite restless, another searched for a cigarette, while others reported hunching toward the screen or wringing their hands. These data agree with the findings of Snyder (1972), who found no correlation between anxiety, as measured by an adapted trait anxiety scale of the STAI, and body movement in a classroom test-like situation. Snyder concluded that the task-irrelevant responses of the high test anxious individual should occur at the internal, cognitive level.

Data from the interview protocols lend support to Snyder's conclusion. When asked to cite any debilitating self-talk that occurred during the CMI tests, several students reported self-disparaging comments which ranged from mild self-recriminations to panicked cursing. Examples of the debilitating self-talk included:

> "You should know the stuff." "What if I fail or get a zero?" "Oh, you dummy!" "I'll bet you're the only one who fails this stupid thing!" "Damn it, I've got to get this thing right."

Whereas only five students reported overt behaviors associated with their anxious feelings, thirty-three students reported self-talk of a self-accusatory or distracting nature.

Not all self-talk was debilitating, however. Eleven students reported self-talk that calmed or managed their test anxiety. Examples of this facilitative self-talk included:

> "Slow down and read it again; it isn't that difficult." "Now don't let this get to you; it's only a test." "If I fail it, I'll just take it over again." "Now be careful--don't go off half-cocked."

A second structured interview was administered to thirty-one students who volunteered to take the computer-based Slosson Intelligence Test (Hedl, 1971). In addition to the appropriate items of the first interview, this second interview asked if the test anxiety elicited during a computer-administered intelligence test was different from test anxiety elicited during a computer-administered course examination.

As mentioned in a previous section, twenty students reported that the CMI tests were more stressful than the computer-based intelligence test, while seven students reacted in an opposite fashion, finding the intelligence test more threatening than the CMI module tests. The principal reason reported by those who found CMI module tests relatively more stressful was the pressure of a course grade, which was not present in the intelligence test. In addition, the thought of a module retest was stressful for some students, since a retest connoted a failure on the first test.

Students who felt that the IQ test was relatively more stressful than the module tests reported that the broad assessment of ability implied in the IQ test threatened them more than the specific assessment of knowledge in each module test. In addition, two students felt anxious because they could not take the IQ test again if they scored poorly the first time; this appeared to be a test-taking strategy they had adopted from the CMI module tests.

Similar to the interviews after the CMI module tests, few overt behaviors--foot tapping, hand wringing, and chair rocking--were reported to be associated with anxious feelings during the IQ test. Most task-irrelevant, debilitating behavior was reported to be at the internal, cognitive level. Twenty-two of the thirty-one students interviewed reported self-talk that was distracting or self-derogatory. Examples of this self-talk included:

> "You ought to know <u>that</u>!" "Why don't you ever look up anything." "When will this be <u>over</u>?" "Everybody told you you were dumb in vocabulary." "Am I really this stupid?"

In summary, the interviews following the CMI module tests and the computer-based intelligence test suggested that:

(1) Anxiety was elicited by a variety of internal and external cues in both testing situations;

(2) Proportionally more students became anxious during a test counting toward a course grade than during an intelligence test; and

(3) Cognitive rather than overt physical behaviors were reported as the predominant form of test anxiety responses in both CMI testing and intelligence testing situations.

GENERAL DISCUSSION

The major concern of the present research was the impact of computer testing procedures on state anxiety. Findings from the initial validation study of administering the Slosson Intelligence Test (Slosson, 1963) via computer indicated that the computer testing procedures resulted in higher levels of anxiety, as measured by the STAI state anxiety scale (Spielberger et al., 1970), in comparison to examiner administration of the same test (Hedl et al., in press). Thus, the control and elimination of human examiner variables did not appear to necessarily result in a concomitant reduction of anxiety within the testing situation.

These affective results from the validation study were interpreted as being a function of certain procedural considerations involved in the computer administration of the test. It was suggested that higher anxiety reactions were in part a function of subjects' lack of familiarity with the terminal and the failure to terminate testing procedures as a result of student performance. The main importance of these findings was that they alluded to the impact of ancillary variables on state anxiety within computer testing, factors not given appropriate emphasis in the earlier computer testing literature.

