THE DEVELOPMENT OF THE TURKISH FORM OF THE SPIELBERGER TEST ANXIETY INVENTORY: A STUDY OF TRANSLITERAL EQUIVALENCE AND RELIABILITY

by

Deniz Albayrak Kaymak

B.A. in Pedagogy, İstanbul University, 1982

Submitted to the Institute for Graduate Students in Social Sciences in partial fulfillment of the requirements for the degree of

Master of Arts

in

Educational Sciences

Boğaziçi University

1985

This thesis, submitted by Deniz Albayrak Kaymak to the Faculty of Education, Department of Educational Sciences of Boğaziçi University in partial fulfillment of the requirements of the Degree of Master of Arts is approved.

Thesis Advisor: _____
Necla Öner, Ph.D.

Committee Member: _____
Ali Baykal, Ph.D.

Committee Member: _____
Kadir Özer, Ph.D.

Date: September 20, 1995

ACKNOWLEDGEMENTS

My indebtedness to my thesis advisor, Necla Öner can never be expressed by words. This study would not even exist without her invaluable contributions. I admired her limitless energy, resourcefulness, critical, but supportive approach, and cooperative attitude that she kept constantly throughout the study. I continuously learned and enjoyed a lot to work with her. I am proud to have the chance of being one of her students

I would like to thank to Ali Baykal, and Kadir Özer for the time they spared for the study, and for their encouraging attitude.

I want to thank to Turhan Oğuzkan, and all members of the Department of Educational Sciences not only for their contributions to my graduate education, but also for the interest and supportive behaviours they showed during this study.

If it wasn't for the efforts and assistance of the counselors, teachers, and my friends who involved in the application part of the research, I could never be able to complete the study. I am happy to express my thankfulness to all these people, and the students participated in the study.

My special thanks are due to my friend Melike, whose companionship is valuable for me, and her younger brother Mahmut Çataklı for their cooperation in computer analysing of the data.

I am profoundly grateful to my husband, who has always supported my academic development, and motivated me towards the better, for his understanding and patience throughout the study. Finally, my sincere appreciation is due to my mother, who has also been my good friend, for her precious efforts to provide the environment where I could study at ease.

Deniz Albayrak Kaymak

ABSTRACT

This thesis is based on a study about the Turkish adaptation of the Test Anxiety Inventory (TAI), developed by Spielberger in 1980. It was carried in two phases. First, the transliteral equivalence of the inventory was established; then, the reliability of the Turkish form of the inventory was studied.

In the transliteral equivalence part of the study, the English and the Turkish forms of the inventory were tested on 164 Turkish bilingual university and highschool students. The analyses including analysis of variance. Product Moment, alpha, item-total, and item-remainder correlation techniques showed that there were no significant differences between the scores obtained from the English and the Turkish language forms of the TAI. This finding indicated that the English and the Turkish forms of the TAI are comparable psycholinguistically.

In the reliability part of the study the stability and the internal consistency of the Turkish form of the TAI (TAI-T) were investigated. For this purpose, 1031 Turkish students from different levels of education (university, senior, and junior highschool) and SES in İstanbul were tested.

The stability of the scores of the TAI-T was found by test-retest correlations with varying time intervals. The

obtained high test-retest correlations attested to the stability of the scale.

The internal consistency of the TAI-T was established using alpha, item-total, and item-remainder correlations. High alpha coefficients, and mediocre to high item-total and item-remainder correlations were obtained. These values indicate that the item consistency and the homogeneity of the Turkish form are very satisfactory in general. One particular finding, however, was that with younger (junior highschool) students some test items need further improvement for better internal consistency. It is recommended that this point is taken up in the following studies. Generally, the study demonstrated supporting data on the stability and the internal consistency of the TAI-T.

ÖZET

Bu tez, 1980 yılında Spielberger tarafından geliştirilen Sınav Kaygısı Envanteri'nin Türkçe uyarlaması üzerinedir. Çalışma iki aşamada yürütülmüştür. İlkin, envanterin dilsel eşiti (karşılığı) oluşturulmuş; daha sonra, bu Türkçe formun güvenirliği üzerinde çalışılmıştır.

Çalışmanın dilsel eşitlik kısmında envanterin İngilizce ve Türkçe formları her iki dili bilen 164 üniversite ve lise öğrencileri üzerinde sınanmıştır. Varyans analizi, Product Moment, alfa, item-toplam ve item-bırakma korelasyon tekniklerini kapsayan analizler envanterin İngilizce ve Türkçe formlarından elde edilen puanlar arasında anlamlı farklılıklar olmadığını göstermiştir. Bu bulgu envanterin İngilizce ve Türkçe formlarının psikolinguistik yönden kıyaslanabilir olduğuna işaret etmiştir.

Çalışmanın güvenirlik kısmında, envanterin Türkçe formunun stabilitesi ve iç-tutarlılığı araştırılmıştır. Bu amaçla İstanbul'un farklı eğitim (üniversite, lise ve ortaokul) ve sosyoekonomik düzeylerinden gelen 1031 Türk öğrencisi teste alınmıştır.

Envanterin Türkçe formundan elde edilen puanların stabilitesi değişik zaman aralıkları ile yapılan test-tekrar - test korelasyonları ile bulunmuştur. Elde edilen yüksek test-

tekrar test korelasyonları ölçeğin stabilitesini kanıtlamıştır.

Envanterin Türkçe formunun iç-tutarlığı alfa, item-toplam ve item-bırakma korelasyonlarını kullanarak oluşturulmuştur. Yüksek alfa korelasyonları ve orta ile yüksek arasında değişen item-toplam ve item-bırakma korelasyonları elde edilmiştir. Bu değerler, genel olarak, Türkçe formunun item tutarlığı ve homojenliğinin oldukça doyurucu olduğunu göstermiştir. Ancak, küçük yaştaki (ortaokul) öğrencilerden elde edilen belli bir bulgu bazı test itemlerinin daha iyi iç-tutarlılık gereksinmesi olduğunu göstermiştir. İlerideki çalışmalarda bu noktanın üzerinde çalışılması önerilmiştir. Genel olarak, bu çalışma envanterin Türkçe formunun stabiliteye ve iç-tutarlılığa sahip olduğunu destekleyen bulgular sergilemiştir.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

INTRODUCTION

In almost every culture children are brought up as hopes and representatives of the future. They are expected to have the best of everything-living standards, education, personal and professional qualifications, and so forth. However, achieving good living standards in today world is difficult. Recent technological advances necessitate high levels of qualification in man. Therefore, in comparison to the past, man needs to be much better trained and educated than ever before.

Generally, good education leads to better professions which facilitate the attainment of good living standards and high social status in a society. It is no different in Turkey where the major goal of the typical, middle-class, educated, urban parents is to provide every possible opportunity for the best education of their children. Such education is being provided by some private and special public schools in which personal and academic skill training of students exceeds that of the public schools. There are very few such schools, however, and the demand is much higher than their actual capacities can meet. Therefore, entrance to these schools is through selective central exams. These exams are very competitive in nature for which many children start preparing at early ages to be successful, and to get ahead. Thus, for some, the process of competition for better

education can be said to begin with pre-school training and it continues through university.

Importance of central exams in the lives of the Turkish children is real, because their results determine both their self-esteem, and their future professions. Through social pressures, the child is made well aware that his value and self-esteem are contingent upon the results of these exams. This awareness makes him anxious about exams. To some degree, the level of anxiety depends on the nature of parent-child interactions in the family. Generally, the more stress he feels coming from his parents about exams, the more anxious he becomes. Therefore, feelings of test anxiety among Turkish students who are about to take central exams, are very real.

Test anxiety is a viable issue in the lives of students during their school period, since student success is mostly determined by the results of classroom and final testing. At all levels of education test anxiety continues to be an important student feeling. Spring is the time when both final and central exams are crucial. It becomes a period of exam crisis both for students and their parents. During this period psychosomatic and psychological disturbances increase, and dramatic incidences such as suicides are seen. Referrals to school psychological services and mental health clinics multiply. Observations also indicate that more students suffer from test anxiety than just those who come in to consult, or who are referred to psychological services.

Considering the implication of test anxiety on the psychological development and the education of the Turkish children and youth, an understanding of the notion and the existing issue of such an anxiety seems vital. To achieve

this understanding, the concept needs to be answered systematically first, and then researched. The present study is the first phase of such an effort. It deals with the test adaptation of a particular topic of test anxiety. The research presented here is on the translation validity, and the reliability of a test anxiety scale on groups of Turkish students.

Significance of the Study

Although there is clearly a real problem of test anxiety among Turkish students, it has not been studied systematically by many researchers so far. One possible reason for this may be the lack of psychometrically well developed test anxiety scales. This lack might have discouraged the researchers in the fields of education and psychology, since in any scientific study objective and valid assessment of a fact is a necessary prerequisite.

In order to be able to study the issues and problems related to test anxiety among Turkish students, first an objective measurement scale is needed. The present research, therefore, addresses itself to fulfill such a need.

Purpose of the Study

The purpose of this study is to develop an adaptation of a test anxiety scale that can be used with Turkish students. For this purpose, a recent test anxiety scale called the "Test Anxiety Inventory" (TAI), developed by Spielberger in 1980, was selected. The original English form of this scale was translated into Turkish. Then the validity of this translation and the reliability of the translated form were studied.

CHAPTER II

CONCEPTUAL BACKGROUND AND REVIEW OF LITERATURE

Conceptual Background

The Concept of Anxiety

The abundance of literature in the area of anxiety implies that anxiety has been one of the major concerns of the behavioural scientist. It is a core construct in every personality theory, and its conceptualization differs from one theory to the other.

In the classical Freudian theory anxiety is conceived as the central problem of neurosis. Freud viewed it as an unpleasant affective state aroused in danger situations and accompanied by motor discharge (in Levitt, 1967). He called anxiety "objective" (reality) if the source of danger was from the external world, and "neurotic" if it was from the internal impulses. Freud's view of objective anxiety is almost identical with the present behaviouristic conceptualization of fear. In the case of objective anxiety, the individual perceives the source of danger consciously. On the other hand, the individual suffering from neurotic anxiety can not perceive the source of danger consciously, because it has been repressed. These repressed feelings are nothing more than the sexual and agressive impulses of the id (in Spielberger, 1966; and in Levitt, 1967).

Unlike Freud, who emphasized the role of impulses in the development of anxiety, the neoFreudian theorists, namely Sullivan, Horney, and Fromm underlined the importance of social processes in originating anxiety in the young child's life. In Sullivan's opinion a major source of anxiety for the child is the disapproval of the significant others.

Later, the child internalizes the values of others, and requires a new source of anxiety, self-disapproval (in Epstein, 1975).

Dollard and Miller, whose theory was based on experimental studies, perceived anxiety as a learned drive originating from the organism's basic tendency to avoid pain. Accordingly, once a pain is associated with a particular stimulus, feelings of anxiety begin. These feelings are generalized to other situations which somewhat resemble the original stimulus (in Levitt, 1967).

In the existentialistic point of view anxiety is aroused by events that have some implications of change in individual's self theory. According to Rogers (1951) threats to one's self-system arouse anxiety. Similarly, May (1950) holds the view that anxiety is the perception of threat to one's existence as a personality (in Levitt, 1967).

The Nature and Types of Anxiety

Until 1960 ies there were some ambiguous and inconsistent uses of the term anxiety among the anxiety researchers. Some studied it as a personality trait, while others dealt with it as a transitory state. With the factor analytic studies of Cattell and Scheier in 1958, it was identified that anxiety had two distinct types, called state anxiety and trait anxiety (in Spielberger, 1966). This distinction was further delineated and researched by Spielberger, Gorsuch, and Lushene, 1970). Their findings and conceptualizations helped to clarify some of the confusion in the research results on anxiety.

Spielberger defined anxiety as a general phenomenon and to be "subjective, consciously perceived feelings of

apprehension and tension, accompanied by or associated with activation or arousal of the autonomic nervous system" (Spielberger, 1966, pp.16-17). However, in his theory two conceptually different anxieties, state and trait anxiety, are recognized. State anxiety is described as a momentary, emotional arousal, and trait anxiety as a relatively stable personality characteristic. In other words, trait anxiety can be thought of as one's anxiety-proneness or predisposition to perceive situations as threatening. The distinction between state and trait anxiety is likened to the distinction between kinetic and potential energy in physics (Spielberger, 1966).

Within the framework of this conceptualization anxiety is an unpleasant emotional reaction that results from the perception of a situation as threatening (Schwarzer, Ploeg, and Spielberger, 1982). A person perceives a situation threatening, if he anticipates that this situation will lead to a personal loss or failure. Defares states that "people are sensitive to stressors, which imply some kind of judgment by significant others, regarding an activity which is highly valued by the person" (Defares, Grossman, and Swart, 1983, p.87).

Another kind of distinction in the conceptualization of anxiety is very recent. This distinction pertains to specific anxieties through which general anxiety is replaced by situation specifics, such as illness anxiety, dental anxiety, death anxiety, mathematics anxiety, computer anxiety, social anxiety, test anxiety, and so on.

Observations and research indicate that one pervasive type of anxiety is test anxiety. In school settings testing frequently evokes anxious responses among students for the reason that its results are vital for students. They can lead

to negative evaluations of an individual if a minimum standard is not attained. From this perspective, test anxiety can be conceived as a situation-specific anxiety.

The Concept of Test Anxiety

In the present study, research related to conceptualization of test anxiety is examined in a historical order. By this way it will also be possible to follow the developmental phases of test anxiety research.

Although the problem of test anxiety is probably as old as the use of tests for evaluative purposes, importance of exam stress was begun to be recognized at the beginning of the 20th century. In 1929, Walter Cannon observed the physiological changes during a real-life stress situation, i.e., academic examination. A Russian physiologist, Alexander Luria, was the first to emphasize the individual differences in test anxiety. He studied the emotional reactions of students and classified them as "stable," and "unstable" (in Spielberger, Gonzalez, and Fletcher, 1979).

The first systematic study related to test anxiety was the work of Seymour Sarason and George Mandler (1950). In their pioneering work at Yale University, Sarason and Mandler observed that test anxiety led to some performance decrements in evaluative conditions (in Sarason, Davidson, Lighthall, Waite, and Ruebush, 1960).

