Spring 2010

Test Anxiety, Working Memory and Verbal SAT Performance

Kareem Al-Khalil

Test Anxiety, Working Memory and Verbal SAT Performance:

by

Kareem Al-Khalil

(Under the direction of Kent Bodily PhD)

ABSTRACT

The present study examined the relationship between test anxiety, working memory, and verbal SAT performance. Test anxiety negatively affects performance because it decreases working memory space available for processing test information. However, the present study wanted to asses this theory with an experimental design to infer causation about test anxiety. Therefore, participants with high and low trait test anxiety were randomly assigned to experimental and control groups. Participants in the experimental group completed the assessments under conditions of elevated stress. However, these conditions had no effect on their performance. The only observed effect was that participants with high test trait anxiety scored lower on the verbal SAT than their lower test trait anxiety counterparts. However, no difference in working memory span was observed between participants with high and low test trait anxiety.
The null results between experimental and control groups were attributed to ineffective stress manipulation. That is, the participants in the experimental group did not feel anymore stress than participants in the control group. Future studies will have to utilize methods directed towards increasing effect size, such as recruiting a larger sample size and using more extreme scores as cutoffs.

INDEX WORDS: Test Anxiety, Working Memory, Reading Span, Verbal SAT, Reading Comprehension, Sentence Completion

Test Anxiety, Working Memory and Verbal SAT Performance:

by

Kareem Al-Khalil

B.S., Emory University, 2005

A Thesis Submitted to the Graduate Faculty of Georgia Southern University in Partial Fulfillment

of the Requirements for the Degree

MASTER OF SCIENCE

STATESBORO, GEORGIA

2010

TEST ANXIETY, WORKING MEMORY AND VERBAL SAT PERFORMANCE

by

KAREEM AL-KHALIL

Major Professor:	Kent Bodily
Committee:	Janie Wilson
	Larry Locker

Electronic Thesis Approved:

May 2010

Table of Contents

List of Tables/Figures

Page

CHAPTER 1

Introduction

Text anxiety is prevalent among students of all academic levels. It can make school very unpleasant and intimidating to the self esteem. Students with severe test anxiety typically display a lack of self-efficacy and motivation in the classroom (Bembenutty, 2008). Test anxiety may also cause students to avoid studying which results in poor test scores.

Additionally, test anxiety may be more prevalent among students taking standardized tests. Standardized tests are widely used to evaluate students from different schools statewide or nationwide. Currently, most colleges evaluate scores on the Scholastic Aptitude Test (SAT) during the admissions process. SAT scores are reliable predictors of academic success in college. In addition to their implications on academic goals, scores on standardized tests, such as the SAT, are regarded by some as reputable measures of intelligence. All things considered, such attributes provide additional pressure for good performance.

Although SAT scores do provide at least some indication of intellect, they are also influenced by test anxiety. In fact, test anxiety may be a bigger problem on the SATs than in regular classroom tests because of the additional pressure from the academic and intellectual implications of this test. Although research on math anxiety is prevalent, research on verbal anxiety is quite sparse. Therefore, the purpose of the present study was to assess the test anxiety theory in terms of performance on the verbal section of the SAT, which is comprised of reading comprehension and sentence completion questions.

Mandler and Sarason (1952) was among the earliest studies to propose and evaluate the test anxiety theory. At that time, the theory held that test anxiety may arouse task relevant or task irrelevant responses. Task-irrelevant responses include feelings of inadequacy,

helplessness, or physiological arousal. On the other hand, task relevant responses are oriented toward task completion. Individuals with high anxiety exhibit a high amount of task-irrelevant responses. Conversely, individuals with low anxiety exhibit only task-relevant responses during tests.

Mandler and Sarason (1952), along with Sarason, Mandler, and Craighill (1952), empirically supported this theory by comparing responses on an anxiety questionnaire with performance on intelligence tests. Mandler and Sarason (1952) used the Kohs block design and the digit symbol subtests from the Wechsler-Bellevue Intelligence Tests whereas Sarason et al. (1952) used the stylus maze test. Kohs block design is a spatial reasoning task in which participants rotate and arrange colored blocks together to reproduce a picture (Gregory, 2007). The digit symbol test requires participants to associate each digit, 0-9, with a specific symbol, and then draw the appropriate symbol for a series of randomly presented digits. Subjects in both studies were instructed that they were about to take intelligence tests and performance was emphasized. They were falsely informed that their scores would be compared to college students for interpretation purposes. Such distressful instructions are known as ego-involving instructions, because they challenge the test taker's self efficacy (Sarason, 1952). The anxiety questionnaire evaluated the participants' subjective test experiences, such as worry, as well as non-subjective experiences, such as perspiration, or accelerated heart beat. To evaluate the differential responses of high and low test anxious participants, two groups were defined in terms of anxiety level. In Mandler and Sarason (1952), the top 20% of scores comprised the high anxiety group and the bottom 20% of scores comprised the low anxiety group. Sarason et al. (1952) also dichotomized extreme anxiety scores, but the 15th percentile was the cutoff criteria. The high anxiety group performed significantly worse than the low anxiety group on the Kohs

block design test (Mandler and Sarason, 1952). They also took significantly longer to complete the task. Although the results were not significant, similar trends were observed for the digit symbol task.

However, there was a noteworthy difference between the designs of these two studies. Sarason et al. (1952) used a 2 x 2 design with anxiety scores and induced stress (from the ego-involving instructions) as predictors of test performance (mean error scores and time). While Mandler and Sarason (1952) used ego-involving instructions for all participants, Sarason et al. (1952) manipulated stressful conditions. Thus, only half of the participants received the ego-involving instructions (Sarason et al., 1952). The other half received non-ego-involving instructions which did not state that the evaluation was an intelligence test and undermined the importance of performing well. The purpose of this factorial design was to determine if the negative effect of test anxiety on performance would be observed in non-ego involving evaluations. High test anxious participants performed more poorly than low test anxious participants in ego-involving condition. However, in the non-ego involving condition, no difference resulted between the scores of the two groups. This experiment showed that test anxiety theory only holds in tests which evaluative stress is replicated. In other words, anxiety only affects performance when test takers realize the implications of their scores, as is normally the case with exams in the classroom.