The research reported in the present manuscript concerned itself with the delineation of stress factors operating within the computer testing situation along a more empirical basis. The studies represent a programmatic research approach in the understanding of computer testing practices and procedures. Study II was concerned with a replication of the previous anxiety research and the establishment of the Slosson Intelligence Test via computer as a viable research instrument. Study III dealt with the effects of computer program procedural revisions on the reduction of stress within the testing situation. Studies IV and V explored the efficacy of prior terminal experience in relation to anxiety reduction effectiveness. The final study also investigated the effects of terminal familiarity variables in addition to delimiting the nature of additional stress cues within the computer test.

The primary purposes of Study II were to further replicate the effects of stress on state anxiety within the computer testing situation and to provide additional information on the viability of the use of computers to administer an individualized intelligence test. With respect to the latter, the results of the study indicated a test-retest reliability of .82, a figure quite reasonable considering the nature of the experimental sample. Concurrent validity indices of .37 and .52 were also found during both the first and second test administrations with college grade point average. Thus, it would appear that the psychometric test data support the efficacy of computer administration of psychological tests. This concurrent validity data is consistent with the initial validation results (Hedl, 1971).

Although the psychometric data appears promising, the impact of the computer testing procedures on state anxiety bears closer scrutiny. Anxiety results from the validation study (Hedl, et al., in press) which were replicated in the second study indicated that the computer test and/or testing procedures were causative in the elicitation of stress, and consequently, high state anxiety in comparison to the available examiner data. Thus, controlling human examiner variables by administering a test via computer did not necessarily result in the reduction of anxiety. The results of this study and the validation study would appear to support an opposite interpretation.

Upon reexamination of the state anxiety data, the importance of procedural testing variables comes to the forefront as an alternate hypothesis, an hypothesis also given support from previous computer and anxiety research (Hansen & O'Neil, 1970). It would appear that two major variables, namely the heightened failure experience in the computer test due to the lack of computer administration algoritnms to cease testing when a test ceiling was reached, and terminal familiarity, needed to be further investigated for anxiety-eliciting affects.

The importance of the former testing procedure is shown by considering the fact that students who showed performance ceilings and were still tested with additional difficult items were found to have higher state anxiety scores than those students who did not show a performance ceiling and thus did not conclude the test with as much of a failure experience. This result was found to be significant in the validation study and marginally supported in the first study. The anxiety results of the first study were possibly attenuated by the differences in trait anxiety sample differences. The validation study included the entire range of the trait anxiety distribution while the first study included students with middle trait anxiety scores falling in the semi-quartile range of the distribution. Thus, the first study did not include students with scores falling at either the upper or lower quartile of the STAI trait anxiety normative distribution.

Suggestive evidence was also found concerning the possible importance of terminal familiarity and subsequent anxiety within the computer test. As the reader may note, the state anxiety data from the second administration of the Slosson Intelligence Test via computer was fairly comparable to the previously reported anxiety data obtained with human examiners suggesting the positive effects of prior terminal experience. The relative importance of such terminal familiarity on anxiety reduction may have been confounded with the observed performance increments over the two testing sessions.

Considering the results of Study II, it becomes apparent that a number of factors were conceivably operative within the computer testing process on state anxiety. Study III was designed to clarify these processes and to evaluate a number of test administrative revisions in the presentation of the Slosson Intelligence Test via computer. An additional dimension within Study III was the inclusion of students differing in trait anxiety levels to investigate the possible aptitude-treatment interactive effects of computer presentation methods and procedures. The test revisions were evaluated within a test-retest design on test-retest reliability, anxiety-reduction effectiveness and reduction of testing time.

The results of Study III indicated a test-retest reliability of .72 for the revised computer administration of the Slosson Intelligence Test. Although this figure is slightly lower in comparison to the results of Study II, fluctuations of this type are not uncommon when dealing with samples with restricted measurement ranges. Thus, it would appear that the administration revisions were not detrimental to the stability of the test scores.

Also of interest in Study III was the effects of program revisions on time to complete the test. It was reasoned that a program that was more response-sensitive in terms of administration strategies would reduce the total number of items presented and consequently lead to a reduction in total testing time, a variable hypothesized to be contributory to anxiety elicitation. Results indicated that there was a reduction in the number of items presented; however, total testing time was not significantly decreased. This effect was apparently mitigated in that students also tended to spend more time answering the questions which had longer permissible time limits.