In the following studies, Sarason (1960) claimed that a test anxious response has two major effects. First, it narrows the perception of the external field; second, it prevents an objective assessment of the nature of problem solving task. In his view, there are some anxiety drives involved in testing situations. These drives are learned,

and can be either relevant or irrelevant to the task at hand. Task-relevant anxiety drives facilitate test performance, while task-irrelevant anxiety drives decrease it (Sarason, et.al., 1960).

Beginning from 1958 till 1965, Irwin G.Sarason studied the situational factors that contributed to differences in the performance of high and low test-anxious students. It was found that high test-anxious individuals were more self - centered and self-critical than their low test-anxious counterparts (in Spielberger, et.al., 1970). The attention hypothesis suggested by Wine (1971) supported this view of Sarason. Wine stated that a high test-anxious individual is occupied with task-irrelevant cognitions which lead to distraction of attention, and result in performance decrements (in Stephan, Fischer, and Stein, 1983).

In a recent article, an important misconceptualization about test anxiety is pointed out by Phillips (1985). This misconceptualization is that in describing test anxiety most people think of the testing environment only. However, test anxious feelings are pervasive in several classroom activities which have some kind of evaluative significance. Some of these test-like activities are reading orally, participating in discussions, answering questions, and doing assignments in the classroom environment (Phillips, 1985). From this perspective, it is not only the testing itself, but also all kinds of evaluative situations in the classroom evoke anxious responses from the students.

The Components of Test Anxiety

While some researchers explored the situational factors in test anxiety, others studied the compositional structure of it. In 1967, Liebert and Morris introduced a two-component

conceptualization of test anxiety. These components were called worry and emotionality. In their initial conceptualization this distinction was based on state anxiety, i.e., the degree of worry and emotionality experienced in specific testing situations. These components were perceived as conceptually independent, but somehow covarying in testing situations (in Morris, Franklin, and Ponath, 1983).

This distinction was later accepted and developed by Spielberger (1980). Unlike the previous conceptualization of test anxiety with state characteristics, he based the new view on trait characteristics. In this approach test anxiety is a situation-specific personality trait and composed of a cognitive component called "worry", and an affective component called "emotionality" (in Schwarzer, Ploeg, and Spielberger, 1982). Schwarzer and his colleagues claim that "In essence, test anxiety refers to individual differences in the disposition to experience feelings of apprehension, and worry cognitions in academic environments where the performance of students is under scrutiny" (Schwarzer, et. al., 1984, p.4).

The worry component of test anxiety is described as one's cognitions about his performance, his poor self - evaluation, and thoughts of consequences of failure. The worrying individual doubts about his performance, thinks how much brighter the others are, and perceives himself as being highly vulnerable to failure. The emotionality component is the affective-physiological arousal, i.e., rapid heart rate, tenseness, nervous feelings, stomach upset, uneasiness, etc., all of which are experienced in evaluative conditions (Minsel and Schwarzer, 1983). However, emotionality is not the arousal itself, but the subjective perception of these internal events (Schwarzer, 1984).

The relationship between the worry and the emotionality components of test anxiety is still a focus of attention for many researchers. In the literature, the reported correlation coefficients between worry and emotionality are found to vary from moderate to high positive (Minsel and Schwarzer, 1983; and Schwarzer, 1984).

It seems that the worry and the emotionality components of test anxiety are interrelated, and their relationship is quite complex. The most consistent finding concerning this worry-emotionality relationship is that the worry component is correlated negatively with the level of academic performance, while the emotionality component does not lend itself to such a relationship (Spielberger, 1980; Becker, 1982; Schwarzer, et.al., 1982; Morris, et.al., 1983; and Schwarzer, 1984).

In explaining these different relationships Spiegler, Morris, and Liebert (1968) state that emotionality is a classically conditioned vegetative, affective response. It is triggered in the beginning of an exam, but its intensity declines throughout the exam. Worry cognitions, however, are related to negative appraisal of performance during an exam, and their effects on performance continue throughout the exam process. Worry is, therefore, closely and negatively related to performance (in Stephan, et.al., 1983). Stephan (1983) does not find this explanation convincing and claims that because the authors do not specify the meaning of conditioned and unconditioned stimulus and response, verification of this hypothesis is impossible.

Although in the test anxiety literature the worry - emotionality distinction is generally accepted and found useful, recent studies have begun to question this distinction. After reviewing an intensive literature, developing a new

model, and testing it by factor and path analyses, Salame' (1984) concluded that what had been conceptualized as "worry" by Spielberger and his co-workers, was, in fact, consisted of four separate factors; "doubts about one's performance," "lack of confidence in ability to succeed," "fear of failure and its consequences to academic career," and "fear of social devaluation." Salame's study also showed that the emotionality component should be divided into two factors; "aaprehension" and "other anxiety reactions." Depending on his new conceptualization Salame' developed a new questionnaire "Salame' Test Attitudes and Reactions" (STAAR).

The Theories on Test Anxiety

In the test anxiety literature two different theoretical formulations are seen. These are the interference model and the learning deficit model. In this study these models are reviewed briefly.

Most studies on test anxiety, particularly the initial ones have been based on a theoretical approach called the *interference model* (Hodapp, 1983). In this model, the highly test-anxious individual is characterized by self-preoccupied (i.e., attending to internal events rather than the task), self-centered worry responses that interfere with the subject's performance. Negative impact of anxiety occurs during testing. Anxious feelings prevent the individual from utilizing his previous knowledge and skills competently, and from using task-relevant cues successfully (Sarason, 1975).

The studies favouring the interference model are mostly based on the findings of the continuous works of I.G. Sarason, and the attention hypothesis suggested by Wine (1971). In these studies the hypotheses derived from the interference model were tested (Deffenbacher, 1978; Galassi,

Frierson, and Sharer, 1981; and Zatz and Chassin, 1983).

Another group of studies focused their efforts on the reduction of the negative effects of test anxiety. In these studies behavioral and cognitive techniques, such as relaxation therapy, systematic desensitization, and cognitive modeling were implemented. However, the results were not satisfactory enough. The reduction obtained in the test anxiety levels of students did not mitigate the observed negative results on performance. The positive changes found in students' performance were not significant (Bruch, Juster, and Kaflowitz, 1983).

Such findings led some researchers to develop more comprehensive theories. They thought that since the existing approaches did not explain all the obtained results, there should be some other variables involved in test anxiety. In 1980, Meichenbaum and Butler developed a new theoretical explanation for test anxiety. In this approach they proposed four sets of variables that are highly interdependent in both test anxiety and its treatment. These variables are (a) the content and patterning of the internal dialogues or covert self-statements; (b) cognitive structures, consisting of the student's broader and more enduring meaning systems concerned with academic evaluation; (c) behavioural acts, such as study and test-taking skills; and (d) behavioural outcomes such as grades, physiological reactions, mood states, and subjective perception of anxiety (in Galassi, et. al., 1981; and in Bruch, et.al., 1983).

At about the same time with Meichenbaum and Butler, Culler and Holahan (1980) proposed a new approach called the *learning deficit model*. Unlike the interference model, this model suggested that anxiety plays an indirect role in testing situations via the impairment of study-related

behaviour during the preparation for a test (in Hodapp, 1983).

This new theoretical formulation was originated from the previous studies which led to question the assumptions of the interference model. One of them was the study of Wittmaier (1972) which showed that high test-anxious students had significantly lower levels of study skills competence when compared to low test-anxious students (in Culler and Holahan, 1980).

The explanation of the interfering effect of study behaviour, which is external to testing situation, on performance is made by the worry responses. It is stated that "the worry responses generated in the testing situation are a product of both the increased anxiety and the knowledge that there has not been adequate preparation for the test" (Culler and Holahan, 1980, p.19).

Theoretical bases of the most recent studies show some differences. Some favour the learning deficit model and discard the interference model, others support both of the models, and finally some suggest totally different, new views.

A recent paper of Covington (1985) sheds a new light on test anxiety conceptualization. Covington asserts that most of the test anxiety literature has been focused on the demonstration of the negative relationship between anxiety arousal and achievement outcome, but very little attention is paid to anxiety as a cluster of interactive mediating factors. Covington claims that test anxiety is not a unified, single reaction to a threat, but a cluster of interrelated factors in the achievement cycle. He states that achievement cycle is consisted of four stages; test

anticipation, test preparation, test taking, and test reaction. Covington adds that the role of test anxiety changes from stage to stage, and it can best be understood if one integrates it into a larger context, i.e., the achievement cycle.

The theoretical background of the present study is based on Spielberger's conceptualization of test anxiety. This conceptualization is mainly derived from the interference model. According to Spielberger (1980) test anxiety is a situation-specific personality trait which is composed of the worry and the emotionality factors.

Review of Literature

Research on test anxiety has a relatively young history, yet its literature is abundant. A group of test anxiety researchers founded their own society in 1980, called the *International Society For Test Anxiety Research*. This society sponsors test anxiety conferences every year and publishes the conference proceedings in a book called *Advances in Test Anxiety Research*. Up to now four volumes of this book have been published, and these have constituted the major source of the literature review presented here.

In this chapter some of the studies on test anxiety are reviewed briefly. These studies are mainly related to (1) the nature of test anxiety according to a particular theoretical model; and (2) the impact of test anxiety on some cognitive and personality variables.

The Studies Based on the Interference Model

The studies based on the interference model generally revealed that high test-anxious individuals show performance decrements under evaluative conditions, and exhibit more task-irrelevant behaviours than low test-anxious individuals.

The pioneering studies derived from this model were carried out at Yale University. In these studies it was observed that the performance levels of anxious students approached the performance level of nonanxious students when the evaluative exam stress was minimized (Sarason, et.al., 1960). Supportive findings were gathered in subsequent studies. For example, Sarason and Harmatz (1965) conducted a study on 144 highschool sophomores which were divided into high, middle, and low test-anxious groups according to their scores on the Test Anxiety Scale (TAS). Subjects learned two difficult lists of dissyllable words under high, low, and neutral motivational or achievement-orienting instructional conditions. It was found that low test-anxious students showed superiority over the high and middle test-anxious subjects. When the results were analysed in terms of errors, it was observed that high test-anxious subjects made fewer errors under the neutral than under the other two experimental conditions. The opposite finding was obtained for low test - anxious subjects. The researchers interpreted this finding that under high motivational on personally threatening conditions the performance of high test-anxious students was more distrupted than those of low test-anxious students.

Sarason (1973) conducted another study as evidence for the view that under evaluative stress high test-anxious subjects are more sensitive to task-related cues than are low test-anxious subjects. A total of 120 high and low test-anxious female university students were given to solve

anagram tasks under four different experimental conditions. In one of these experimental conditions an experimenter solved the anagrams while the subjects observed. In the second condition only the verbal description of the manipulation of the anagrams was told. In the third, the experimenter stated the problem solving principles of the anagram tasks. The last was the control group who took the anagram problems without any kind of assistance. In the results it was observed that under the experimental condition in which the problem solving principles were told, this was called "cognitive modeling," high test-anxious subjects solved the anagrams more quickly than did low test-anxious subjects. These findings suggested that high test-anxious subjects show significant improvements as a function of modeling reassurance, and psychological support (Sarason, 1973; and 1975).

The study carried out by McCoy (1965) also yielded similar results supporting the interference model. In this study 28 high, and 28 low test-anxious fourth grade boys were selected as subjects on the basis of their scores on the Test Anxiety Scale for Children (TASC), and the Defensiveness Scale for Children (DSC). Half of the subjects were given the tasks, which were related to tracing and drawing, with test instructions, while the other half were given the same tasks with game instructions. On the tracing task, high test-anxious children made fewer errors than low test-anxious children when the tasks were given in game-like conditions. On the drawing task, high test-anxious subjects made smaller drawings and less verbalization under the test condition and tended to use more colours and materials under the game-like condition.

Similarly, Smith, Ascough, Ettinger, and Nelson (1971) divided 215 university students into low, moderate, and high

test-anxious groups according to their scores on the TAS. The subjects were administered a course examination in which half of the students received a form of examination containing humorous test items, while the others received a nonhumorous form of the examination. It was found that the high test - anxious subjects who received the nonhumorous form performed significantly lower than did the low or moderate test-anxious subjects. On the other hand, the high test-anxious subjects receiving the humorous form not only outperformed the high test-anxious subjects with the nonhumorous form, but also equaled the performance level of the low test-anxious group.

Ganzer (1968) conducted a study in which females with different TAS scores were given 25 learning trials on a list of nonsense syllables under the presence and absence of an observer. He found that observer presence was detrimental both for high and middle anxious groups. However, observer presence did not affect the performance of the low anxious group. This finding is consistent with the theoretical view of the interference model that high test-anxious individuals are very sensitive to evaluation.

Although nonevaluative, supportive experimental behaviours do not always mitigate the detrimental effects of test anxiety (Marso, 1970; and Öner, 1971), in general, findings favour the view that high test-anxious subjects perform well, if evaluative nature of the conditions is reduced to minimum.

Another characteristics of test anxious subjects is their task-irrelevant behaviours. In a study of Nottelmann and Hill (1977) 48 fourth and fifth grade children, assigned to low, middle, and high test-anxious groups by their scores on the TASC, were given to solve anagram tasks with the presence of an experimenter who also dealt with anagrams, and

their off-task behaviours were measured by recording the frequency and direction of off-task glancing. Their results showed that high test-anxious children had the lowest performance scores and exhibited substantially more off-task glancing than the low and middle test-anxious children. The authors proposed that the history of failure of the anxious child makes him more sensitive to external feedback, and less likely to solve problems on his own. The finding that high test-anxious children have more off-task glancing also supports the view that high test-anxious subjects exhibit more task-irrelevant behaviour than do low test-anxious subjects.

The attention hypothesis, related to task-relevant or irrelevant behaviours of anxious individuals, proposed by Wine (1971) was tested by various researchers. For instance, Deffenbacher (1978) administered the TAS to 185 male and female sophomore university students. From this pool 68 volunteers, half of these were high test-anxious, and the half were low test-anxious, took part in the study. The subjects were given to solve difficult anagram tasks under high stress (evaluative) and low stress (nonevaluative) conditions. The findings showed that the high test-anxious group (a) reported more anxiety during testing; (b) rated their abilities and the task negatively; (c) solved fewer anagrams; (d) estimated spending less time on task; (e) experienced more interference from anxiety; and (f) reported greater distraction of attention than the low test-anxious group.