The processing efficiency theory (Eysenck and Calvo, 1992) considers the dynamics of task-irrelevant responses and task-relevant responses (Mandler and Sarason, 1952) in relation to working memory; the processing efficiency theory adds a cognitive factor to the previous test anxiety theory. Test takers with high test anxiety worry about their performance and/or the aversive consequences of poor performance on their self efficacy and social image during

examinations. These worrisome thoughts or feelings occupy mental resources in the working memory system, the portion of human cognitive capacity concerned with information processing and temporary storage. Thus, test takers with worrisome thoughts have reduced storage and processing capacity relative to test takers without worrisome thoughts. However, the processing-efficiency theory also states that anxious test takers will use their anticipation of aversive outcomes as motivation to improve their level of performance. Motivated test takers with high test anxiety may increase effort and time devoted to a reading task to obtain similar levels of comprehension to their low anxious counterparts. Thus, test anxiety does not necessarily deteriorate effectiveness, but may deteriorate efficiency.

Calvo and Carrerias (1993) and Calvo, Eysenck, Ramos, Jimenez (1994) empirically demonstrated that high test anxious individuals tend to devote more effort and time to the task. Subjects in both studies used a moving frame method to measure reading times of given passages. In this procedure, subjects move a cursor across a computer screen to reveal one word of the passage at a time (once a new word in the passage was revealed, the previous word was replaced with dashes). Groups of high anxiety participants and low anxiety participants were evaluated with the moving frame method on a computer screen. Although no difference in reading comprehension resulted between high and low test anxious participants, high test anxious individuals took longer to read the passage than their low anxious counterparts. Calvo et al. (1994) empirically attributed the difference in reading speed between the high and low test anxious participants to the increased number of regressions, rereading previous parts of the passage, made by high test anxious students. Regressions were detectable when the participants moved the cursor back to reveal previous words of the passage (this feature was unavailable with Calvo and Carrerias, 1993). Test anxious students perform regressions to preserve performance,

or accuracy, at the expense of reduced processing efficiency. Of course, the use of regressions requires more time and thus, anxious test takers read more slowly, or take more time to read the passage. However, under timed conditions the use of regressions is restricted. Therefore, the motivational aspect of the process efficiency theory will not bolster the test scores of individuals with high test anxiety. Subsequently, high test anxious individuals will score lower than low test anxious on verbal reasoning tasks.

Evaluative stress was employed in these studies by recording the subjects during the exam. Subjects were also provided ego-involving instructions. They were informed that their reading comprehension would be evaluated and that reading comprehension performance was a reliable measure of intellect. Thus, subjects were asked to obtain a high level of comprehension while reading quickly. Lastly, subjects wrote their names on separate pieces of paper so that they could be misled to believe that their scores would be compared to other students. Although Calvo et al. (1994) manipulated evaluative stress, Calvo and Carrerias (1993) placed all their participants under evaluative stress. Under the non-evaluative conditions, no camera was present and the subjects were not informed that reading comprehension is an indicator of intellect. Subjects were asked to obtain an adequate level of reading comprehension, and they were told to read at their own pace. In the evaluative conditions, subjects with high test anxiety made many more regressions than those with low test anxiety, and spent a greater amount of total time performing regressions. The high anxiety group also required more time to read the passages. In the non-evaluative conditions, subjects with high test anxiety also performed more regressions than their low anxiety counterparts, but the difference was much smaller than the evaluative conditions. There was no difference between the two groups in the non-evaluative condition in total regression time or in reading speed.

High test anxious participants, in both studies, were free to reread the passage or parts of the passage to perform well on the reading comprehension evaluations. This was possible because the experiment did not include a time limit. Another noteworthy reservation regarding the methodology of these studies is that they did not examine reading comprehension across passages of various difficulty levels.

Waid, Kanoy, Blick, and Walker (1978) assessed the effects of trait anxiety, levels of disposition for anxiety, and state anxiety, current levels of anxiety, on reading comprehension and reading speed. They assessed these effects on more difficult reading passages. Trait and state anxiety levels were assessed using the State Trait Anxiety Inventory (Spielberger, 1977). Passages were presented to participants at a rate of 200 words per minute. Each participant was then required to answer a multiple choice test based on reading comprehension on the passage.

Waid et al. (1978) did not find that test anxiety significantly affected performance on a reading comprehension exam during in their first experiment. This was attributed to a lack of sufficient stress present in the first study. However, the second study used more difficult passages and ego-involving instructions to make the testing situation more stressful. The instructions labeled the reading comprehension exam as an indication of participants' "ability to do college work". The second study did not manipulate passage difficulty. Participants with high trait anxiety obtained poorer comprehension scores. Thus, the difference in performance between high and low test anxious participants will not be evident unless the exam is sufficiently stressful.

Calvo et al. (1994) assessed whether high and low test anxious students differ in working memory spans, as proposed by the processing efficiency theory of test anxiety. Participants were administered a reading span measure (Daneman and Carpenter, 1980) that involved

administration of increasingly longer sets of sentences at a rate of 400ms per word. After reading each set of sentences, participants had to recall the final word of each sentence in the order the sentences appeared. Reading span was defined as to the maximum number of sentences in which the task can be completed. The reading span measure was only administered in non-evaluative situations, however. Therefore, it should be no surprise that reading span did not vary between low and high anxious individuals, given that previous findings demonstrated that stressful, evaluative situations must be present for anxiety to have a sufficient effect on reading comprehension (Waid et al., 1978). Given that passages were displayed at a fixed rate 200 words/minute, regressions were restricted as would be the case with timed reading comprehension tests. Therefore, test takers with low trait anxiety scored better than those with high trait anxiety. To confirm the processing-efficiency theory of test anxiety, tests of working memory space must also be conducted under evaluative situations to obtain differential working memory spans between high and low anxious test takers. When stressful conditions were implemented differential reading spans were obtained (Calvo, Ramos, & Estevez, 1992).