Further analysis of the data indicated that the relationships between testing times and state anxiety were non-significant for both administrations of the Slosson Intelligence Test via computer. These data support the hypothesis that the amount of time spent in the evaluative situation is not a causative stress variable. This hypothsis is also suggested in CAI research, which reported that length of instructional time is not the critical variable for reducing state anxiety and improving performance (Leherissey, O'Neil, & Hansen, 1971b; Leherissey, O'Neil, Heinrich, & Hansen, 1971).

Of more primary interest within Study III was the effects of program revisions on state anxiety within the computer testing situation. Comparisons of the state anxiety scores from Study III with those obtained in Study II and the human test administration scores suggested that the more response-sensitive administration procedures lead to lower mean state anxiety scores than Study II. However, it should be remembered that the revised computer program still produced higher anxiety levels in comparison to the Slosson Intelligence Test administered via examiners. Thus, although the revisions of the computer administration algorithms appeared to be effective for anxiety reduction when compared to the first program (Study II), the revised test still compared less favorably to the anxiety levels observed with human examiners. This interpretation should only be considered suggestive due to the experimental sample differences.

Again, suggestive evidence was noted concerning the possible importance of terminal familiarity and observed anxiety levels. The state anxiety indices from the first administration of the test were significantly higher than those observed from the second administration. The data from the second administration compared favorably to the previously reported human examiner results. Whether these positive transfer effects in terms of the reduction of stress and subsequent state anxiety are attributable to observed performance increases or to terminal familiarity with the test and testing procedures is indeterminable from the research design in Study III. Given that similar results were obtained in both Study II and Study III, the importance of terminal familiarity as an anxiety reduction variable bears further experimentation.

Study III was also concerned with the effects of trait anxiety on state anxiety within the computer testing situation. Results from this study failed to reveal interactions between trait anxiety and the pre-post measurement periods. High anxious students were found to exhibit higher levels of state anxiety, higher hostility scores, and lower attitude scores in comparison to their low anxious counterparts. The lack of statistical interactions indicated that there were no significant differential stress reactions between the high anxious and low anxious students.

In reviewing the results of Study III, it appeared that the more response-sensitive computer administration was somewhat effective for the reduction of anxiety within the computer experience. These test revisions were also found to produce a test-retest reliability index consistent with the previous research on the same test lending more evidence concerning the viability of the Slosson Intelligence Test via computer as a research instrument. Although the test revisions did not result in total testing time reductions due to these effects being lost by item limit considerations, the revisions appeared to have anxiety benefits. The duration of testing time was also not found to have important anxiety correlates.

The major importance of Study III was that it further alluded to the importance of possible pre-task experiences in the form of terminal interactive familiarity and the nature of the interactive testing experience. Studies IV and V sought to more adequately evaluate the importance of pre-task terminal experiences on subsequent stress and state anxiety in the testing situation. Study VI was concerned not only with pre-task terminal experiences, but also with an attempt to determine the nature of the stress cues, if any, in the testing situation.

The results of Studies II and III replicated the affective reactions from the initial validation study. These indices also indicated the possible operation of additional stress variables in comparison to examiner-administrations of the same test. However, the experimental design of these studies (both test-retest in nature) does not allow for the determination of the importance of terminal familiarity in the reduction of subsequent anxiety. It could have been in large part determined by the increase in test performance, therefore causing the second testing to be easier and thus less stressful. Study IV was directed toward an evaluation of the effects of prior terminal experience in the form of a computer game on the reduction of anxiety levels within the evaluative computer test. It was reasoned that the lack of performance evaluation would make the terminal experience non-stressful within the game itself and have carry-over effects to the computer testing.

The results of Study IV clearly confirmed the predicted overall difference in state anxiety elicited during the game and testing conditions, with the former leading to lower levels of state anxiety. The effectiveness of prior game experience was inferred from comparisons of state anxiety data from Study III and the examiner data from the program validation study. Students who received terminal experience in the form of a game exhibited lower state anxiety scores during the following computer testing than did those students in Study III who completed the revised computer test but with no prior terminal experience. However, the more important comparison with the examiner data revealed that anxiety levels were still marginally higher with the computer testing procedures. Equivalent results were obtained from similar comparisons of the attitude data.