Evidence supporting the hypotheses that high test - anxious subjects experience more negative than positive thoughts than low test-anxious subjects is also derived from the interference model. In a recent experimental study (Galassi, Frierson, and Sharer, 1981) 231 low, moderate, and

high test-anxious college students were included as subjects. These students were observed and given a Checklist of Positive and Negative Thoughts, a Subjective Unit of Disturbance Scale, and a Checklist of Bodily Sensations at the beginning, middle, and end of an actual exam. The findings revealed that high test-anxious students experienced significantly fewer positive thoughts, and more negative thoughts than low test-anxious students. Interestingly, the most frequent thought of the high test-anxious group was about escaping from the testing situation. Other frequent thoughts were about the test being hard, insufficient time to finish, and the likelihood of obtaining a poor grade. On the other hand, the most frequent thought of the low test-anxious group was "will do all right on test." In addition, high test-anxious students had lower grades, and experienced more disturbance and bodily sensations indicating arousal than their low test-anxious counterparts.

Similar results were obtained by Zatz and Chassin in 1983. Their study consisted of 294 sixth graders who were screened on the TASC and the DSC measures and then divided into low, moderate, and high test-anxious groups. After being given the instructions intended to evoke anxiety, subjects performed on two experimental tasks, the Coding-B subtest of the Wechsler Intelligence Scale for Children - Revised (WISC-R) and an anagram task adapted from Stevenson and Odom (1965). After the completion of tasks, subjects were given the Children's Cognitive Assessment Questionnaire (CCAQ) to report their thoughts during testing. It was found that high test-anxious subjects reported significantly more task-debilitating cognitions than either moderate or low test-anxious subtects, including negative evaluations and off-task thoughts.

In sum, all these reported studies, supporting the

interference model, revealed that high test-anxious individuals are very sensitive to evaluation and show performance decrements, task-irrelevant behaviours, and negative thoughts under conditions that imply some kind of judgement. These studies also showed that performance level of high test-anxious students can be increased, if the evaluative nature of the situation is reduced or removed.

The Studies Based on the Learning Deficit Model

The studies based on the learning deficit model suggest that performance decrements of high test-anxious individuals can be explained by their ineffective acquisition of academic skill, particularly the study skills (Wittmaier, 1972; Culler and Holahan, 1980; and Hodapp, 1983).

Kirkland and Hollandsworth (1980) conducted different treatment groups to alleviate test anxiety levels of 231 college students. The treatment groups constituted a skills-acquisition group, cue-controlled relaxation group, mediation group, and a practice-only group. They were compared in terms of GPAs, performance on the Otis-Lennon Mental Abilities Test-Advanced Level, and on a difficult 10-item anagram task. No performance differences were obtained on the Otis-Lennon Test, but the other two measures, GPA and the experimental task, yielded some differences among groups. It was observed that the skills-acquisition group in which the training was focused on effective test-taking strategies, adaptive self-instructional statements, and attentional - control skills, was superior to the other groups in both GPA and experimental task measures. In this group no emphasis was given to arousal or anxiety reduction. Instead, skills - acquisition was stressed and it was underlined that the above skills-acquisition would allow the individual to utilize his arousal or anxiety effectively to increase

performance. The investigators discussed the results and suggested that test anxiety could be defined as a specific skill deficit. They concluded that "it is time to give the phrase *test anxiety* a respectful burial and talk about inadequate test performance in terms that more accurately describe what it is, namely, ineffective test taking" (Kirkland and Hollandsworth, 1980, p.438).

Culler and Holahan (1984) investigated the relationships among test anxiety, academic performance, and study-related behavious on 96 first-semester freshmen college students. Subjects were regarded as high or low test-anxious according to their TAS scores. The two groups were compared on their study hours, cramming, number of missed classes, late exams, and GPAs. The study behaviours of subjects were also assessed by the Survey of Study Habits and Attitudes measure of Brown and Holtzman. The comparisons showed that high test-anxious students had lower GPAs and poorer study-skills than low test-anxious students. There was a significantly positive correlation between study skills and GPA. It was also found that high test-anxious students spent more study time than low test-anxious students. Culler and Holahan interpreted this finding to mean that high test-anxious students attempted to compensate for their lower study competence by increasing the total study time.

Studies related to the relationship between test anxiety and academic skills did not always give consistent results. For example, in their study Brown and Nelson (1983) failed to find significant relationship between test anxiety and academic skills. Seventy two undergraduates, selected on the basis of their TAS scores, participated in this study. Several measures were administered to assess cognitive - attentional and somatic components of test anxiety and test - taking and study skills. The results showed that high test -

anxious students differed from their low test-anxious peers on traditional cognitive and somatic indicators of test anxiety, but not on any measure of study or test-taking skills. It was observed regardless of their anxiety scores, students with high grades scored higher on academic skills measures. The authors concluded that academic performance may be more strongly associated with skills involved in studying and taking tests than with cognitive and somatic variables related to test anxiety.

As a source for construct validity of his newly developed test anxiety scale Spielberger (1980) reports low to moderate negative correlation coefficients between test anxiety and study skills. These correlations were somewhat higher for males and highschool students than females and college students, and can be interpreted to provide data for the learning deficit model.

The theoretical formulation of the research conducted by Benjamin, McKeachie, Lin, and Holinger (1981) seems to incorporate the ideas of both the interference and the learning deficit models. This formulation was based on the information processing model of cognitive psychology. In this model, information is processed in stages. The first stage is encoding, the second is storage and organization, and the last stage is retrieval. Depending on this model, Benjamin and his colleagues hypothesized that performance deficits of high test-anxious subjects could be explained in terms of problems in encoding, organizing, and retrieving the information in testing situations. To test their hypothesis they conducted research consisting of two studies, and a total of 194 undergraduates scoring low, medium, and high on the Worry and Emotionality Questionnaire (WEQ) of Liebert and Morris (1967). Difficulties in the course, study hours, study habits as assessed by the Survey of Study Habits and Attitudes

Scale of Brown and Holtzman (1967), GPAs, performance on course exams, and performance on a take-home exam were recorded. The results showed that high test-anxious subjects had lower GPAs and reported more problems in learning and reviewing the course material. Though it did not reach significance, high test-anxious subtects also reported more study time than low test-anxious subjects. This finding was consistent with the findings of Culler and Holahan (1980). In addition, performance of high test-anxious students on course exams and take-home exam revealed that these students had problems not only in retrieval of information during the testing situation, but also in encoding (learning) the information beforehand.

In brief, the studies supporting the learning deficit model showed that academic skills, particularly study skills of students are important in the performance deficients of high test-anxious individuals. The most important implication of these studies is that any treatment group designed to alleviate test anxiety should also include teaching study or test-taking skills techniques.

Impact of Test Anxiety

Research on the impact of general anxiety demonstrates both facilitating and debilitating effects on individual's behaviour and performance. The majority of the findings, however, seems to focus on its debilitating effects. The contradictory data may be due to differences in approaches derived from different theoretical models.

Studies on the positive or facilitative impact of anxiety are based on behaviouristic, Hull-Spence Drive Model in which anxiety is conceptualized as an energizing drive (in Öner, 1981). On the other hand, studies based on its negative

or debilitating effects are based mainly on cognitively and psychoanalytically oriented models.

The studies related to facilitative effects of anxiety showed that high anxious subjects do better than low anxious subjects on simple on easy tasks. The reverse was true for the complex tasks. Moreover, it was found that high anxious subjects were more easily conditioned than low anxious subjects. The observed effect of manifest anxiety was generally positive for highly intelligent individuals, but negative for those with low intelligence (Spielberger, 1966). Several reviews of literature on the facilitative effects of general anxiety (Sarason, et.al., 1960; Öner, 1977; and 1981) have indicated that the supporting data were not too strong for this view.

Considering the moderately positive correlations between general anxiety and test anxiety (Sarason, et.al., 1960; and Spielberger, 1980) it might be expected that all these findings in general anxiety literature have also certain implications for test anxiety. Similar to the general anxiety literature, most studies in the test anxiety literature dwell on the negative or debilitating effects of test anxiety. Hence, in the present study the literature focusing on the negative impact of test anxiety are reviewed. Rather than being causal, these studies are generally correlational in nature. Therefore, it seems useful to review some major correlates of test anxiety separately.

Test anxiety and academic performance. Test anxiety literature is abundant with studies trying to depict the relationship between test anxiety and academic performance. It seems that one of the major correlates of test anxiety is academic performance. This is probably because the concept of test anxiety itself is most viable in academic settings.

The general finding of the studies exploring the relationship between test anxiety and academic performance is that test anxiety is negatively related to academic performance.

In reviewing the anxiety research Tobias cited Spielberger expressing that "more than 20 % of students who were characterized as highly anxious dropped out of school because of academic failure whereas only 6 % of a low anxious group left school for such reasons" (Tobias, 1979, p.573).

The majority of the findings revealed that test anxiety levels and GPAs of students are significantly and negatively correlated. This finding has been repeatedly confirmed by various studies over a twenty three year span (Sarason, et. al., 1960; Spielberger, 1967; Pulvino and Hansen, 1972; Wittmaier, 1974; Spielberger, 1980; Culler and Holahan, 1980; Benjamin, et.al., 1981; Morris, et.al., 1983; and Minsel and Schwarzer 1983). Recent studies have also showed that it is the worry component, and not the emotionality component of test anxiety, which correlates negatively with school achievement as assessed by grades (Becker, 1982; Deffenbacher, 1980 and Hodapp, Laux, and Spielberger, 1982 Minsel and Schwarzer, 1982; Morris, et.al., 1983; and Schwarzer, 1984).

Test anxiety and intelligence. The data confirming the significant negative relationship between test anxiety and intelligence are based on psychoanalytically oriented Yale studies conducted by S.Sarason and his colleagues in early 1960 ies (in Öner, 1971).

In 1956 Zweibelson proposed that the negative correlations between test anxiety and intelligence are due to test-like, threat producing aspects of intelligence tests,

and claimed that these negative correlations could be reduced considerably by removing the test-like characteristic of intelligence measures. He selected the Otis Beta intelligence Test as having the most typical test-like, threatening quality of an assessment instrument. He also used the Otis Alpha Tests and the Davis-Eells games as having game-like characteristic. On a group of 258 fifth grade elementary school children he found significantly negative correlations ($r = -.24$) between a test anxiety scale, the TASC, and the Otis Beta Test, while the correlations of the TASC with other intelligence tests which seemed game-like in nature did not yield significant results. In this finding the hypothesis that negative correlations between intelligence and test anxiety could be reduced by the removal of the test-like characteristic of the instrument was supported (in Öner, 1971).

As evidence for the validity of the TASC, Sarason (1960) correlated the TASC scores with Thurstone's (1954) tests of Primary Mental Abilities (PMA) on 553 third and fourth grade children. He found a significantly negative overall correlation ($r = -.24$), and observed that correlations were higher on the test-like subtects of the PMA than on its non-test-like subtects. These data are consistent with Zweibelson's (1956) findings. Significant negative correlations between test anxiety and intelligence were further supported in later studies (Sarason, Hill, and Zimbardo, 1964 and Feldhusen and Klausmeier, 1962 in Öner, 1971).

Öner (1971) found significant negative correlations between test anxiety and intelligence. In her study the correlations between test anxiety and performance on a learning material decreased when intelligence was held constant, indicating the importance of intelligence as a

variable in the level of test anxiety.

Though there were some contradictory data in test anxiety literature Sarason (1960) concluded that the determining factor in the negative test anxiety-intelligence relationship was the level of test anxiety, rather than the level of intelligence. When the intelligence of subjects was controlled, there were still performance differences among subjects' levels of anxiety. This finding supported the view of Sarason.

Mandler and Sarason (1952) hypothesized that correlations between intelligence and test anxiety would approach zero in intellectually superior students. They administered a test anxiety scale and the Henmon-Nelson Test of Mental Ability on a group of highly intelligent undergraduate Yale students. Their findings, however, did not confirm their expectations, as they found a correlation of -.21 between the two measures. Sarason (1960) interpreted this finding to indicate that level of intelligence was not a plausible explanation for the negative relationship between test anxiety and intelligence.

In his recent study on the relationship between test anxiety and intelligence Ploeg (1984) administered the adaptations of Spielberger's Test Anxiety Inventory and the Stait Trait Anxiety Inventory (STAI) on 154 secondary school children. The intelligence levels of students were also assessed. In the analysis of the results Ploeg found that girls with lower level of intelligence were hardly influenced by higher levels of test anxiety, but in the groups with higher level of intelligence, high test anxiety produced large performance decrements.

Ploeg obtained similar findings in his subsequent

studies. He found that the debilitating effects of test anxiety were greater among brighter students compared to its relative impact on less able individuals (Ploeg, 1985).

Most of the studies designed to explore the relationship between test anxiety and intelligence revealed significant, negative correlations. Somehow, there were other studies which failed to show significant relationship between test anxiety and intelligence. For instance, Doyal and Forsyth (1972) assessed the test anxiety and intelligence levels of 37 third grade students, and obtained nonsignificant relationship between these measures.

Spielberger (1980) investigated the test anxiety - intelligence relationship to constitute data on the validity of the TAI. He correlated the TAI scores with the Otis-Lennon Intelligence Test scores of 117 highschool students. He found that although the correlations did not reach significance they were in the negative direction.

To summarize all these findings it can be stated that although some studies failed to depict significant relationship, the majority revealed significant negative relationship between test anxiety and intelligence. It seems that the level of intelligence, is responsible for this negative relationship. It is also observed that the correlations between test anxiety and intelligence increase with the evaluative nature of the testing situation.

Test anxiety and some cognitive variables. Some studies in test anxiety research are directed to explore the relationship between test anxiety and some cognitive variables, such as problem solving, different kinds of learning, level of processing, and convergent-divergent thinking.

Harleston, Smith, and Arey (1965) investigated the *test anxiety-problem solving* relationship on 42 female college students. The subjects were characterized as low, medium, and high test-anxious according to their scores on the test-anxiety questionnaire developed by Harleston (1962). Sixty anagrams selected from high and low frequency words as rated by Thorndike and Lorge (1944) were given to students and their heart rates were recorded. It was found that though the differences did not reach statistical significance, the high test-anxious group solved fewer anagrams than the low and medium test-anxious groups. The low test-anxious group had shorter solution times. It was also found that high test-anxious subjects produced significantly larger increases in heart rate than low test-anxious subjects. Large increases in heart rate were consistently associated with poorer anagram solving.