Test anxiety is now defined as the tendency to be overly and anxiously preoccupied with test performance, rather than focusing one's attention on the test (Bembenutty, 2008). During academic exams, test anxiety hinders a test taker's information processing by interfering with his attention, encoding, and retrieval processes. Test anxious students tend to allow distracting thoughts to interfere with concentration during test situations, and sometimes these thoughts include ways of avoiding the test. Test anxious students may also generate negative thoughts during tests. These thoughts may be self-condescending with regards to one's intelligence. Additionally, anxious test takers may fearfully anticipate social situations in which the subject of

test scores may arise. Test anxiety is not only manifested cognitively, but socially, emotionally, and physiologically as well (Lowe 2008).

Over 50 years of test anxiety research cohesively suggests that under stressful, timed conditions, students with high test anxiety will score lower on tests than their low anxiety counterparts. This is due to reduced working memory space available for individuals with high test anxiety. The working memory hypothesis regarding test anxiety suggests that individuals with high test anxiety have fewer working memory, or attention, resources available during testing situations than individuals with lower test anxiety, as a portion of their working memory space is consumed by worrisome thoughts (Lee, 1995). Sarason et al. (1952) originally demonstrated, however, that high test anxious individuals will not perform more poorly than low test anxious individuals unless they feel ego-involved. However, Mandler and Sarason's (1952) account of test anxiety was learning-based and lacked cognitive factors. Calvo et al. (1994) contributed empirical evidence in support of the cognitive dynamics of test anxiety and working memory. Calvo et al. (1994) also demonstrated the time limit constraints of test anxiety. That is, anxiety will only have an influence on performance when tests are timed and regressions cannot be used.

Spielberger (1980) developed the Test Anxiety Inventory (TAI) to serve as a measure of predisposition toward exhibiting anxiety symptoms during tests. The TAI consists of items assessing how often participants exhibit the physical, emotional, and cognitive manifestations of test anxiety. The physical symptoms of test anxiety include sweaty palms, racing pulse, trembling limbs, jittery feeling, and dizziness (Lowe 2008). The cognitive manifestations refer to the interference or obstruction of attention to the actual test questions. As with the previously discussed methods, the cognitive manifestations are reflected by worrying about academic

performance or self-condescending consequences (Spielberger, 1980). The TAI has test-retest reliability coefficients for high school, college, and graduate students of about .80, .81, and .82 when the pretest and post tests are administered within one month. TAI was also found to have strong correlations with other measures of anxiety. The TAI has a correlation of .82 with Sarason's (1978) Test Anxiety Scale (TAS), and correlations of .73 and .77 with the worry scale and emotionality scales, respectively, of Liebert and Morris's (1967) Worry and Emotionality Questionnaire (WEQ).

Daneman and Carpenter's (1980) reading span task was constructed such that the task-demands would be compatible with working memory characteristics. The concept of working memory developed as a more active form of short term memory. Working memory is assumed to involve processing as well as storage functions (Baddeley, 1992). Working memory is capable of parallel processing tasks, such as temporarily storing information while making inferences from that information. Such parallel processing tasks are demanded by reading comprehension tasks on scholastic tests. Measures of short term memory, such as the word span test, alternatively do not significantly predict reading comprehension scores (Daneman and Carpenter, 1980). Short term memory measures are not taxing enough because no information processing, only information storage, is required. Thus, a measure of working memory, such as the reading span, may be a more accurate predictor of reading comprehension score. The reading span and the test anxiety inventory (TAI; Spielberger, 1980) likely capture the relationship between anxiety and verbal reasoning.

The literature examining the relationship between test anxiety and reading skills is still lacking in some ways. Only Waid et al. (1978) assessed reading comprehension with a time limit. Although anxiety has been induced with ego-involving instructions, in most studies all

participants were administered these instructions. Therefore ego-involving anxiety has been rarely manipulated in research that assessed its effects on reading comprehension. Furthermore, although Waid et al. (1978) used difficult passages, passage difficulty was not manipulated, nor was it manipulated in any other study.

The purpose of this study was to examine the effects of test anxiety on verbal working memory and verbal test performance in an experimental design. An evaluative stress manipulation was included such that half of the participants completed the assessments under the administration of ego-involving instructions. The effects of test anxiety were evaluated using a verbal section of the SAT, a timed standardized test in which question have the competing response attribute. In other words, this test is difficult because many questions have incorrect answer choices that closely resemble the correct choices. In accordance with the working memory hypothesis (anxiety restricts working memory space), individuals with high trait anxiety were predicted to perform worse on the reading span test and the verbal SAT tests than individuals with low trait anxiety. However, this prediction should only be observed under the ego-involving conditions (Sarason et al., 1952; Waid et al., 1978; Calvo et al., 1994). Additionally, participants with low test anxiety taking the assessments under ego-involving conditions should perform better than participants under non ego-involving conditions, as demonstrated by Mandler and Sarason (1952).

Research assessing the relationship between test anxiety and other forms of verbal reasoning, such as sentence completion, which tests the student's vocabulary proficiency and verbal reasoning skills, is scare. Like reading comprehension, fill-in-blank/sentence completion questions are very prevalent in the classroom. Sentence completion questions are also included in the verbal SAT along with reading comprehension questions. Therefore, it would greatly behoove college-bound students to understand the effects of test anxiety on accuracy for both

types of questions in a time-constrained situation. An adjunct purpose was to compare the effects of test anxiety on the accuracy of both types of questions, sentence completion and reading comprehension. All subjects will receive composite SAT scores, sentence completion scores, and reading comprehension scores. In accordance with the processing-efficiency theory of test anxiety, it was predicted that subjects with high test anxiety would obtain lower scores on both question types than subjects with low test anxiety. It was also predicted that higher anxiety levels, influenced by test trait anxiety and the stress manipulation, would be more detrimental to accuracy on reading comprehension questions than sentence completion questions. The reading comprehension passages contain much more information that must be stored in working memory than the individual sentences in the sentence completion section. Additionally, test takers have to remember the information discussed in the passage while answering questions regarding the content of the passage and/or make appropriate inferences from the passage. Since reading comprehension requires the parallel processing of storage and reasoning tasks, reading comprehension questions require more working memory capacity than sentence completion questions. Furthermore, scores on reading comprehension should be more affected by the time limit because the use of regressions is restricted for an entire passage, rather than for individual sentences. Lastly, blood pressure and heart rate measures were obtained to verify that elevated anxiety levels were induced by ego-involving instructions.