These findings more firmly entrench the importance of prior terminal experience for anxiety-reduction within computer testing, a suggestive hypothesis of Studies II and III. However, the effectiveness of the prior experience must be tempered by the fact that the state anxiety levels within the computer test were still high in comparison to the examiner data.

Previous research in the areas of transfer of training, simulation, and modeling has suggested that positive transfer effects of practice are partially dependent upon the similarity of the important task conditions. (Sarason, et al., 1968). In viewing these findings within the present context, it became apparent that the terminal experience provided by the computer game was minimal in terms of the typewriter interactive requirements (typing practice) of the computer test. In the game situation, students received only minimal exposure to typing and response-entering practice while the intelligence test required much more interaction with the terminal. However, as the reader may note, even this minimal experience appeared to have anxiety-reduction benefits within the computer tests.

Study V was concerned with evaluating the hypothesis that prior terminal experience which more closely matched the interactive terminal requirements of the computer test would be even more beneficial for anxiety-reduction purposes. Results of this experiment confirmed this expectation. Students with prior terminal experience in the form of a CAI instructional program exhibited lower state anxiety scores in comparison to students with no prior experience (Study III).

In addition, these results suggested that the type of terminal experience may be important for anxiety-reduction effectiveness. Experience in the form of a computer game led to anxiety results similar to the examiner results, although slightly more elevated. However, the CAI instructional program (constructed response) seemed to produce even lower state anxiety scores within the following computer test in comparison to the examiner testing results. Attitude score comparisons revealed similar findings. These data further suggest that the effectiveness of prior experience would partly be a function of the degree of task similarity between the prior terminal procedures and the nature of the procedures required in the testing situation.

Furthermore, the results of Study V clearly indicated the differential impact of stress on state anxiety in the CAI instructional program and testing situation. Clearly, state anxiety scores declined in the learning situation of easy materials and increased in the testing situation, a finding supporting predictions from Trait-State Anxiety Theory (Spielberger, 1972). Of more practical interest, the study indicated that terminal experience prior to computer-based testing can lead to lower anxiety levels within the latter period. Thus, these findings which replicated and extended those of Study III demonstrated the efficacy of prior terminal experience within automated intelligence testing.

In line with the results of Studies IV and V suggesting the effectiveness of prior terminal experience for anxiety-reduction within computer-based psychological testing, Study VI was directed toward evaluating the effectiveness of prior experience in the form of computer-managed instructional testing on later state anxiety within computer-based intelligence testing. An additional aim within the final investigation was the determination via student interviews of specific cues within both testing environments perceived to be causative in the elicitation of test anxiety.

The results of Study VI were found to replicate those of Studies IV and V in that state anxiety scores were found to increase over the computer-administered intelligence test (pretest X = 7.94; posttest X = 10.95) substantiating the theoretical prediction of the impact of stress on state anxiety (Spielberger, 1972). Of further interest within this study were the positive benefits in terms of anxiety reduction as a result of previous computer-managed testing experience in the form of course module examinations. Indirect evidence concerning this can be inferred by comparing the data from Study VI with the prior data. Table 42 presents such a comparison.

As the reader may note, these data would appear to suggest that the type of prior terminal experience is related to the stress levels experienced during the computer-based intelligence test administration. This is especially true for the pretest state anxiety mean scores. As the type of terminal experience more closely simulates the interactive components of the actual testing situation, there appears to be a corresponding decrease in state anxiety. The same holds true for the posttest state anxiety scores with the exception of the final study in which the posttest state anxiety scores approximated those of Study III (no prior experience) although they were marginally lower.

Additional interest in Study VI was the determination via student interviews of the nature of the specific cues within both the CMI testing and computer intelligence testing that elicit test anxiety. Among the students who reported feeling anxious during the CMI tests, several noted that noise in the testing room was distracting, especially when test items were difficult.