Marlett and Watson (1968) carried out a study on 56 subjects from a pool of 220 ninth grade males. On the basis of their scores on the Achievement Anxiety Test (AAT) of Alpert and Haber half of the sample were assessed as high, and the other half as low test-anxious groups. The task which was related to learning series of correct responses was given to subjects under two experimental conditions, delay or no delay. In other words, subjects were observed under immediate and delayed failure feedback conditions. The initial response latency of the high test-anxious group was the greatest. It was found that after repeated failures the response latency of the high test-anxious group was increased. It was also observed that high test-anxious subjects were more redundant in their problem solving.

In their study related to interactions among test anxiety, intelligence, sex, and problem solving ability Doyal and Forsyth (1972) found that the problem solving ability of high test-anxious girls was interfered with the blocking effect of anxiety. However, the opposite was observed in case

of boys. The high levels of test anxiety had facilitative effects on their performance levels.

Similar findings were obtained by Deffenbacher (1978) who compared the problem solving levels of high and low test-anxious 185 undergraduates. The tasks were 13 high-difficulty anagrams. As a result it was found that in comparison to low test-anxious students, high test-anxious students solved fewer anagrams.

The studies focused on the relationships between *test anxiety and different kinds of learning*. For example, Öner (1971) dealt with the learning of addition, subtraction, multiplication, and division in decimals. One hundred and sixty sixth grade children were included in this study. They were selected on the basis of their scores on the Lie Scale for Children (LSL), the Defensiveness Scale for Children (DSC), and the TASC. The subjects learned the material under four instructional styles, namely; (1) feedback-supportive; (2) feedback-non-supportive; (3) no feedback-supportive; and (4) no feedback-non-supportive. She found that under all conditions low test-anxious subjects, especially girls, demonstrated a higher level of performance than high test-anxious subjects. On a retest by which retention was assessed, low test-anxious students also scored significantly better than high test-anxious students.

The relationship between test anxiety and learning a complex verbal task was investigated by Ray, Katahn, and Snyder (1971) on a group of 122 male university subjects. Subjects were divided into low, medium, and high test-anxious groups according to their scores on the STAI modified by Snyder and Katahn (1970) to estimate test anxiety before a classroom. The learning task which consisted of a set of 16 information items and two sets of 14 questions, was administered during an actual class. Then, retention and generalization tests were given to these subjects two days later. Significant differences were found both on the

retention and the generalization tests between the test anxiety groups. These differences were in favour of the low test-anxious students.

In a study of Weinstein, Cubberly, and Richardson (1982) the relationship between *levels of processing and test anxiety* was explored. Ninety female university students whose test anxiety levels were assessed by the Factor 1 subscale of the Test Anxiety Scale (Richardson, O'Neil, Whitmore, and Judd, 1977) were randomly assigned to either the superficial level or the deep level processing group. The learning task which was a lesson about a learning strategy that could be used for learning lists of paired associates, and the direction designed to arouse anxiety in subjects were given. The students in the superficial processing group were required to examine only structural similarities or differences between the two words in a pair, whereas the students in the deep level processing group were required to perform more meaningful analyses of the words. After receiving these instructions the subjects were given an opportunity to practice their learning strategy on a list of 10 word pairs for about 10 minutes. Then, one list of 30 word paired-associate learning task was given to subjects and their performance levels were evaluated. The results revealed that the performance of low test-anxious students was not significantly different from high test-anxious students' performance on the superficial level processing task. However, on the deep level processing task the performance of low test-anxious students was significantly higher than the performance of high test-anxious students.

Similarly, Fransson (1977) found that strong interest in task, together with low test-anxiety, produced deep level of learning, while lack of interest and high test anxiety produced surface level of learning.

Test anxiety-convergent and divergent thinking relationships were investigated by Vidler (1974) on 212

undergraduate students. The study was based on Guilford's model of intelligence in which convergent thinking was identified with the kind of thinking that involves logical reasoning toward single "right" answers. In this model divergent thinking was used as a synonym for "creativity," and described as generation of a variety of responses. The Test Anxiety Questionnaire (TAQ) and convergent-divergent thinking measures as well as a curiosity measure were administered to subjects. Vidler's findings revealed that both convergent and divergent thinking correlated negatively with test anxiety.

Briefly stated, the studies concerning the relationships between test anxiety and school achievement, intelligence, and other cognitive variables indicated a negative direction. These findings seem to imply that the more anxious the individual the poorer he performs on tasks that require intellectual performance.

Test anxiety and some personality variables. The relationships between test anxiety and some personality variables, such as learned helplessness, defensiveness, curiosity, locus of control, self-concept, and self-esteem have been investigated by various researchers.

Lavalle, Metalsky, and Coyne (1979) studied the *test anxiety-learned helplessness* relationship on 72 university students. The TAQ was used to select high and low test-anxious students. The subjects received either controllable or uncontrollable noise in a typical helplessness induction. Half of the subjects subsequently received an acknowledgment of contingencies in the induction task, and the others did not. Then an anagram task was given to subjects. It was found that only high test-anxious subjects were debilitated by the helplessness induction.

Schwarzer and Cherkes-Julkowski (1982) conducted a study in which 484 fifth grade students served as subjects. By several measures, test-anxiety, self-concept of ability,

lack of self-confidence, self-esteem, hope for success, and social anxiety levels of students were assessed. The correlations showed that these factors are interrelated. In explaining the complex interrelationships among these factors, the authors stated that "Clearly the concepts of test anxiety, learned helplessness, and academic self-concept are closely related. Anxious self-preoccupation and self-doubt can be associated with moderate state of test anxiety. Low self-esteem can be viewed as an indicator of helplessness. These factors seem to be associated in a series of convoluted, multidirectional relationship" (Schwarzer and Cherkes-Julkowski, 1982, p.35).

The negative relationships between *test anxiety and self-concept and self-esteem* are consistent with both the previous (Levitt, 1966) and the recent (Mueller and Thompson, 1984) findings.

Levitt (1966) reviewed some of the relevant literature and identified that anxious people are less motivated by curiosity than non-anxious people. This was verified by a later study (Vidler, 1974) which explored the relationship between test anxiety and convergent-divergent thinking on 212 undergraduates. Vidler also investigated the relationship between curiosity and convergent-divergent thinking. It was found that though they were not always significant, the correlations between curiosity and convergent-divergent thinking were in the positive direction. Considering the significantly negative relationship between test anxiety and convergent-divergent thinking (Vidler, 1974) one can speculate that subjects, *test anxiety and curiosity* levels were in the opposite direction. In fact, the data reported by Spielberger (1980) supported this speculation. On a group of 246 community college students he found small but statistically significant negative correlations between the TAI and the

Trait scale of the Stait-Trait Curiosity Inventory scores of students.

Some researchers investigated the relationship between *test anxiety and locus of control*. Locus of control is an important personality variable related to one's perception of his own role in controlling the environment. When the orientation of control is internal, it is called internal locus of control, and when the orientation is external, it is called external locus of control. In the literature locus of control was found to be correlated with some variables, such as socioeconomical status, level of aspiration, academic achievement, parental child-rearing attitudes, and anxiety.

On a group of 2438 ninth grade highschool students Bar-Tal and his colleagues (1980) investigated the inter-relationships among locus of control, academic achievement, test anxiety, and level of aspiration. Locus of control, socioeconomic status, academic achievement, anxiety, and level of aspiration were assessed by a closed-ended questionnaire. It was found that students with internal locus of control had less anxiety, higher academic achievement and higher level of aspiration than external students. "External students who feel that they can not change exents, tend to be worrisome, anxious, and to have low expectations for future success. Thus, perception of locus of control appears to be a powerful variable related to the above-mentioned psychological variables and academic achievement even when socioeconomic status is controlled" (Bar-Tal, Kfir, Bar-Zohar, and Chen, 1980, p.58).

The finding of significant positive relationship between test anxiety and external locus of control in the study of Bar-Tal and his colleagues is consistent with the previous finding of Watson (1967). A group of 648 university

students who scored in the external direction on the Locus of Control Scale (Rotter, Seeman, and Liverant, 1962) were given the AAT and the Manifest Anxiety Scale (Taylor, 1953). External locus of control was found to be correlating positively with both test anxitey and manifest anxiety. External locus of control and test anxiety correlations were positive for the AAT Debilitating Scale, and negative for the AAT Facilitating Scale.

Similarly, Ross (1978) found moderately positive correlations between external locus of control and test anxiety. However, the correlations of the Worry subscale of the TAI were higher than those for the Emotionality subscale of the TAI (in Spielberger, 1980).

As another personality variable defensiveness was correlated with test anxiety. Sarason (1960) used lie scales which were found helpful in detecting the use of defensiveness in taking the anxiety scale, as a measure reflecting the defensiveness of an individual. He obtained moderately negative correlations between *test anxiety and defensiveness*.

The Defensiveness Scale for Children (DSC) was developed by Ruebush (1960) to measure the tendency to deny experiences of negative feelings, such as inadequacy, hostility, and anxiety. Highly significant negative correlations were obtained between the DSC and test anxiety (in Öner, 1971).

Öner (1971) observed that the children who scored high on the DSC showed a tendency to have very low test anxiety (the TASC) scores.

The negative relationship between test anxiety and defensiveness was verified by a recent study of Eaton (1980).

He used the TASC and a lie scale (LSC) scores obtained from 670 third and fourth grade students participated in the Hill and Sarason (1966) study. Significantly negative correlations were obtained between the TASC and the TSC scores. Eaton proposed that the causal linkage between test anxiety and defensiveness was probably related more to the impact of defensiveness on test anxiety than the impact of test anxiety on defensiveness.

In short, the studies on the relationships between test anxiety and several personality variables indicated that high test-anxious subjects tend to be more occupied with cognitions of helplessness, have lower self-concept and self-esteem, are less curious, more defensive, and more controlled by external locus than their low test-anxious counterparts.

Test anxiety and sex differences. Both the earlier (Sarason, 1960; and Öner, 1971) and the recent (Doyal and Forsyth, 1972; and Spielberger, 1980) studies in test anxiety literature indicate that there are significant sex differences in levels of test anxiety. In all studies including cross-cultural research females have been found to have higher scores on test anxiety measures than males (Schwarzer and Kim, 1984). The same tendency was observed for Turkish subjects (Öner, 1977).

Sex differences were found in defensiveness as well as in test anxiety. It was observed that defensiveness scores of males were higher than females (Sarason, 1960; and Öner, 1971).

Sarason (1960) observed the sex differences in test anxiety and defensiveness and offered that females had higher levels of anxiety, probably because they were less defensive than males in the admission of anxious feelings. This may

further be explained by sex-role identification. Based on cultural differences boys are taught and expected not to admit feelings of weaknesses while girls can easily admit these feelings (Öner, 1971).

Measurement of Test Anxiety

Several measurement techniques have been developed to assess test anxiety levels of individuals. Here, the major self-report, paper-pen instruments of test anxiety are briefly reviewed.

The first known test anxiety scale was developed by Brown (1938) at the University of Chicago. This scale aimed to identify the test-anxious students using items that tapped subjective feelings such as nervousness, irritability, and worry about examinations (in Spielberger, et.al., 1979).

The most widely used questionnaire on test anxitey is the *Test Anxiety Questionnaire* (TAQ) developed by Handler and S.Sarason in 1952. The original form of the scale consisted of 42 questions. Later it was revised and the number of items was reduced to 37. These items deal with reactions to facing or taking a course examination. The respondent indicates his reaction by marking a point on a line which is 15 centimeters long, and anchored at the ends, and in the middle. The TAQ can be administered both individually and in group.

The scoring of responses of this questionnaire was somewhat time-consuming and tedious. Therefore, some authors developed new scoring methods.

In 1958, Sarason revised tha TAQ as a 17-item true/false version. The TAQ has a split-half reliability of .91

and a test-retest reliability of .82 over a 6-week period (in Tryon, 1980).

Alpert and Haber criticized the construction of the TAQ, because, they claimed, it could only assess the extreme cases; the highly debilitated anxious individuals or the nonanxious individuals. They stated that the TAQ was unable to identify the individuals whose performances were facilitated by test-taking stress. In order to identify these people Alpert and Haber developed a new test called the *Achievement Anxiety Test* (AAT) in 1960. The AAT consisted of two independent inventories, each with a 5-point response scale. The total AAT consisted of 28 items; a 10-item Debilitating Scale designed to measure the interfering effects of anxiety on test performance, and a 9-item Facilitating Scale designed to measure the positive effects of anxiety. The remaining 9 items are neutral, buffer items. The test-retest reliabilities of the Debilitating Scale were .87 over 10 weeks, and .87 over 8 months. For the Facilitating Scale these reliability coefficients were .83 over 10 weeks, and .75 over 8 months. The Debilitating and Facilitating scales correlate negatively and significantly (r=-.37). The TAQ-AAT correlations are moderately positive (r=.64) for the Debilitating Scale, and negative (r=-.40) for the Facilitating Scale (in Tryon, 1980).

The distinctive nature of the AAT comes from its two characteristics. First, the subscales can be used separately as well as together. Second, the statements of the AAT were selected from a pool of items designed to predict students' GPAs. Therefore, the AAT is basically related to academic performance and it is not suitable to be used in studies that investigate the academic performance-anxiety relationship, because the test itself is based on the assumption that such a relationship already exists (in Levitt, 1967).

The most popular test anxiety instrument has been the *Test Anxiety Scale* (TAS), first developed by I.G.Sarason in 1958. The original form consisted of 21 true/false items which had been rewritten from the TAQ. The original TAS underwent a number of revisions. The first revision was made by I.G.Sarason and Ganzer in 1962. It consisted of 16 items, and correlated highly (r=.93) with the current true/false 37 item form. This current form of the TAS was presented by I.G.Sarason in 1972. The items of the scale were designed to tap the irrelevant thinking processes and the autonomic arousal aspect of anxiety evoked in different kinds of testing situations (in Tryon, 1980).

The *Test Anxiety Scale for Children* (TASC) was developed by S.Sarason and his colleagues at Yale University in 1960. This scale consisted of 30 yes/no items and can be used with children in first through sixth grade. S.Sarason (1960) suggests that the questions should be read to children by the experimenter during test administration. S.Sarason and his Yale team also developed an adult form of this test which could succesfully be used with college students.