CHAPTER 2

Method

Participants

Georgia Southern Undergraduates were administered the Test Anxiety Inventory (TAI). Participants were between the ages of 18-39 (m=19.7, SD=12.64, n=193) and were recruited via the Georgia Southern online sona-system. All results were associated with an experimenter ID number randomly assigned to each participant to ensure anonymity. Participants were awarded one research credit for participating in the TAI.

Participants who scored in the upper and lower 25 percentiles were invited to participate in the experiment. This sample consisted of 86 participants with TAI scores above 53 and below 33. Forty-five of the participants had high TAI scores and 41 participants had low TAI scores. Fifty participants completed the study. Approximately half of the high TAI (M=26.8, SD=3.65) and low TAI (M=60.4, SD=5.92) participants were randomly assigned to the control group (High n=11, Low n=11) and the other halves were randomly assigned to the experimental group (High n=13, Low n=15) which received evaluative stress. Participants were awarded an additional research credit for completing the experiment.

Materials

Participants used a 19" Dell flat screen monitor to complete the assessments online at www.surveygizmo.com.

TAI/STAI. The Test Anxiety Inventory (TAI; Spielberger, 1980) is a 20 item survey that measures trait anxiety during tests. This assessment was administrated as the "Test attitude inventory" rather than the "Test anxiety inventory". In fact, the word anxiety was not used throughout administration of the TAI until the debriefing section. The State-Trait Anxiety

Inventory (STAI: Spielberger, 1980) also consists of a 20 item questionnaire that measures state anxiety, or current/instant levels of anxiety. The instructions for both assessments state that participants are to indicate how they generally feel (during tests on the TAI). Both inventories use likert scales. Students are to indicate how often they have certain feelings, bodily symptoms, or thoughts (during exams on the TAI) on a scale from 1 (almost never) to 4 (almost always). An overall TAI/STAI score is calculated from the sum of the individual item scores. The minimum score is a 20 (very low anxiety) and the maximum score is an 80 (very high anxiety). Example items on the TAI are, "I feel anxious during tests," and on the state scale of the STAI are, "I feel calm", or "I feel pleasant".

Scholastic Aptitude Test. Participants completed a verbal section of the Scholastic Aptitude Test. This was a practice test from the Barron's study guide to the SAT. The test was comprised of 24 multiple choice questions: 12 reading comprehension questions and 12 sentence completion questions. For reading comprehension questions, participants read a passage (74 lines, 800 words) and answered questions regarding information presented in the passage. The passage was an excerpt about Charles Darwin's experience on the Galapagos Islands. For sentence completion questions, participants selected the answer choice with the word or words that best completed a given sentence.

Composite SAT scores were calculated by counting the number of questions answered correctly. The maximum possible score was 24 and minimum score was 0. Reading comprehension and sentence completion scores were obtained by counting the number of correct answers in each section. The maximum possible score for each section was 12.

Reading Span Task. Daneman and Carpenter (1980) created a reading span measure to assess the working memory capacity dedicated to verbal reasoning tasks. This reading span has

a correlation with, and is predictive of, Verbal SAT scores. This test was administered via note cards containing one sentence each. Participants initially read five sets of two sentences aloud and then recalled the last word of each sentence in the order they were presented. Participants then attempted to complete this task with longer sets of sentences (three, then four, then five) until they fail at three out of five sets. Participants are instructed to begin reading the sentences aloud as soon as the card is flipped over and each sentence is revealed. This strict procedure limits rehearsal of the last words.

Blood pressure/Heart rate. Three measures of systolic blood pressure, diastolic blood pressure, and heart rate were taken throughout the session. An automatic blood pressure wrist cuff was placed on the participant's left wrist as he was seated.

Evaluative stress. Evaluative stress was manipulated between the experimental and control groups. The experimental group completed the session with a powered on, video camera directed at them. The experimental group also received ego-involving instructions with a fictitious retest policy to further invoke evaluative stress. The ego-involving instructions stated that the SAT is a reputable measure of intellect and scholastic ability. The ego-involving instructions also contained a fictitious retest policy for those who failed to answer 70% of the SAT questions correctly. Conversely, participants in the control group, were not placed in front of the camera, and were instructed to try their best even though their performance was not crucial. Appendix A shows the instructions for the experimental and control group.

Procedure

Most participants (n=130) were provided the link to the online TAI on the sona-systems website. These participants were free to complete the assessment at any place or time before a deadline. Some participants (n=63) completed the TAI during a class session. From this initial

pool of participants, the upper and lower 25th percentile were invited to participate in an experiment. Participants who accepted the invitation were taken to a lab to complete the rest of the study one at a time. Participants were provided written informed consent to settle their blood pressure and heart rate which may have been elevated from walking to the lab. After they signed the informed consent, a baseline blood pressure and heart rate measurement was taken. They were read the instructions, which corresponded to their group assignment (see Appendix A). The experimental group was informed that they would be video taped to further evaluate their reactions to the questions on the STAI, SAT, and the reading span assessments. The camera was then turned towards them and powered on. A second blood pressure and HR reading were taken after the instructions were read. Then, they proceeded to complete the state scale of the STAI online and then clicked on a second link to complete a verbal SAT section online. They were allotted 25 minutes to complete the verbal SAT exam. After the SAT, participants were read the instructions for the reading span test. Participants in the experimental group were told that the reading span test was also a reputable measure of intellect and scholastic ability. Finally, they were debriefed and a final blood pressure/HR measure was taken. The debriefing is stated in Appendix A. The experimental design is illustrated in table 1.