Table 42

Comparison of State Anxiety Measures as a Function of Prior Terminal Experiences

Prior Experience	Study	N	State Anxiety	
			Pretest	Posttest
None	III	24	11.00	11.10
Game	IV	60	9.63	10.40
CAI Instruction	V	30	8.60	9.80
CMI Testing	VI	42	7.95	10.74

They also reported that the vibration or flickering of the letters on the cathode ray tube screen made reading sometimes difficult or upsetting. Most interesting, however, was their interpretation of randomly varying intervals between student response and presentation of the next test item as indicative of the correctness of their response. In reality, the varying systems representation was due to time-sharing considerations of the processing algorithms of the IBM 1500 computer, not to the correctness of a respondent's answer at an individual computer terminal. This interpretation of a neutral stimulus as an evaluative cue, however, is consistent with Sarason, Pederson, and Nyman's (1968) description of a high test-anxious individual, who is particularly vigilant for any cue that might be interpreted as evaluative. In this case the differential systems response time was interpreted as evaluative in nature, although it had nothing to do with the correctness of the responses.

Similar considerations were found to be existent within the computer administration of the Slosson Intelligence Test. Students reported that the feedback statements instructing them to further describe their initial answer to a question and to answer within the allowable time limits per question were somewhat stressful in nature. Although these procedures occur within typical examiner administrations, the same procedures were seen as anxiety-eliciting within a computer context. It would appear that the wording of these feedback statements should be rewritten to reflect more neutral requests. However, it could be that regardless of the manner in which these statements are written, certain high-anxious students will interpret these cues as being evaluative. Of course, this finding and interpretation is subject to further experimentation.

The major limitation of the findings and conclusions of Studies II through VI concerns the experimental sample differences which may have influenced the outcomes in lieu of the discussed variables of terminal familiarity and testing procedures. Table 43 presents a comparison of individual sample data across the validation study and Studies II to VI.

One individual difference variable that could have determined the experimental outcomes in lieu of the particular treatments is anxiety-proneness of the students as measured by the STAI Trait Anxiety scale. In other words, overall decreases in the mean trait anxiety levels over the five studies could have caused the observed decreases in observed state anxiety scores. Inspection of trait anxiety levels negates this alternative hypothesis by considering that the mean trait anxiety levels tended to increase over the experimentation. This fact would make it difficult to show state anxiety decreases as merely a function of trait anxiety, at least from the theoretical position of Trait-State Anxiety Theory (Spielberger, 1972). Thus, it would appear that the experimental treatments were indeed effective for anxiety-reduction purposes. The Test

Table 43

Comparison of Experimental Samples According to Trait Anxiety, IQ, Test Anxiety Scale, and Sex of Student

Study	STAI Mean Trait Anxiety	Test Anxiety Scale	IQ	N	Sex
Hedl, 1971*	34.63	-	127.1	16	8 Males 8 Females
Study II	37.13	-	118.2	30	Females
Study III	36.75	-	119.5	24	Females
Study IV	38.13	14.6	118.5	60	25 Males 35 Females
Study V	40.27	14.1	121.9	30	10 Males 20 Females
Study VI	-	18.8	120.5	42	15 Males 27 Females

*Denotes students with the Slosson Intelligence Test administered by examiners.

Anxiety Scale mean scores would also tend to support this interpretation. Although Test Anxiety Scale data are not available for Studies II, III, and the Hedl (1971) study, the available Test Anxiety Scale data on Studies IV, V, and VI show a similar tendency. Test Anxiety Scale scores were roughly equivalent for Studies IV and V with a marked rise in Study VI. This latter fact makes the findings of Study VI more credible considering the increased test-anxiety proneness of the sample in this study.

Another alternative explanation that might have given rise to the observed data concerned the general ability levels of the experimental groups. If the ability level was found to increase over the reported studies, then the observed decreases in state anxiety levels could have been a function of the task becoming less difficult, and therefore, less stressful in nature. Inspection of the IQ data in Table also negates the forcefulness of this argument. As the reader may note, with the exception of the Hedl (1971) study, the mean IQ scores for the five studies were almost identical. Considering the fact the previous data on the computer-based Slosson Intelligence Test indicated that the administration via computer was approximately five points more difficult than the examiner tests, the observed IQ scores would then tend to more closely match the IQ scores obtained with examiners. The minor IQ differences between experimental samples then fall within the standard error of measurement of the test itself. It should be pointed out that this finding has precedent in other automated testing research. For example, Elwood (1969) found that Wechsler Adult Intelligence Scale scores were lower with the automated presentation method in comparison to scores obtained from examiner administrations of the Wechsler Adult Intelligence Scale. Thus, the experimental groups in the present research appear equivalent in terms of the ability criterion.