In 1967, Liebert and Morris developed a questionnaire which was called the *Worry-Emotionality Questionnaire* (WEQ). This scale consisted of 5 emotionality and 5 worry items which were modified from the TAQ. Items are based on a 5-point scale. In 1970, Liebert and Morris reported test-retest reliabilities of .83 for Emotionality, and .69 for Worry when given to college students; .76 for Emotionality, and .68 for Worry when given to highschool seniors (in Tryon, 1980). The revised WEQ was based on item-factor analyses and designed to allow for more discriminative measures of the two components, and a better test of the state-trait relationships (Morris, et.al., 1983).

Another test anxiety scale was developed by Suinn, called the *Suinn Test Anxiety Behavior Scale* (STABS) in 1969. The STABS consisted of 50 items which describe behavioural situations that arouse test anxiety. It was specifically designed to assess the effectiveness of behavioural therapy in the treatment of test anxiety. The items of the STABS can also be used in constructing the systematic desensitization hierarchies. Subjects respond to the STABS on a 5-point scale. It is suitable for both group and individual use. It has norms available for college students, and adult non-students (in Tryon, 1980; and in Sweetland and Keyser, 1984).

The *Test Anxiety Profile* (TAP) was developed by Oetling and Cole (in Sweetland and Keyser, 1984)[1]. It is a 77-item scale designed to assess a person's feelings and thoughts in regard to six academic testing situations, such as multiple-choice exams, mathematics exams, essay exams, unannounced tests, talking in front of a class, and tests with time limits. Each test item consists of a pair of bipolar adjectives separated by a 7-point Likert scale. The subject rates himself on every pair of adjectives in each of the testing situation. Two anxiety scores are derived for each testing situation: Feelings of Anxiety (FA), and Thought In-fer-ence (TI). The TAP is suitable for use with highschool and college students.

In 1972 Osterhouse developed a scale called the *Inventory of Test Anxiety* (ITA). It is a measure of emotionality and worry, and consisted of 16 items developed using items from the WEQ, and other test anxiety inventories. Some items were written by Osterhouse himself. All items refer

---

[1]The publication date of this test is not given in the source.

to classroom exams only, and subjects respond to each one on a 5-point scale. Split-half reliability for a true/false version of the scale over a 7-week period is .92. Test-retest reliabilities for the Emotionality and the Worry scales over 7 weeks are .68 and .72, respectively (in Tryon, 1980).

The most recent test anxiety scale was developed by Spielberger in 1980. It is called the *Test Anxiety Inventory* (TAI). Since this inventory is also the subject of the present study, it is described in detail.

Description of the Test Anxiety Inventory. The Test Anxiety Inventory (TAI) is a self-report psychometric scale to measure individual differences in test anxiety as a situation-specific personality trait.

The TAI was developed by Spielberger through a five-year research program at the University of South Florida. The studies on the development of the TAI began in 1974 and were completed by the end of 1979. The preliminary handbook of the TAI was published in 1980.

In its preliminary form the TAI consisted of 32 items. Twenty six of these items were revised from Sarason's Test Anxiety Scale, three were adapted from Spielberger's A-State scale of the STAI, one was adapted from Osterhouse's Inventory of Test Anxiety, and one item was written by Spielberger himself.

The Preliminary Form went through several revisions and modifications according to the results of factor analyses. Finally, a 20-item scale was obtained as the last form. An example of the original form of the TAI is presented in Appendix A.

The TAI is composed of two subscales, called "Worry" and "Emotionality" both of which were determined by factor loadings. There are four additional items with moderately high loading on both factors, and these are taken into consideration for the computation of the TAI Total score. Thus the test is consisted of eight Worry, eight Emotionality and four Total, adding up to 20 items.

Although the TAI was basically developed to measure the test anxiety levels of highschool and college students, it was successfully used with Junior highschool students as well (Spielberger, 1980).

Administration. The TAI can be administered individually or in group. There is no time limit, but most students complete answering it in eight to 10 minutes.

Spielberger (1980) notes that while administering the inventory, the examiner should avoid the term "anxiety" and use the word "attitude" instead, as it is printed on the test form.

Respondents are asked to report how frequently they experience specific symptoms of anxiety before, during, and after examinations. They are expected to blacken the appropriate circle to the right of each item on the test form, and are encouraged to respond to all the items without skipping any.

During the administration of the TAI, the examiner should read the direction aloud, while students follow it from the test pages silently. The test administrator's response to students' questions during administration should be supportive, but noncommittal.

Scoring. The TAI is a four-point scale inventory. The four point choices are (1) almost never, (2) sometimes, (3) often, and (4) almost always. The scoring weights for each item are one through four. Only the first item is scored in the reverse direction. Since the responses are weighted from one to four, the minimum TAI Total score can be as low as 20, while the maximum as high as 80. Similarly, the possible minimum and maximum scores for both the Worry and the Emotionality subscales are eight and 32, respectively.

For hand scoring, there are templates available. To score each scale, the appropriate template is placed on the test form, and the weight value printed on the test form, for the response to each item is simply added. The test form is also suitable for computer scoring.

When a student does not respond to one or more items, it is possible to compute a prorated TAI Total score. For this, first, the mean of the items is determined, then, this mean is multiplied by 20; and the product is rounded to the next higher whole number. If a student omits answering more than two items of the inventory or more than one item from either subscale, the validity of the scores is questionable (Spielberger, 1980).

Reliability. The stability of responses to the TAI items were determined by test-retest reliability coefficients. These reliability coefficients were .80 for shorter periods (two weeks), and .62 for longer periods (six months).

The alpha coefficients computed by a generalized form of the Kuder-Richardson Formula 20 were .92 and higher for both females and males, indicating high internal consistency of the scale.

Further evidence of internal consistency was provided by item-remainder correlations. The median item-remainder correlations for the TAI Total Scale ranged from .61 to .69; for the TAI Worry subscale (TAI/W), the medians ranged from .58 to .72; and for the TAI Emotionality subscale (TAI/E) they ranged from .57 to .74 (Spielberger, 1980).

Validity. The concurrent validity of the TAI was based on the correlations of the TAI scores with different test anxiety measures. The correlations of the TAI with Sarason's (1978) "Test Anxiety Scale" (TAS) were found to be .82 for males, and .83 for females. The correlations with Liebert and Morris's (1967) "Worry and Emotionality Questionnaire" (WEQ) ranged from .69 to .85.

As evidence of the construct validity of the scale, the correlations of the TAI with the STAI A-Trait and the STAI A-State were computed, and found to be moderately positive, but generally lower than the correlations with other test anxiety measures. The TAI-STAI A-Trait correlations ranged from .41 to .54, and the TAI-STAI A-State correlations ranged from .28 to .67. Higher correlations were obtained between the TAI and the STAI A-State scale, when special directions were used in test administration. The students were asked to respond to the STAI A-State scale according to how they would feel just before taking an exam or during an exam. This kind of administration of the STAI A-State was called "Exam A-State." Correlations obtained from the Exam A-State and the TAI ranged from .61 to .86 and provided data for the construct validity of the latter scale (Spielberger, 1980).

Further support for the construct validity of the scale was obtained from the correlations of the TAI with measures of study skills, intelligence, and academic

achievement which were generally low-to-moderate and negative.

Negative correlations were obtained between GPA and the TAI/W subscale. They were consistently higher than those with the Total TAI scale, whereas the TAI/E subscale and GPA essentially did not correlate.

CHAPTER III

METHOD

METHOD

The purpose of the present study is twofold, and therefore it consists of two parts. The first part is on the transliteral equivalence, and the second part is on the reliability of the Turkish form of the Test Anxiety Inventory (TAI-T). These parts will be presented separately in this chapter.

The Transliteral Equivalence Study

The Translation Procedure

The translation of the Original English TAI into Turkish was previously done by Öner. In this study, the validity of the translation was tested by use of a back-translation technique. For this purpose, two American (a male and a female) teachers who know Turkish well and live in Turkey were given the Turkish translation of the TAI to be translated back into English. Their translations were then compared with the original TAI and revisions were made. After several revisions and a modification of the order of items for a better face validity, the final revision was obtained.

Sample

One hundred and sixty four bilingual Turkish university and

highschool senior students were selected from two Istanbul schools (Boğaziçi University and Robert College Highschool). The language of instruction is English in both schools. After one year of preparatory English the majority of courses in highschool, and almost all courses in this particular university are taught in English. The English language background of the highschool students is similar to each other. They all had at least six years of education in English. However, the English background of the university group is variable. Some learned the language earlier and attended a highschool where instruction was in English, but some had just one year of preparatory English before their regular university education. Therefore, the English proficiency of the highschool group is high, while that of the university group is variable.

Though the sample was composed of both male and female students, sex could not be equally represented because of limitations in the selection procedure. Hence, the natural sex distribution of the groups which formed the sample was kept.

The exact composition of the sample is shown in Table 1.

Table 1: Composition of the Sample in the Transliteral Equivalence Study

	Boğaziçi University	Robert College Highschool	Total
males (N)	16	50	66
females (N)	44	54	98
total (N)	60	104	164

Instrument

The Spielberger Test Anxiety Inventory (TAI) was the main instrument. In addition to this original English form of the inventory, three experimental versions were developed. Altogether there were four forms which were English, Turkish, and two mixed-language forms. The first form was the original English scale, and called "Form A." The second form was the translated Turkish scale, and called "Form B." The remaining two forms, "Form C" and "Form D," were split-language scales. Half (or ten) of the items which were selected randomly were in English, and the other half were in Turkish. The Turkish items in one split-language form (e.g., Form C) were presented in English in the other split-language form (Form D).

Procedure

The four forms (A, B, C, and D) of the test were administered to four different groups in a counter-balancing design. Each group responded to the inventory twice in a different language each time, and within an interval of approximately two weeks. In this way, each subject responded to every item both in English and in Turkish. The order of presentation of four experimental forms can be seen in detail in Table 2.

The researcher was not able to administer every test because of time limitations, but instead trained the examiners who were school counselors, graduate counseling students, and faculty members in education and psychology. All of them were skilled in administering psychological tests.

Table 2: The Order of Presentation of Four Experimental Forms

Group	Sex		First Testing	Second Testing
	male (N)	female (N)		
I	21	20	A	B
II	19	23	B	A
III	15	25	C	D
IV	11	30	D	C

Expectations

The criterion for transliteral equivalence was taken to be the similarity of scores obtained from the Turkish and the English forms of the TAI. If the items are translated well and are meaningful for the Turkish bilingual respondents, then there would be no significant differences among the mean scores of the four experimental groups. It was also expected that there would be high correlations between the English and the Turkish TAI scores. High correlations would indicate the equivalence of the English and Turkish scales.

Statistical Analysis

In order to test for the transliteral equivalence of the Turkish form, the following analyses were carried:

(1) Differences among the four experimental forms (A, B, C, and D) of the inventory were analysed by use of two way (sex x experimental forms) analysis of variance both for the first and the second testing, separately.

(2) Differences among the four experimental groups (A-B, B-A, C-D, and D-C) were analysed by use of two way (sex x experimental groups) analysis of variance.

(3) Difference between single (A and B) and split language (C and D) forms was analysed by use of two way (sex x language forms) analysis of variance.

(4) Difference between the English and the Turkish language items (for Turkish, Form B and half of C and D; for English, Form A and half of C and D) was analysed by use of correlated sample t-test.

(5) The relationships between the first and the second test scores were determined by use of the Pearson Product Moment correlation technique for each of the experimental group.

(6) The internal consistency of the translated Turkish form of the TAI was found by use of alpha coefficients, which were computed by a revised form of the Kuder Richardson Formula 20. Item-total and item-remainder correlations[2] of the English and the Turkish forms were also computed for the TAI Total, the TAI/Worry (TAI/W), and the TAI/ Emotionality (TAI/E) scales separately.

The Reliability Study

This study is a replication of the study carried out by Spielberger (1980). Therefore, similar procedures were used.

Sample

A total of 1031 students constituted the sample of this part of the study. These students were drawn from Istanbul schools

---

[2]The formulas used in computing the item-total, and item-remainder correlations are given in Appendix C.

representing different levels of education, namely, university, senior highschool, and junior highschool. Students were selected from different schools in different districts of Istanbul so that low and high SES levels could be represented. Tables 3 and 4 give a detailed description of the sample. Table 4 also shows the drop in the number of students on second testing.

Table 3: A Description of the Reliability Sample

Level of Education	Name of School and SES	
	Low	High
Preparatory	-	Robert College H.
Junior Highschool	İbrahim Turhan L.	Şişli Terakki L.
Senior Highschool	İbrahim Turhan L.	Şişli Terakki L. Kadıköy Anadolu L.
University		
Liberal Arts and Sciences		İstanbul University
Education		Boğaziçi University
Medicine		Çapa University

Table 4: Schools and Subjects Constituting the Reliability Sample

Level of Education	Name of School	Number of Subjects	
		First Testing	Second Testing
Junior Highschool	Robert College H.	109	107
	Şişli Terakki L.	114	102
	İbrahim Turhan L.	204	182
	Total	427	391
Senior Highschool	Şişli Terakki L.	115	45
	Kadıköy Anadolu L.	41	34
	İbrahim Turhan L.	175	156
	Total	331	235
University	L.Arts and S.	108	85
	Education	102	46
	Medicine	63	40
	Total	273	171
	TOTAL	1031	797

Instrument

The newly developed Turkish version of the Test Anxiety Inventory (TAI-T) was used in this part of the study. An example of the test form of the TAI-T is given in Appendix B.

Procedure

In order to test the stability of the TAI-T scores the students were given the inventory twice with varying time intervals. These intervals were the same day, one day, one week, two weeks, and three weeks. Longer time intervals could not be taken because the academic year ended before the desired intervals could be attained.

The tests were administered by the researcher, the the school counselors, and in some necessary cases, by the graduate counseling students, and faculty members in education and psychology. Before test administration all examiners were trained by the researcher.

Statistical Analysis

The analyses were based on computations for the stability of the TAI-T scores, and the internal consistency of the TAI Total, the TAI/W and the TAI/E scales.

The stability of the scores was found by test-retest correlations. For this, the Pearson Product Moment correlation coefficients were computed for each time interval, separately.