Table 1

Experimental Design

	Evaluative Stress	Non Evaluative Stress
Sentence completion	½ High TAI ½ low TAI	½ High TAI ½ low TAI
Reading Comprehension	½ High TAI ½ low TAI	½ High TAI ½ low TAI

CHAPTER 3

Results

A 2 (High TAI, Low TAI) x 2 (experimental, control) between-subjects MANOVA, $F(3, 44) = .027$, $p = .994$, was conducted. The model did not significantly account for variation in composite SAT scores, $F(3, 49) = 2.48$, $p = .073$, or reading span scores, $F(3, 49) = 1.10$, $p = .358$, but did account for variation in state anxiety, $F(3, 49) = 3.5$, $p = .023$.

Regarding composite SAT scores as the dependent variable, a main effect was observed for TAI scores. Participants with high TAI scores had lower SAT scores, $F(1, 49) = 6.59$, $p = .014$. However, group assignment (experimental/control) did not affect SAT scores, $F(1, 49) = .656$, $p = .422$, nor was there an interaction between group assignment (experimental/control) and test trait anxiety $F(1, 49) = .081$, $p = .778$. Figure 1 displays a graph of the relationship between evaluative stress and TAI scores on composite SAT scores.

Group assignment (experimental/control) had no effect on reading span scores, $F(1, 49) = 1.52$, $p = .224$. Trait anxiety did not affect reading span scores either, $F(1, 49) = 1.68$, $p = .201$. There was no interaction between these two predictors, $F(1, 49) = .012$, $p = .914$. Figure 2 displays a graph of the relationship between evaluative stress and TAI scores on reading span scores.

As expected, high trait anxiety was predictive of higher state anxiety, $F(1, 49) = 10$, $p = .003$. However, group differences (experimental/control) did not lead to differences in state anxiety $F(1, 49) = .463$, $p = .499$. There was also no interaction between TAI and stress group on STAI scores $F(1, 49) = .001$, $p = .974$. If the manipulation had worked, the experimental group would have yielded a significant difference in state anxiety between high and low TAI individuals. The lack of interaction shows that the manipulation was ineffective. Figure 3

displays a graph of the relationship between evaluative stress and TAI scores on state anxiety scores. According to the results of the STAI, participants did not respond to the evaluative stress.

Finally, A 2 (high/low TAI) x 2 (experimental/control) x 2 (question type) mixed analysis of variance was conducted to compare the effects of reading comprehension and sentence completion. Reading comprehension scores were lower across all groups $F(1, 49) = 37.65$, $p < .001$. However, no interactions were present. Performance between question types did not differ between experimental and control groups, $F(1, 17) = .245$, $p = .627$, high TAI and low TAI groups, $F(24, 17) = 1.21$, $p = .351$, nor was there an interaction between these two independent variables, $F(7, 17) = .87$, $p = .549$. Figure 4 displays a graph of the relationship between evaluative stress and TAI scores on reading comprehension and sentence completion scores.

A 2 (high/low TAI) x 2 (evaluative stress/no stress) x 3 (initial/after manipulation/ final) mixed MANOVA was conducted for each of the three vital signs (systolic, diastolic and HR) across both stress conditions and high and low TAI groups. A within subjects variable was added to this analysis based on the time the vital signs were taken: either following the informed consent, the instructions, or the debriefing. Stress had no main effect. Individuals with high trait anxiety should have shown a greater increase in blood pressure and heart rate, from baseline to after the instructions, than those with low trait anxiety. Furthermore, this disparity should be even greater among high and low TAI individuals in the evaluative stress group. However, no interaction was found between systolic blood pressure and evaluative stress, $F(2, 170) = 1.014$, $p = .367$. However, there was an interaction between high and low TAI groups and systolic blood pressure, $F(2, 17) = 4.413$, $p = .015$. The low TAI group exhibited a significant increase in systolic blood pressure from the baseline after the instructions, $t(24) = -2.69$, $p = .013$.

Additionally, systolic blood pressure was different after the instructions to the end, t(24)= 3.95, p=.001, and differed from baseline to the end, t(24)= 2.64, p=.014. Nevertheless, an interaction between both TAI and experimental/control group was not found, F(2, 17) =.152, p=.859. Also, no interactions were found between diastolic blood pressure and evaluative stress, F(2, 170) =.102, p=.903, TAI, F(2, 17) =2.26, p=.111, or both, F(2, 17) =.830, p=.440. Finally, no interactions were found between heart rate and experimental/control group assignment, F(2, 17) =.106, p=.900, high and low TAI groups, F(2, 170) =.052, p=.949, nor was there an interaction between TAI and experimental/control group assignment, F(2, 17) =.693, p=.506.