A third major alternate explanation could center on the differences in sex of the students. However, even though there were differences from study to study in the male-female proportion of students, correlations of sex with state anxiety were nonsignificant (near zero) for all of the reported studies. Therefore, the results of these studies cannot be explained on the basis of sample differences in sex.

All of these considerations support the major conclusions from this research concerning the impact of computer testing procedures on state anxiety and the importance of terminal familiarity for anxiety reduction. These findings were consistent across the majority of the studies providing replication data for the stability of the important variable effects. In addition, the findings appeared to be consistent across different academic institutions, negating this as a plausible alternative explanation. However, it must be noted that the comparison was post hoc.

The results of these studies also reveal the complex interactive nature between student and computer testing. The issue of the affective causes and implications for computer testing remain unresolved from the present research, with the research findings alluding to the difficulty in formulating precise statements concerning the effectiveness and/or benefits of computers in psychological evaluation, along an affective dimension.

In order to more adequately investigate the importance and benefits of terminal familiarity on anxiety reduction, studies need to be designed to evaluate the anxiety variable as a function of anxiety proneness of the student and the type and amount of terminal exposure on subsequent affective indices within computer testing. An additional independent variable worthy of further investigation would be the type of psychological test. In this manner, certain of the larger questions raised from the present research may be resolved. Only in this fashion can more precise comparisons between computer and examiner testing systems become meaningful.

REFERENCES

Berlyne, D. E. Conflict, arousal and curiosity. New York: McGraw-Hill, 1960.

Berlyne, D. E. Arousal and reinforcement. Lecture presented at Florida State University, February, 1971.

Campeau, P. L. Test anxiety and feedback in programmed instruction. Journal of Educational Psychology, 1968, 59, 159-163.

Day, H. I. Anxiety, curiosity and arousal. OPA Quarterly, 1967, 20, 11-17.

Day, H. I. The measurement of specific curiosity. Paper read in the symposium on Intrinsic Motivation in Education. June, 1969.

Elwood, D. L. Automation of psychological testing. American Psychologist, 1969, 24(3), 287-289.

Fishman, E. J., Keller, L., & Atkinson, R. C. Massed versus distributed practice in computerized spelling drills. Journal of Educational Psychology, 1968, 59, 290-296.

Hansen, D. N., & O'Neil, H. F., Jr. Empirical investigations versus anecdotal observations concerning anxiety in computer-assisted instruction. Journal of School Psychology, 1970, 8, 315 316.

Hedl, J. J., Jr. An evaluation of a computer-based intelligence test. Technical Report No. 21. Tallahassee, Florida: Computer-Assisted Instruction Center, Florida State University, 1971.

Hedl, J. J., Jr., O'Neil, H. F., Jr., & Hansen, D. N. Affective reactions toward computer based intelligence testing. Journal of Consulting and Clinical Psychology, in press.

International Business Machines Corporation. 1401, 1440 or 1460 operating system: Computer-assisted instruction. Endscott, New York: IBM, 1965.

International Business Machines Corporation. IBM Coursewriter II author's guide, Part I: Course planning. New York: IBM, 1967.

Leherissey, B. L. The effects of stimulating state epistemic curiosity on state anxiety and performance in a complex computer-assisted learning task. Unpublished doctoral dissertation, The Florida State University, 1971.