The internal consistency of the TAI-T was tested by use of alpha coefficients, item-total, and item-remainder correlations. These computations were carried for the TAI Total, the TAI/W, and the TAI/E, independently.

CHAPTER IV

FINDINGS AND DISCUSSION

FINDINGS AND DISCUSSION

In this chapter data are presented in two sections. In the first section the findings of the transliteral equivalence study are presented and discussed. In the second session the data on the reliability of the Turkish form of the TAI are presented and discussed.

The Transliteral Equivalence Study

This particular study was aimed to demonstrate that the translated Turkish form is not different from the English form of the TAI. Therefore, various comparisons are conducted among the experimental versions of the TAI focusing on language differences. The reliability of the Turkish and English TAI scores of the same individuals were also computed and compared.

For an overall picture the means and standard deviations obtained from four experimental groups (A-B, B-A, C-D, and D-C) at first and second testing for males and females are presented in Table 5.

Table 5: Means and Standard Deviations Obtained from Experimental Groups in First and Second Testing

Experimental Groups	Experimental Forms	Males			Females			Total		
		N	M	SD	N	M	SD	N	M	SD
I	A	21	33.62	9.49	20	43.45	10.73	41	38.41	11.16
	B	21	31.62	8.90	20	40.40	11.15	41	35.90	10.88
II	B	19	31.05	10.67	23	37.17	9.76	42	34.40	10.52
	A	19	29.89	13.29	23	35.04	10.45	42	32.71	11.95
III	C	15	38	11.93	25	36.40	9.92	40	37	10.59
	D	15	37	11.59	25	36.92	11.34	40	36.95	11.28
IV	D	11	38.73	13.08	30	37.90	10.04	41	38.12	10.77
	C	11	37.09	11.61	30	36.27	10.55	41	36.49	10.70

This table indicates that the means obtained from first testing are consistently higher than the means obtained from second testing. It is also seen that in groups where the number of females and males are close (as pure language forms, i.e., Forms A and B) the female subjects tend to have higher scores than the male subjects.

Comparisons Among Experimental Groups, Forms and Between Languages

The results of the two way (sex x experimental forms) analysis of variance yielded no significant differences among experimental forms (A, B, C, and D) in either the first or the second testing scores (see Tables 6 and 7).

Table 6: Analysis of Variance for Four Experimental Forms at First Testing and Sex

Source of Variation	SS	DF	MS	F
Sex(A)	21.21	1	21.21	3.57
Forms(B)	23.49	3	7.83	1.32
A x B	42.68	3	14.23	2.39
Error	-	156	5.94	-

Table 7: Analysis of Variance for Four Experimental Forms at Second Testing and Sex

Source of Variation	SS	DF	MS	F
Sex(A)	19.53	1	19.53	3.03
Forms(B)	22.81	3	7.60	1.18
A x B	30.05	3	10.02	1.55
Error	-	156	6.45	-

The differences among experimental forms were also nonsignificant in both the first and the second testing ($F(3,156)=1.32$, and 1.18). Although the F values for sex were relatively high in these analyses ($F(1,156=3.57$, and 3.03), they did not reach significance either.

The results of the two way analysis of variance for experimental groups (A-B, B-A, C-D, and D-C) and for sex are given in Table 8. In this analysis, neither the experimental groups ($F(3,156)=1.36$), nor sex ($F(1,156)=3.72$) showed significant differences. The latter, however, fell very short of significance.

Table 8: Analysis of Variance for Four Experimental Groups and Sex

Source of Variation	SS	DF	MS	F
Sex(A)	21.89	1	21.89	3.72
Groups(B)	23.98	3	7.99	1.36
A x B	37.76	3	12.59	2.14
Error	-	156	5.88	-

Another analysis of variance computation was carried out for differences among the pure-split language forms and sex (Table 9). No difference was found between the pure and the split language forms ($F(1,160)=1.55$), but sex was significant ($F(1,160)=5.37$, $p<.05$).

Table 9: Analysis of Variance for Pure and Split Language Forms and Sex

Source of Variation	SS	DF	MS	F
Sex(A)	15.62	1	15.62	5.37*
pure and split language(B)	4.50	1	4.50	1.55
A x B	10.64	1	10.64	3.66
Error	-	160	2.91	-

*$p<.05$

The difference between the English and the Turkish items, obtained by using the means of the experimental forms A and B, and including the composite means[3] for the English and

[3]The composite means for the two languages were computed by combining ten items answered in Turkish by one group with the remaining ten in Turkish by the comparable group receiving the other form in which the items were reversed. Same procedure was followed for the composite mean score of English items.

the Turkish items compiled from the mixed-language forms (Form C and D), was analysed by a correlated sample t-test. The results showed non-significant differences between the English and Turkish items. These data are presented in

Table 10: Means, Standard Deviations, and t-Values for the English and the Turkish Items of the Inventory

	Language of Items	N	M	SD	t Values
Males	English	66	34.26	11.61	1.09
	Turkish	66	33.42	10.90	
Females	English	98	37.83	10.61	0.049
	Turkish	98	37.72	10.49	
Total	English	164	36.39	11.13	0.74
	Turkish	164	35.98	10.85	

Sex differences in test anxiety scores were significant at the .01 level ($t(98,66)=2.03$) for the English items, and at the .005 level ($t(98,66)=2.53$) for the Turkish items. The female subjects scored higher than the male subjects in both languages.

The last comparison was carried out by use of correlation coefficients between the scores obtained in two testing, each time using a different language. The Product Moment correlations between the first and second testing for each experimental groups are shown in Table 11. These were .87 (A-B), and .92 (B-A) between the pure language forms, and .94 (C-D), and .85 (D-C) between the mixed language forms. All were significant beyond the .001 level.

Table 11: Correlations Between the First and Second Testing in the Experimental Groups

	A-B	B-A	C-D	D-C
N	41	42	40	41
r	.87	.92	.94	.85

Reliability of the Experimental Turkish TAI and the Original English TAI

The reliability of the English and the Turkish TAI items were determined by use of test-retest, alpha, item-total, and item-remainder correlations.

The two-week interval test-retest correlations provided data on the stability of scores obtained from the English and the Turkish items of the TAI. These correlations range between .85 and .94 as shown in Table 11.

The analysis for the internal consistency of the experimental forms were based on the scores obtained from the pure language forms, i.e., Forms A and B. For this, alpha coefficients shown in Table 12 were computed by a generalized form of the Kuder-Richardson Formula 20 for the TAI Total, TAI/W, and TAI/E scales.

It is seen from this table that the alpha coefficients of the English and the Turkish TAI are very close to each other for the total sample. The Total scale alpha values are all above .91 for both sample groups and the total group in both languages. The Emotionality subscale alphas ranged between .85 and .94. The worry subscale coefficients were, however, lower (.73 and .76 for the English and the Turkish, respectively) for the highschool sample than the university

sample (.91 and .85), as well as the total sample (.86 and .82). In all these values the two language forms did not seem to differ considerably.

Table 12: Alpha Coefficients Obtained from the English and the Turkish Forms of the TAI for the Worry, the Emotionality, and the Total TAI Scores

Level of Education	Language of Forms	N	Worry	Emotionality	Total
University	English	30	.91	.94	.96
	Turkish	30	.85	.93	.94
Highschool	English	53	.73	.87	.91
	Turkish	53	.76	.85	.92
Total	English	83	.86	.91	.94
	Turkish	83	.82	.90	.93

Further information about the internal consistency of the English and the Turkish forms of the TAI were obtained from item-total and item-remainder correlations. The item-total correlations for the TAI Total scale, and for the Worry and Emotionality subscales are presented in Tables 13 and 14[4].

[4]For the face validity of the TAI-T, the original order of items was changed. Therefore, the order of items is different in the two TAI forms (A and B). For purposes of comparison, the original order of items was adjusted according to the order of the Turkish form.

Table 13: Item-Total Correlation Coefficients for the TAI Total Scores Obtained from the English and the Turkish Forms of the TAI

Item	University N=30		Highschool N=53		Total N=83	
	English	Turkish	English	Turkish	English	Turkish
1	.47	.63	.74	.78	.65	.72
2	.71	.73	.48	.49	.63	.57
3	.74	.72	.74	.59	.74	.63
4	.66	.38	.43	.60	.59	.51
5	.65	.35	.33	.38	.52	.39
6	.77	.73	.61	.71	.68	.74
7	.78	.67	.81	.63	.78	.68
8	.86	.67	.56	.53	.71	.63
9	.74	.85	.54	.54	.68	.72
10	.74	.73	.70	.54	.73	.67
11	.88	.72	.54	.64	.68	.70
12	.68	.48	.23	.56	.52	.56
13	.70	.84	.78	.72	.71	.79
14	.89	.87	.79	.74	.86	.81
15	.77	.64	.71	.75	.76	.73
16	.79	.78	.72	.76	.78	.78
17	.70	.55	.54	.68	.67	.65
18	.87	.79	.70	.66	.80	.72
19	.71	.55	.51	.62	.62	.56
20	.82	.69	.79	.62	.79	.69
Medians	.74	.70	.65	.62	.69	.68

Table 14: Item-Total Correlation Coefficients for the Worry and the Emotionality Scores Obtained from the English and the Turkish Forms of the TAI

Item	University N=30		Highschool N=53		Total N=83	
	English	Turkish	English	Turkish	English	Turkish
Worry						
2	.71	.79	.64	.58	.69	.65
3	.70	.74	.76	.63	.73	.67
4	.83	.57	.52	.68	.72	.63
5	.80	.62	.46	.58	.65	.58
8	.87	.82	.67	.58	.77	.72
12	.80	.49	.32	.65	.62	.61
17	.79	.77	.55	.72	.73	.76
20	.87	.73	.78	.67	.81	.73
Medians	.80	.73	.59	.64	.72	.66
Emotionality						
6	.75	.75	.64	.75	.69	.77
7	.87	.79	.83	.67	.83	.75
9	.80	.88	.59	.66	.72	.80
11	.89	.80	.55	.68	.69	.75
13	.76	.91	.81	.75	.76	.83
14	.93	.81	.83	.71	.89	.77
16	.85	.82	.82	.74	.85	.79
18	.89	.80	.79	.73	.85	.76
Medians	.86	.80	.80	.72	.79	.77

The obtained item-total correlations for the English and the Turkish scales seem very similar in general. The median correlations are Just a few points lower for the TAI-T, but they are all above .62 for the Total TAI (TAI/T); and above .64 for the TAI/W; and above .72 for the TAI/E subscales.

The median correlations for the original English TAI were above .65 for the TAI/T; .59 for the TAI/W, and .79 for the TAI/E.

The corrected item-total correlations are presented in the form of item-remainder correlations. These are given in Tables 15 and 16.

Table 15: Item-Remainder Correlation Coefficients for the TAI Total Scores Obtained From the English and the Turkish Forms of the TAI

Item	University N=30		Highschool N=53		Total N=53	
	English	Turkish	English	Turkish	English	Turkish
1	.43	.60	.72	.76	.62	.70
2	.69	.71	.43	.45	.60	.54
3	.72	.70	.72	.55	.72	.60
4	.64	.34	.40	.59	.57	.48
5	.62	.31	.27	.33	.48	.34
6	.75	.71	.58	.69	.66	.72
7	.77	.64	.79	.60	.76	.66
8	.85	.65	.52	.50	.69	.60
9	.72	.84	.50	.51	.65	.71
10	.72	.70	.67	.51	.71	.64
11	.88	.69	.49	.60	.65	.66
12	.66	.46	.17	.54	.49	.54
13	.69	.82	.76	.70	.69	.77
14	.88	.86	.77	.71	.84	.79
15	.75	.60	.67	.72	.73	.70
16	.77	.75	.69	.73	.77	.76
17	.68	.51	.51	.66	.65	.62
18	.86	.78	.67	.62	.79	.69
19	.69	.52	.46	.57	.59	.53
20	.80	.66	.77	.60	.77	.67
Medians	.72	.67	.62	.60	.67	.66

Table 16: Item-Remainder Correlation Coefficients for the Worry and the Emotionality Scores Obtaines from the English and the Turkish Forms of the TAI

Item	University N=30		Highschool N=53		Total N=83	
	English	Turkish	English	Turkish	English	Turkish
Worry						
2	.65	.74	.51	.45	.61	.56
3	.64	.68	.68	.52	.66	.58
4	.80	.47	.43	.64	.67	.56
5	.76	.56	.29	.44	.57	.49
8	.85	.78	.57	.46	.72	.65
12	.76	.42	.16	.60	.55	.55
17	.75	.73	.47	.63	.68	.70
20	.84	.67	.72	.59	.77	.67
Medians	.76	.67	.49	.55	.66	.57
Emotionality						
6	.71	.71	.57	.70	.64	.72
7	.85	.76	.79	.60	.80	.70
9	.76	.87	.51	.62	.66	.77
11	.87	.76	.44	.59	.62	.69
13	.74	.89	.77	.70	.72	.80
14	.92	.78	.80	.65	.87	.73
16	.82	.78	.78	.68	.82	.74
18	.87	.77	.74	.67	.82	.71
Medians	.83	.77	.75	.66	.76	.72

It is seen from these tables that the median item-remainder correlations do not fall below .60 in any of the scales for any sample group. The median item-remainder correlations for the English and the Turkish forms of the TAI are very close to each other.

A general view at the item-total and item-remainder

correlations is that in the Total TAI scales and its subscales all of the correlations, except for items 4, 5, and 12, fall well above .50. These correlations are higher for the Emotionality, than the Worry subscale.

Discussion

The analyses of variance showed no significant differences among the experimental forms, the experimental groups, or between the pure and split language forms. The *t*-test results also revealed non-significant difference between the English and the Turkish TAI items. The English and the Turkish forms of the TAI are highly correlated.

These findings are interpreted as evidence for the adequacy of item translation, and the equivalence of the two language forms of the inventory. These findings are consistent with the expectations of the study. Thus, depending on these data it can be stated that the Turkish form of the TAI is transliterally equivalent to the original TAI and can be used with Turkish subjects for further investigations.

In addition, the reliability of the English and the Turkish forms of the TAI are determined in terms of the stability and the internal consistency of forms. The stability of scores are reflected in the high correlations (ranging from .85 to .94) between the English and the Turkish language forms. These high correlations are very close to those reported in the TAI Manual (Spielberger, 1980).