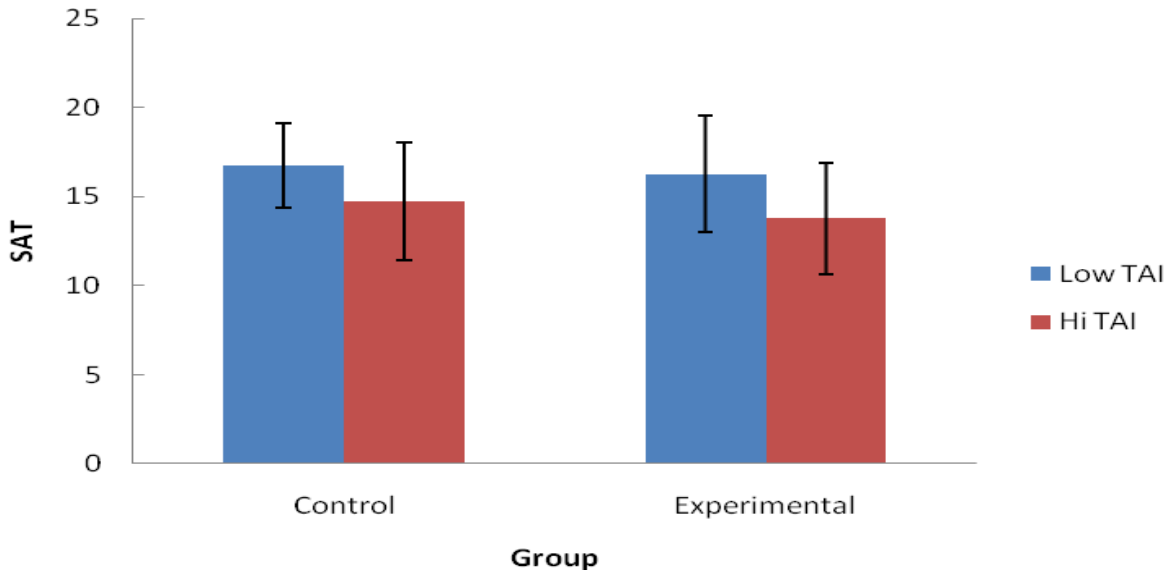

Figure 1. Effects of experimental manipulation and TAI on verbal SAT. This graph illustrates mean SAT scores for each of the four groups (High TAI-Exp, High TAI-Control, Low TAI-Exp, Low TAI-Control). Error bars represent standard deviations.

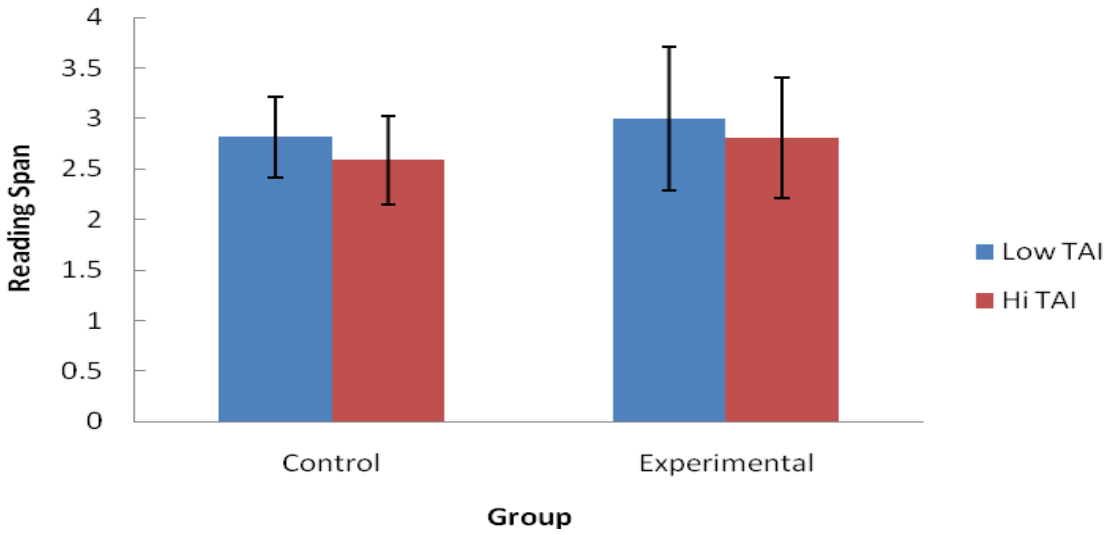

Figure 2. Effects of experimental manipulation and TAI on reading span. This graph illustrates mean reading span scores for each of the four groups (High TAI-Exp, High TAI-Control, Low TAI-Exp, Low TAI-Control). Error bars represent standard deviations.

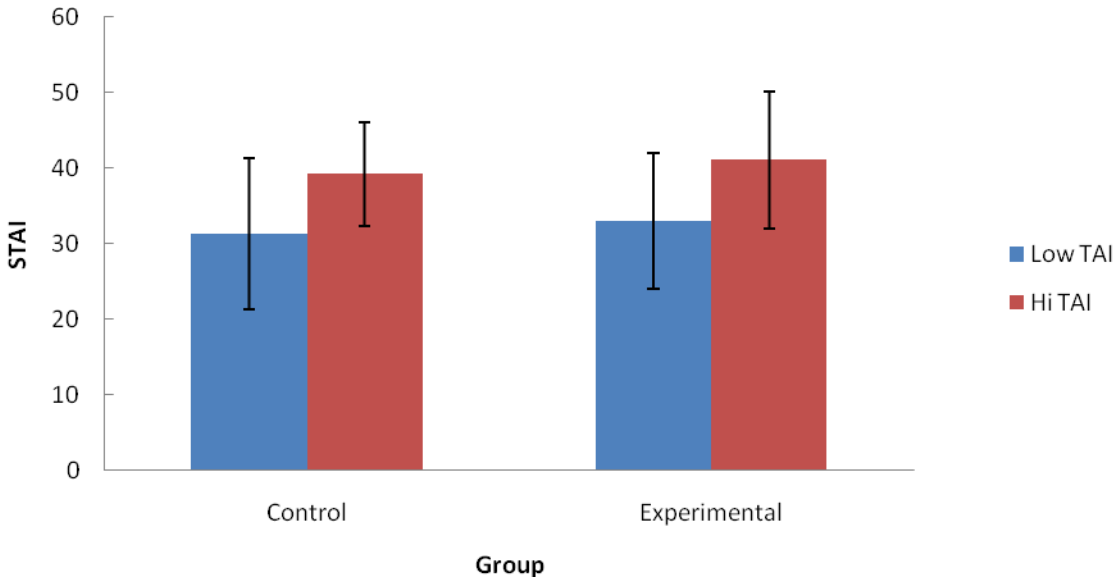

Figure 3. Effects of experimental manipulation and TAI on STAI scores. This graph illustrates mean scores for the state scale of the STAI for each of the four groups (High TAI-Exp, High TAI-Control, Low TAI-Exp, Low TAI-Control). Error bars represent standard deviations.