Leherissey, B., O'Neil, H. F., Jr., & Hansen, D. N. Effects of memory support upon anxiety and performance in computer-assisted learning. Journal of Educational Psychology, 1971, 62, 413-420. (a)

Leherissey, B. L., O'Neil, H. F., Jr., & Hansen, D. N. Effect of anxiety, response mode, and subject matter familiarity on achievement in computer-assisted learning. Technical Memo No. 41. Tallahassee, Florida: Computer-Assisted Instruction Center, The Florida State University, 1971. (b)

Leherissey, B. L., O'Neil, H. F., Jr., Heinrich, D. L., & Hansen, D. N. Effect of anxiety, response mode, subject matter familiarity, and learning time on achievement in computer-assisted learning. Technical Memo No. 42. Tallahassee, Florida: Computer-Assisted Instruction Center, The Florida State University, 1971.

Lester, D. The effect of fear and anxiety on exploration and curiosity: Toward a theory of exploration. Journal of General Psychology, 1968, 79, 105-120.

Liebert, R., & Morris, L. Cognitive and emotional components of test anxiety: A distinction and some initial data. Psychological Reports, 1967, 20 975-978.

Mathis, A., Smith, T. A., & Hansen, D. N. College students' attitudes toward computer-assisted instruction. Journal of Educational Psychology, 1970, 61(1), 46-51.

O'Neil, H. F., Jr. Effects of stress on state anxiety and performance in computer-assisted learning. Journal of Educational Psychology 1972, 63(5), 473-481.

O'Neil, H. F., Jr., Hansen, D. N., & Spielberger, C. D. Errors and latency of response in computer-assisted learning as a function of anxiety and task difficulty. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, February, 1969.

O'Neil, H. F., Jr., Spielberger, C. D., & Hansen, D. N. Effects of state anxiety and task difficulty on computer-assisted learning. Journal of Educational Psychology, 1969, 60(5), 343-350.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. The measurement of meaning. Urbana, Illinois: University of Illinois Press, 1957.

Paulson, F. L. The effects of memory support on the concept learning of test anxious children. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, February, 1969.

Penney, R. K. Reactive curiosity and manifest anxiety in children. Child Development, 1965, 36, 697-702.

Sarason, I. G. Effects of verbal learning of anxiety, 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, 403-415.

Sarason, I. G., Kestenbaum, J. M., & Smith, D. H. Test anxiety and the effects of being interviewed. Journal of Personality, 1972, 40, 242-249.

Sarason, I. G., Pederson, A. M., & Nyman, B. Test anxiety and the observation of models. Journal of Personality, 1968, 36, 493-511.

Sieber, J. E., & Kameya, L. I. The relation between anxiety and children's need for memory support in problem solving. Revised Research Memorandum No. 11. Stanford: Stanford Center for Research and Development in Teaching, Stanford University, 1967.

Sieber, J. E., Kameya, L. I., & Paulson, F. L. Effects of memory support on the problem solving abilities of test-anxious children. Journal of Educational Psychology, 1971, 61(2), 159-168.

Sinclair, K. E. The influence of anxiety on several measures of classroom performance. Australian Journal of Education, 1969, 13, 296-307.

Slosson, R. L. The Slosson Intelligence Test for Children and Adults. East Aurora: Slosson Educational Publication, 1963.

Smith, W. F., & Rockett, F. C. Test performance as a function of anxiety, instructor and instructions. Journal of Educational Research, 1958, 52, 138-141.

Snyder, C. R. Amount of body movement as an indicant of task-irrelevant test anxious responses. Journal of Psychology, 1972, 80, 147-150.

Spielberger, C. D. Theory and research on anxiety. In C. D. Spielberger (Ed.), Anxiety and behavior. New York: Academic Press, 1966.

Spielberger, C. D. Anxiety as an emotional state. In C. D. Spielberger (Ed.), Anxiety: Current trends in theory and research. New York: Academic Press, 1972.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. Manual for the State-Trait Anxiety Inventory. Palo Alto, California: Consulting Psychologist Press, 1970.

Terman, L. M., & Merrill, M. A. Revised Stanford-Binet intelligence scale, 3rd revision. Boston: Houghton Mifflin, 1960.

Tobias, S. Effect of creativity, response mode, and subject matter familiarity on achievement in programmed instruction. Journal of Educational Psychology, 1969, 60, 453-460.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

Zuckerman, M., & Lubin, B. Manual for the Multiple Affect Adject Check List. San Diego, California: Educational and Industrial Testing Service, 1965.