The internal consistency of both the English and the Turkish TAI is reflected in the alpha, item-total, and item-remainder correlations. The alpha coefficients are high (.91 or higher) and also close to the findings of Spielberger (1980).

The item-total and item-remainder correlations of the Turkish TAI are also similar to the correlations of the English TAI. Generally, both the item-total and the item-remainder correlations of the Turkish form are a bit lower than the English form, but in some items (i.e., items 1, 6, 9, 11, 12, and 13) the Turkish form has higher correlations.

All these indicate that the experimental Turkish form has proven to be internally consistent and stable on this particular Turkish sample.

Other important findings of this study are related to sex differences and the effect of testing order on test scores. In their raw scores females consistently scored higher than males. This is consistent with the previous findings of Sarason (1960), Öner (1971), and Spielberger (1980). Also subjects scored higher in the first testing when compared to second testing scores. This findings seems supportive of Spielberger et.al.'s (1970) suggestion on the administration of the STAI, and the data reported by Öner (1977) on trait anxiety scores.

Limitations

The limitations of this part of the study are related to the composition and the selection of the sample. (1) The sample was composed of bilingual university and highschool students, but the English background of the university group was not as homogeneous in terms of high level of proficiency as that of the highschool group. The findings would be considered more reliable, if the bilingual sample was more homogeneous. (2) Because of practical limitations in the selection of the size of the university and highschool groups, the groups could not be represented equally. The university group was smaller than the highschool group. For the same reason equal sex

distribution could not be obtained. The interpretation of the findings would, therefore, be considered tentative and not generalizable. If both the size and the sex distribution of the sample groups were equal, the findings could probably be generalized with more certainty.

The Reliability Study

This part of the investigation aims to demonstrate the stability and the internal consistency of the TAI-T. Also for the comparability of the TAI-T scores with those of the English TAI scores, the means and standard deviations of the first test scores of the subjects on the TAI-T are obtained (see Table 17).

Table 17: Means, and Standard Deviations Obtained from the TAI-T Scores of the Reliabilty Sample

		N	M	SD
UNIVERSITY	Male	130	34.54	8.22
	Female	143	37.45	9.59
	Total	273	36.06	9.06
SENIOR HIGHSCHOOL	Male	197	39.29	9.77
	Female	134	44.61	10.24
	Total	331	41.45	10.28
JUNIOR HIGHSCHOOL	Male	270	38.01	8.97
	Female	157	39.66	9.34
	Total	427	38.61	9.13
TOTAL	Male	597	37.68	9.24
	Female	434	40.46	10.12
	Total	1031	38.86	9.76

The Stability of the TAI-T

In determining the stability of the TAI-T, test-retest correlations with varying time intervals were obtained. The time intervals ranged from the same day to three weeks. About one fourth student attribution rate is observed on retesting. Thus the stability coefficients are based on three fourths of the original sample. Table 18 shows the size of the sample and test-retest correlations over the designated time intervals.

Table 18: Test-Retest Correlation Coefficients Obtained from the TAI-T

Time Interval	Male		Female		Total	
	N	r	N	r	N	r
Same day	-	-	-	-	28	.91
One day	-	-	-	-	50	.93
One week	162	.89	110	.91	272	.90
Two weeks	153	.76	135	.85	288	.81
Three weeks	86	.77	70	.62	156	.72

It is seen from this table that with the passage of time, some decrements in the stability coefficients occured. The correlation was .91 for the same day interval, but it reduced to .72 for the three weeks interval.

The Internal Consistency of the TAI-T

The internal consistency of the TAI-T was determined by two different kinds of analyses, namely, the alpha coefficients as computed by Kuder-Richardson Formula 20, and the item-total (and item-remainder) correlations. Table 19 presents the alpha coefficients obtained from several subject groups

for the TAI/T, the TAI/W, and the TAI/E scores.

Table 19: Alpha Coefficients Obtained from Different Sample Groups for the Worry and the Emotionality Subscales, and the Total TAI Scale

Sample	N	Worry	Emotionality	Total
Grand Total	1031	.74	.79	.87
University Total	273	.69	.84	.89
İstanbul U.	108	.68	.82	.87
Çapa U.	102	.71	.80	.86
Boğaziçi U.	63	.69	.68	.90
Senior Highschool	331	.76	.79	.87
İbrahim T.L.	175	.75	.75	.85
Şişli Terakki	115	.76	.80	.88
Kadıköy A.L.	41	.79	.89	.92
Junior Highschool Total	427	.73	.75	.84
İbrahim T.L.	204	.76	.71	.84
Şişli Terakki	114	.64	.75	.82
Robert College	109	.67	.83	.86

As it can be read from this table, the alpha coefficients increased with the age of subjects, and all were quite high. They ranged from .82 to .92 for the TAI/T scale, and from .64 to .79 for the TAI/W, and from .71 to .89 for the TAI/E subscales.

The item-total correlations for the TAI/T scale, and for the TAI/W and the TAI/E subscales are presented in Tables 20 and 21.

Table 20: Item-Total Correlation Coefficients for the TAI Total Scores Obtained from the TAI-T

Item	University N=273	Senior Highschool N=331	Junior Highschool N=427	Total N=1031
1	.64	.49	.31	.46
2	.40	.49	.45	.45
3	.47	.51	.51	.52
4	.37	.43	.39	.41
5	.42	.43	.40	.41
6	.65	.64	.55	.61
7	.65	.58	.59	.60
8	.53	.58	.56	.56
9	.64	.57	.55	.59
10	.55	.49	.46	.49
11	.57	.47	.44	.50
12	.45	.58	.50	.53
13	.71	.61	.48	.58
14	.63	.62	.62	.63
15	.59	.44	.43	.48
16	.65	.59	.55	.59
17	.58	.55	.57	.58
18	.61	.57	.54	.57
19	.54	.51	.47	.52
20	.64	.68	.63	.66
Medians	.58	.56	.50	.54

Table 21: Item-Total Correlation Coefficients for the Worry and the Emotionality Scores Obtained from the TAI-T

Item	University N=273	Senior Highschool N=331	Junior Highschool N=427	Total N=1031
Worry				
2	.58	.61	.53	.57
3	.52	.55	.57	.56
4	.48	.52	.56	.55
5	.54	.50	.51	.50
8	.62	.68	.65	.65
12	.51	.66	.58	.61
17	.65	.65	.64	.66
20	.61	.73	.64	.68
Medians	.56	.63	.57	.59
Emotionality				
6	.69	.66	.59	.65
7	.73	.67	.65	.67
9	.69	.65	.60	.64
11	.66	.56	.54	.58
13	.74	.64	.60	.65
14	.63	.69	.61	.65
16	.70	.63	.66	.66
18	.69	.62	.62	.63
Medians	.69	.64	.60	.65

Just like in the alpha values, the item-total correlations increase with the age of the sample group. For instance, while the median TAI Total correlations are above .56 and .58 for the senior highschool and the university groups, it is .50 for the junior highschool group. Five items fall below .45 (range from .31 to .44) in this group, while only three such items are noted in each of the older groups. All item-total correlations for the Worry and the

Emotionality subscales are above .50 in all three groups, however, with one exception (.48 for item 4 in the university group). Correlations for items 4 and 5 are consistently lower than other items (range from .37 to .42).

In Tables 22 and 23 item-remainder correlations are presented for the TAI/T scale, and for the TAI/W, and the TAI/E subscales.

Table 22: Item-Remainder Correlation Coefficients for the TAI Total Scores Obtained from the TAI-T

Item	University N=273	Senior Highschool N=331	Junior Highschool N=427	Total N=1031
1	.61	.45	.24	.42
2	.33	.44	.40	.40
3	.43	.47	.47	.47
4	.32	.37	.32	.35
5	.37	.38	.33	.35
6	.61	.61	.51	.57
7	.61	.54	.54	.55
8	.48	.54	.50	.51
9	.61	.53	.50	.55
10	.51	.44	.41	.44
11	.52	.42	.38	.45
12	.41	.54	.45	.49
13	.69	.57	.43	.54
14	.60	.58	.58	.59
15	.53	.36	.34	.40
16	.61	.55	.50	.55
17	.54	.50	.52	.53
18	.57	.52	.49	.52
19	.49	.46	.41	.46
20	.62	.65	.58	.62
Medians	.53	.51	.46	.50

Table 23: Item-Remainder Correlation Coefficients for the Worry and the Emotionality Scores Obtained from the TAI-T

Item	University N=273	Senior Highschool N=331	Junior Highschool N=427	Total N=1031
Worry				
2	.44	.51	.42	.46
3	.39	.45	.47	.46
4	.36	.40	.44	.43
5	.39	.37	.38	.37
8	.49	.59	.54	.55
12	.43	.58	.48	.53
17	.54	.55	.53	.56
20	.50	.66	.53	.59
Medians	.43	.53	.47	.49
Emotionality				
6	.62	.58	.50	.57
7	.66	.59	.56	.59
9	.63	.57	.50	.56
11	.58	.44	.41	.47
13	.68	.57	.50	.57
14	.56	.61	.51	.57
16	.64	.53	.57	.58
18	.62	.52	.50	.53
Medians	.62	.57	.50	.57

Since the item-remainder correlations are corrected for artifact, they are slightly lower than the item-total correlations. Again, these values are lower for the junior highschool group than those for the senior highschool and university groups. The median item-remainder correlations of the TAI Total scale were .46 (range from .24 to .58) for the junior highschool group, .51 (range from .36 to .65) for the senior highschool group, and .53 (range from .32 to .69) for the university group. The median item-remainder correlations

were above .36 for the TAI/W, and above .41 for the TAI/E subscales.

The Emotionality subscale yielded somewhat higher item-total (and item-remainder) correlations than the Worry subscale. This finding is consistent with the data obtained from the first part of the present study.

Discussion

Test-retest correlations obtained from the Turkish subjects of this study are similar to those reported by Spielberger (1980). In the TAI Manual, the stability coefficient of .80 was reported for a two-week interval. In the present study, .81 was found for the same time. For three-weeks interval a correlation of .80 was reported in the handbook. In this study it was .72 for the same interval. These are considered as evidence for the high stability of the TAI-T.

The data on the internal consistency of the TAI-T were based on alpha, item-total, and item-remainder correlations. Although the obtained correlations were not as high as those reported by Spielberger (1980) they were still considered satisfactory. For example, the alpha coefficients for the TAI/T ranged from .92 to .96 in the original handbook. The obtained alpha coefficients of the present study ranged from .82 to .92. The median item-remainder correlations for the TAI Total were reported to be ranging from .61 to .69 in the TAI Manual, while they were found to be ranging from .46 to .53 in the present study.

The internal consistency of the Emotionality subscale of the TAI-T was much higher than that of the Worry subscale. The same finding was also obtained in the first part of the present study. Interestingly, Spielberger (1980) reported

similar findings. Though the differences were not as high as these of the present study, he also observed that the Emotionality subscale had higher alpha coefficients than the Worry subscale.

Parallel to the findings of the first study, and to those of Spielberger (1980), the female subjects in this sample group scored consistently higher than the male subjects (see Tables 5 and 17).

In sum, the TAI-T was found to be a stable and internally consistent scale which could be used in research with Turkish subjects. It should be noted, however, that the alpha, item-total and item-remainder correlations were somewhat lower for the younger sample groups. It was generally observed that the older the sample group the higher the internal consistency of the TAI-T. Further refinement might be seemed necessary before its use with younger Turkish subjects.

Limitations

The major limitation of the present study was the selection of the sample. For several reasons, this sample can not be considered a good representative of the Turkish student population.

First of all, the subjects were selected only from Istanbul schools. Second, sampling could not be randomized. To provide some heterogeneity, the subjects were taken from schools with low and high SES levels. These schools were the ones in which officials expressed cooperation in test administration. Because of various difficulties different age groups and sexes were not equally represented in the sample.

Another limitation comes from the shortage of time. The longest time-interval of the stability coefficients in this study was three weeks, because the academic year had ended before longer desired time intervals could be actualized.

Conclusions and Recommendations

The concern of the present research is the development of the Turkish adaptation of the TAI. For this purpose, the transliteral equivalence and the reliability of the Turkish form of the TAI are studied.

The findings on the transliteral equivalence part of the study supported the expectations of no differences among experimental groups, experimental forms, and between languages. High correlations are obtained between the first and the second testing scores in the experimental groups as evidence for the stability of the experimental forms. The obtained alpha, item-total, and item-remainder correlations of the Turkish TAI are high and very close to those of the English form. All these are interpreted to indicate that the adapted Turkish TAI is transliterally equivalent to the original English TAI.

The findings related to the reliability of the Turkish version of the TAI showed high stability scores. Though they are not as high as these obtained from the first part of the study, the alpha, item-total, and item-remainder correlations are still satisfactory, and generally higher than those obtained from some other versions, such as the Hindi, Hungarian, and Italian forms (in Comunian, 1985; and Sipos, Sipos, and Spielberger, 1985).

In both parts of the present research lower internal

consistency data are obtained from the younger sample groups. Similar to some cross-cultural studies (Comunian, 1985; and Sipos, Sipos, and Spielberger, 1985), it is also observed that the internal consistency of the Emotionality subscale is higher than that of the Worry subscale. Although the alpha coefficients for the Worry subscale of the TAI-T are not as high as those reported in the TAI Manual, they are higher than the values obtained on the Hungarian (.65 to .69) or the Italian (.62 to .74) versions of the TAI (Comunian, 1985; and Sipos, Sipos, and Spielberger, 1985).

When the normative data are taken into consideration females are found to score higher on the TAI-T than males. This is parallel to the data generally reported in test anxiety literature. Other normative data are related to the cross-cultural comparison of the mean score of the Turkish subjects on the TAI-T. In relation to the other adaptations of TAI, the mean of the Turkish samples fall in intermediate position, and are very similar to the mean of the American and German samples (Schwarzer, and Kim, 1984).

These cross-cultural comparisons provide supporting evidence for the reliability as well as the criterion validity of the newly developed Turkish form.

The present research is limited with sampling and time variables. Therefore, cross-validation with larger and different student samples and retesting over longer periods are recommended.

In this study the internal consistency of the TAI-T is found somewhat high in the older groups, but low in the younger groups. For this reason, the internal consistency of the TAI-T for younger students should be further investigated.