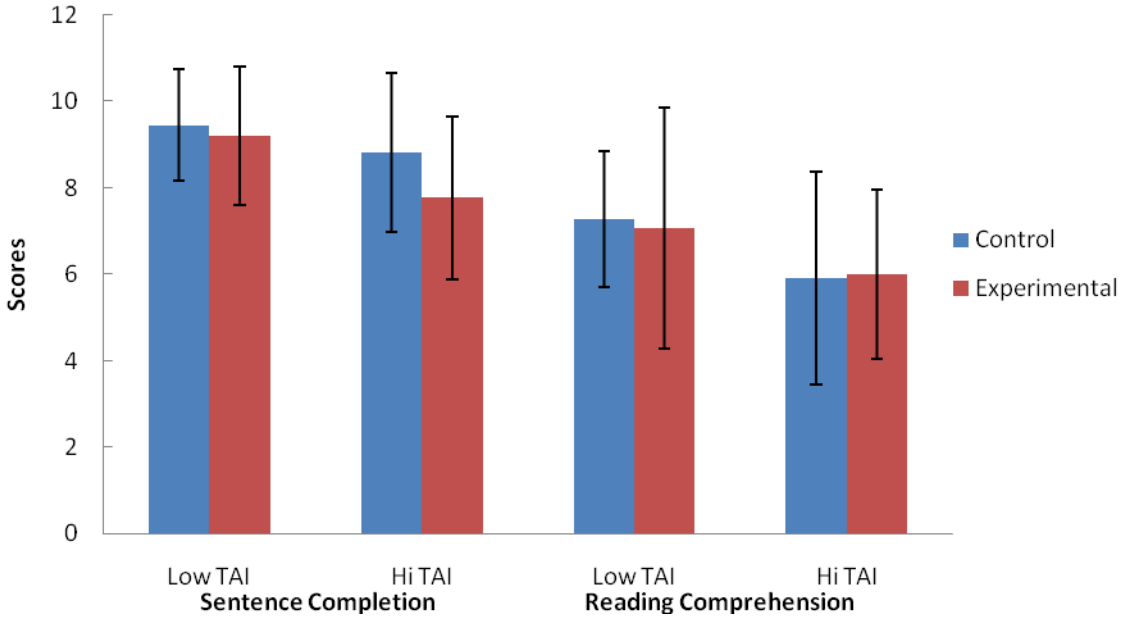

Figure 4. Effects of experimental manipulation and TAI on different question types. This graph illustrates mean SAT scores for reading comprehension and sentence completion questions for each of the four groups (High TAI-Exp, High TAI-Control, Low TAI-Exp, Low TAI-Control). Error bars represent standard deviations.

CHAPTER 4

Discussion

The results revealed that TAI had an effect on SAT, reading span, and STAI scores across experimental and control groups. None of the predicted interaction effects were found. SAT and reading span scores were unaffected by evaluative stress. This is probably because the manipulation (evaluative stress) was not perceived by the participants, which is suggested because STAI scores did not differ among high and low TAI participants in the evaluative stress. No differential effects were found between reading comprehension and sentence completion questions. Overall, these results suggest that although test anxiety is predictive of performance, test anxiety was not successfully manipulated in the present study. Such a manipulation may be more difficult than anticipated because test anxiety is a product of many factors such as previous testing experience, test trait anxiety, and even the participants' current health conditions. For example, participants who are sleep deprived, or take the test hungry, will be more irritable and be more susceptible to uncontrollable test anxiety. It may be difficult to manipulate evaluative stress if it's caused or precipitated by many extraneous factors.

Alternatively, the manipulation may have failed to increase state anxiety due to a ceiling effect. Participants were informed that the experiment would involve completing a verbal SAT section. They probably knew the intellectual implications of the SAT given that they were all college undergraduates who recently took this test. The manipulation, which was administered after the instructions, may not have increased the participants' state anxiety which may have elevated after reading the informed consent.

We drew our predictions from Calvo et al. (1994) who showed that high anxious participants read more slowly than low anxious participants and this disparity was greater under

evaluative stress. The difference in reading speed was attributed to a greater number of regressions performed by high TAI individuals. Under timed conditions, anxious participants would not be able to rely on regressions to achieve the same level of comprehension as non-anxious participants. Thus, anxious participants should score lower on reading comprehension tasks than non-anxious participants. Test trait anxiety and evaluative stress should have predicted lower SAT performance. However, only test trait anxiety did so. These results could be attributed to certain differences between the methodology of Calvo et al. (1994) and our methodology. They used high and low TAI scores and State STAI scores as cutoff criterion for the means end split. In other words, Calvo et al. (1994) selected participants with TAI scores in the top and bottom quartiles only if the high TAI scores corresponded with significantly higher STAI scores than the lower TAI scores. They administered the STAI scale during an inductive reasoning task under test instructions to create a stressful situation. However, our criterion was only based on TAI scores, not STAI scores. Calvo et al. (1994) were more likely to obtain reliable STAI results because their test takers already demonstrated higher state anxiety in response to stress. Additionally, they administered an abridged 10-items STAI scale after participants read the reading comprehension passages, but before they answered the questions. Thus, participants had more time to process the evaluative stress which may have been facilitated by exposure to complicated passages. Such methods would be more likely to result in a significant difference of STAI scores between high and low TAI participants in the evaluative stress group.

Calvo et al. (1994) also demonstrated that differential performances in reading comprehension appear only under conditions of evaluative stress. Furthermore, Calvo et al. (1992) demonstrated that differential performance in reading span only results under evaluative

stress. Thus, under evaluative stress conditions, high TAI participants should score lower on the reading span and reading comprehension than lower TAI participants. However, no differences between reading comprehension and sentence completion questions were observed under evaluative stress or control conditions in the present study.

Further limitations of this study include the small number of participants we obtained, thus limiting our power. Calvo et al. (1994) obtained 18 high TAI participants and 18 low TAI participants in each condition (evaluative stress and no stress), but their cutoffs were higher than the cutoff scores used in the present study.

Future directions include testing a greater number of participants and running methods more in line with previous research (Calvo et al. 1994; Calvo and Carrerias, 1993). The TAI could be administered to a larger pool of participants so that higher cutoff scores could be used as criterion to participate in the experiment. For example, if the TAI was administered to 400 participants, the criterion cutoffs could be set to the top and bottom 10% to obtain 80 participants for the experiment, (20 in each of the four groups). Such methods would give the present study more power by increasing the effect size of test anxiety because there would be more disparity between anxiety scores in the high and low TAI groups. Additionally, STAI measures can be obtained from participants during a brief, stressful evaluation and used as an initial criterion to screen out participants between the extremes. Such methods would increase the power of the manipulation.

References

Baddeley, A. (1992) Working Memory. *Science, 255*, 556-559.

Bembenutty, H. (2008) Self Regulation of Learning and Test Anxiety. Psychology Journal, 5 122-139.

Calvo, M. G., & Carreiras, M. (1993). Selective influence of test anxiety on reading processes. *British Journal of Psychology*, 84, 375-388.

Calvo, M. G., Eysenck, M. W., Ramos, P. M., & Jimenez, A. (1994) Compensatory Reading Strategies in Test Anxiety. *Anxiety, Stress, and Coping*, 7, 99-116.

Calvo, M.G, Ramos, P.M., & Estevez, A. (1992) Test anxiety and comprehension efficiency: the role of prior knowledge and working memory deficits. *Anxiety, Stress, and Coping*, 5, 125-138

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466

Eysenck, M.W & Calvo M.G. (1992) Anxiety and performance: The processing efficiency theory. *Cognition and Emotion*, 6, 409-434

Gregory, R.J. (2007) Individual tests of intelligence and achievement. In Psychological Testing: History, Principles, and Applications. Wheaton, Illinois: Pearson.

Lee, J. (1995). The information processing of test anxiety and its effect on the speed-accuracy tradeoff (the spreading activation explanation of ACT). Dissertation Abstracts International, 56, 2176

Mandler, G., & Sarason, S. B. (1952). A study of anxiety and learning. *Journal of abnormal and social psychology*, 47, 166-173.

Sarason, S. B., Mandler, G., & Craighill, P. G. (1952). The effect of differential instructions on anxiety and learning. *Journal of abnormal and social psychology*, 47, 561-565.

Spielberger, C. D., Gonzales, H. P., Taylor, C. J., Anton, W. D., Algaze, B., Ross, G. K., & Westberry, L. G. (1980). *Test Anxiety Inventory.* Palo Alto, CA: Consulting Psychologists Press

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P.R., & Jacobs, G.A. (1977). State-Trait *Anxiety Inventory for adults.* Palo Alto, CA: Consulting Psychologists Press

Waid, L., Kanoy, R., Blick, K., & Walker, W. (1978). Relationship of state-trait anxiety and type of practice to reading comprehension. *Journal of Psychology: Interdisciplinary and Applied*, *98*, 27-36.

Appendix A (Instructions & Debriefing)

Evaluative Stress Instructions

The purpose of the following study is to assess the influential cognitive factors of critical reading skills. In this study, you will complete a brief questionnaire regarding your attitudes towards the exam followed by one SAT verbal section. You will be allotted 25 minutes to complete the exam which consists of reading comprehension and sentence completion questions. Keep in mind that the SAT is an aptitude test, meaning that it measures reasoning ability. It is a reputable measure of intelligence and scholastic ability. Your performance on the SAT is indicative of your ability to have success in college classes. Thus, it is essential that you put forth your best effort on these exams. To ensure reliable results, you will be given a retest if you fail to answer 70% of the questions correctly.

I will also videotape you to further evaluate your reactions to the questions. (I turn the camera to point to the participant and turn it on).

Control Instructions

The purpose of the following study is to assess the influential cognitive factors of critical reading skills. In this study, you will complete a brief questionnaire regarding your attitudes towards the exam followed by one SAT verbal section. You will be allotted 25 minutes to complete the exam which consists of reading comprehension and sentence completion questions. We are mainly interested in how questions may be presented in a certain way to correspond to human cognition. Although your performance is not a crucial part of this study, please try your best on this test.

Debriefing

The actual purpose of this study was to investigate the effects of test anxiety on test performance. More specifically, my main interest was to assess whether high test anxiety was a reliable predictor of poor performance on a critical reading evaluation. Since most high school students have taken the Scholastic Aptitude Test (SAT), a discussion of the effects of test anxiety on SAT scores would benefit the high school student population. The first questionnaire you completed, approximately 1-2 weeks ago, was actually the Test Anxiety Inventory (TAI), created by Charles Spielberger (1980). The second questionnaire you completed was the state scale of the state trait Anxiety Inventory (STAI), created by Charles Spielberger (1980). To ensure the validity of the study, I did not inform you of the actual purpose. If I did, your knowledge of the purpose beforehand may have influenced your results on the SAT or the TAI.

Test anxiety is detrimental to performance because it constricts working memory space. Therefore, after the SAT, I also administered a reading span (Daneman and Carpenter, 1980) to assess your working memory space.

The effects of test anxiety could be generalized to other standardized tests, especially those with similar formats as the SAT. Additionally, research has shown that test anxiety may have detrimental effects on test scores in general, making the classroom setting very stressful and intimidating. If you feel that you may have high test anxiety, you may see your test results by emailing me your participant ID number. Furthermore, if you feel test anxiety has been detrimental on your academic performance throughout your lifetime, I encourage you to contact the Georgia Southern Counseling Center: 912-478-5541, PO Box 8043, Statesboro, GA 30460-8043.