Although the general finding of the present research indicated that the TAI-T is a promising instrument, the scale needs to be validated before its use with Turkish students. Further research on the validation of the TAI-T should include the factor analytic structure of the scale, as well as its predictive power.

REFERENCES

Bar-Tal,D., Kfir,D., Bar-Zohar,Y., and Chen,Y. The Relationship Between Locus of Control and Academic Achievement, Anxiety, and Level of Aspiration. *British Journal of Educational Psychology,* 1980, *50*, 53-60

Becker,P. Towards a Process Analysis of Test Anxiety: Some Theoretical and Methodological Observations. In Schwarzer,R., Ploeg,H.M., van der, and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.1), Lisse: Swets and Zeitlinger B.V., 1982.

Benjamin,M., McKeachie,W.J., Lin,Yi-Guang, and Holinger,D.P. Test Anxiety: Deficits in Information Processing, *Journal of Educational Psychology,* 1981, *73*(6), 816-824.

Brown,S.D., and Nelson,T.L. Beyond the Uniformity Myth: A Comparison of Academically Successful and Unsuccessful Test-Anxious College Students. *Journal of Counseling Psychology,* 1963, *30*(3), 367-374.

Bruch,M.A., Juster,H.R., and Kaflowitz,N.G. Relationships of Cognitive Components of Test Anxiety to Test Performance: Implications for Assessment and Treatment. *Journal of Counseling Psychology,* 1983, *30*(4), 527-536.

Comunian,A.L. The Development and Validation of the Italian Form of the Test Anxiety Inventory. In Ploeg,H.M., van der, Schwarzer,R., and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.4), Lisse: Swets and Zeitlinger B.V., 1985.

Covington,M.V. Test Anxiety: Causes and Effects over Time. In Ploeg,H.M., van der Schwarzer,R., and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.4), Lisse: Swets and Zeitlinger B.V., 1985.

Culler,R.E., and Holahan,C.J. Test Anjiety and Academic Performance: The Effects of Study-Related Behaviors. *Journal of Educational Psychology,* 1980, *72*(1), 16-20.

Defares,P.B., Grossman,P., and Swart,H.C.G., de. Test Anxiety, Cognitive Primitivation and Hyperventilation. In Ploeg,H.M., van der, Schwarzer,R., and Spielberger, C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.2), Lisse: Swets and Zeitlinger B.V., 1983.

Deffenbacher,J.L. Worry, Emotionality, and Task-Generated Interference in Test Anxiety: An Empirical Test of Attentional Theory. *Journal of Educational Psychology,* 1978, *70*(2), 248-254.

Doyal,G.T., and Forsyth,R.A. The Effect of Test Anxiety, Intelligence, and Sex on Children's Problem Solving Abilities, *The Journal of Experimental Education,* 1972, *41*(2), 23-26.

Eaton,W.O. Cross-Lagged Analysis of Children's Test Anxiety and Defensiveness. *Journal of Educational Psychology,* 1980, *72*(3), 404-407.

Epstein,S. *Anxiety, Arousal, and the Self-Concept.* Paper presented at Conference on Anxiety and Stress, Oslo, June 29-July 3, 1975.

Fransson,A. On Qualitative Differences in Learning: IV-Effects of Intrinsic Motivation and Extrinsic Test Anxiety on Process and Outcome. *British Journal of Educational Psychology,* 1977, *47,* 244-257.

Galassi,J.P., Frierson,H.T., Jr., and Sharer,R. Behavior of High, Moderate, and Low Test Anxious Students During an Actual Test Situation. *Journal of Consulting and Clinical Psychology,* 1981, *49*(1), 51-62.

Ganzer,V.J. Effects of Audience Presence and Test Anxiety on Learning and Retention in a Serial Learning Situation. *Journal of Personality and Social Psychology,* 1968, *8*(2), 194-199.

Garrett,H.E., and Woodworth,R.S. *Statistics in Psychology and Education.* New York: Longmans, Green and Co., 1959.

Geen,R.G. Test Anxiety, Observation, and Range of Cue Utilization. *British Journal of Social Clinical Psychology,* 1976, *15,* 253-259.

Harleston,B.W., Smith,M.G., and Arey,D. Test-Anxiety Level, Heart Rate, and Anagram Problem Solving. *Journal of Personality and Social Psychology,* 1965, *1*(6), 551-557.

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

Hodapp,V. Test Anxiety, Study Habits, and Academic Performance. In Ploeg,H.M.,van der, Schwarzer,R., and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.2), Lisse: Swets and Zeitlinger B.V., 1983.

Kirkland,K., and Hollandsworth,J.G., Jr., Effective Test Taking: Skills-Acquisition Versus Anxiety-Reduction Techniques. *Journal of Consulting and Clinical Psychology,* 1980, *48*(4), 431-439.

Lavelle,T.L., Metalsky,G.I., and Coyne,J.C. Learned Helplessness, Test Anxiety, and Acknowledgments of Contingencies. *Journal of Abnormal Psychology,* 1979, *38*(4), 381-387.

LeCompte,W.A., and Öner,N. Development of the Turkish Edition of the State Trait Anxiety Inventory. In Spielberger,C.D., and Diaz-Guerrero,R. (Eds.), *Cross-Cultural Anxiety* (Vol.1), New York: Hemisphere Publishing Corporation, 1976.

Levitt,E.E. *The Psychology of Anxiety.* Indianapolis: The Bobbs-Merrill Company Inc., 1967.

Marlett,N.J., and Watson,D. Test Anxiety and Immediate or Delayed Feedback in a Test-Like Avoidance Task. *Journal of Personality and Social Psychology,* 1968, *8*(2), 200-203.

Marso,R.N. Classroom Testing Procedures, Test Anxiety, and Achievement. *The Journal of Experimental Education,* 1970, *38*(3), 54-58.

McCoy,N. Effects of Test Anxiety on Children's Performance As a Function of Instructions and Type of Task. *Journal of Personality and Social Psychology.* 1965, *2*(5), 634-641.

Minsel,B., and Schwarzer,C. Nonlinear Relationships of Worry and Emotionality to School Achievement. In Ploeg,H.M., van der, Schwarzer,R., and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.2), Lisse: Swets and Zeilinger B.V., 1983.

Morris,L.W., Franklin,M.S., and Ponath,P. The Relationship Between Trait and State Indices of Worry and Emotionality. In Ploeg,H.M., van der, Schwarzer,R., and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.2), Lisse: Swets and Zeitlinger B.V., 1983.

Muller,J.H., and Thompson,W.B. Test Anxiety and Distinctiveness of Personal Information. In Ploeg,H.M., van der, Schwarzer,R., and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.3), Lisse: Swets and Zeitlinger, B.V., 1984.

Nottelmann,G.D., and Hill,K.T. Test Anxiety and Off-Task Behavior in Evaluative Situations. *Child Development,* 1977, *48,* 225-231.

Nunnally,J.C. *Psychometric Theory.* New York: McGraw-Hill, Inc., 1967.

Öner,N.P. *Impact of Teacher Behavior and Teaching Technique on Learning by Anxious Children.* Unpublished doctoral dissertation, Minnesota University, 1971.

Öner,N. *Durumluk-Sürekli Kaygı Envanterinin Türk Toplumunda Geçerliği.* Basılmamış doçentlik tezi, Hacettepe Üniversitesi, 1977.

Öner,N. Kaygı, Yeni Davranışların Kazanılmasında Güdüleyici, Uyumu Kolaylaştırıcı Olabilir mi? *Boğaziçi Üniversitesi Dergisi* Vols. *8-9*, (ayrı baskı), 1980-81.

Öner,N., and LeCompte,A. *Süreksiz, Durumluk/Sürekli Kaygı Envanteri El Kitabı.* İstanbul: Boğaziçi Üniversitesi Yayınları, 1983.

Phillips,B.N. Test Anxiety in School Settings: Application of the Aptitude-Treatment Interaction Concept in Research on Test Anxiety, Instruction, and Minimum Competency Testing. In Ploeg,H.M., van der, Schwarzer R., and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.4), Lisse: Swets and Zeitlinger B.V., 1985.

Ploeg,H.M., van der. Worry, Emotionality, Intelligence, and Academic Performance in Male and Female Dutch Secondary School Children. In Ploeg,H.M., van der, Schwarzer,R., and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.3), Lisse: Swets and Zeitlinger B.V., 1984.

Ploeg,H.M., van der. In Schwarzer,R. Personal Communication, 1985.

Pulvino,C.J., and Hansen,J.C. Relevance of "Needs" and "Press" to Anxiety, Alienation, and GPA. *The Journal of Experimental Education,* 1972, *40*(3), 70-75.

Ray,W.J., Katahn,M., and Snyder,C.R. Effects of Test Anxiety on Acquisition, Retention, and Generalization of a Complex Verbal Task in a Classroom Situation. *Journal of Personality and Social Psychology,* 1974, *20*(2), 147-154.

Salamé,R.F. Test Anxiety: Its Determinants, Manifestations and Consequences. In Ploeg,H.M., van der, Schwarzer,R., and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.3), Lisse: Swets and Zeitlinger B.V., 1984.

Sarason,I.R. Test Anxiety and Cognitive Modeling. *Journal of Personality and Social Psychology,* 1973, *28*(1), 58-61.

Sarason,I.R. Anxiety and Self-Preoccupation. In Sarason, I.R., and Spielberger,C.D. (Eds.), *Stress and Anxiety* (Vol.2), Washington,D.C.: Hemisphere Publishing Corporation, 1975.

Sarason,I.R., and Harmatz,M.G. Test Anxiety and Experimental Conditions. *Journal of Personality and Social Psychology,* 1965, *1*(5), 499-505.

Sarason,S.B., Davidson,K.S., Lighthall,F.F., Waite,R.R., and Ruebush,B.K. *Anxiety in Elementary School Children.* New York: John Wiley and Sons, Inc., 1960.

Schwarzer,C., and Cherkes-Julkowski,M. Determinants of Test Anxiety and Helplessness. In Schwarzer,R., Ploeg,H.M., van der, and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.1), Lisse: Swets and Zeitlinger B.V., 1982.

Schwarzer,C., and Kim,Moon-Joo. Adaptation of the Korean Form of the Test Anxiety Inventory: A Research Note. In Ploeg,H.M., van der, Schwarzer,R., and Spielberger, C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.3), Lisse: Swets and Zeitlinger B.V., 1984.

Schwarzer,R. Worry and Emotionality As Separate Components in Test Anxiety. *International Review of Applied Psychology,* 1984, *33*, 205-220.

Schwarzer,R., Ploeg,H.M., van der, and Spielberger,C.D. Test Anxiety: An Overview of Theory and Research. In Schwarzer,A., Ploeg,H.M., van der, and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.1), Lisse: Swets and Zeitlinger B.V., 1982.

Sipos,K., Sipos,M., and Spielberger,C.D. The Development and Validation of the Hungarian Form of the Test Anxiety Inventory. In Ploeg,H.M., van der, Schwarzer,R., and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.4), Lisse: Swets and Zeitlinger B.V., 1985.

Smith,R.E., Ascough,J.C., Ettinger,R.F., and Nelson,D.A. Humor, Anxiety, and Task Performance. *Journal of Personality and Social Psychology,* 1971, *19*(2), 243-246.

Spielberger,C.D. *Anxiety and Behavior.* New York: Academic Press, 1966.

Spielberger,C.D., Gonzalez,H.P., and Fletcher,T. Test Anxiety Reduction, Learning Strategies, and Academic Performance. In O'Neil,H.F., Jr., and Spielberger,C.D. (Eds.), *Cognitive and Affective Learning Strategies,* New York: Academic Press, 1979.

Spielberger,C.D., Gorsuch,R.L., and Lushene,R.E. *Manual for State-Trait Anxiety Inventory.* California: Consulting Psychologists Press, 1970.

Spielberger,C.D. *Preliminary Professional Manual for the Test Anxiety Inventory,* Palo Alto, CA: Consulting Psychologists Press, 1980.

Stephan,E., Fisher,M., and Stein,F. Self-Related Cognitions in Test Anxiety Research: An Empirical Study and Critical Conclusions. In Ploeg,H.M., van der, Schwarzer,R., and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.2), Lisse: Swets, and Zeitlinger B.V., 1983.

Sweetland,R.C., and Keyser,D.J. *Tests: A Comprehensive Reference for Assessments in Psychology, Education, and Business.* Kansas City: Test Corporation of America, 1984.

Tobias,S. Anxiety Research in Educational Psychology. *Journal of Educational Psychology,* 1979, *71*(5), 573-582.

Tryon,G.S. The Measurement and Treatment of Test Anxiety. *Review of Educational Research* 1980, *50*(2), 343-372.

Ulek,C., and Pruyn,A. Evaluative Stress and Performance: The Case of Test Anxiety. In Ploeg,H.M., van der, Schwarzer,R., and Spielberger,C.D. (Eds.), *Advances in Test Anxiety Research* (Vol.3), Lisse: Swets and Zeitlinger B.V., 1984.

Walker,H.M., and Lev,J. *Statistical Inference.* New York: Holt, Rinehart and Winston, 1953.

Watson,D. Relationship Between Locus of Control and Anxiety. *Journal of Personality and Social Psychology,* 1967, *6* (1), 91-92.

Weinstein,C.E., Cubberly,W.E., and Richardson,F.C. The Effects of Test Anxiety on Learning at Superficial and Deep Levels of Processing. *Contemporary Educational Psychology*, 1982, *7*, 107-112.

Winer,B.J. *Statistical Principles in Experimental Design*. New York: McGraw-Hill Book Company, 1971.

Wittmaier,B.C. Test Anxiety, Mood and Performance. *Journal of Personality and Social Psychology*, 1974, *29*(5), 664-669.

Vidler,D. Convergent and Divergent Thinking, Test-Anxiety, and Curiosity. *The Journal of Experimental Education*, 1974, *43*(2), 79-85.

Zatz,S., and Chassin,L. Cognitions of Test-Anxious Children. *Journal of Consulting and Clinical Psychology*, 1983, *51*(4), 526-534.

Zülemyan,A. *A Comparison of Cognitive Therapy, Modified Desensitization, Cognitive Behavior Modification and Control Groups for Reducing Test Anxiety*. Unpublished Master's Thesis, Boğaziçi University, 1979.

APPENDIX A

TEST ATTITUDE INVENTORY

NAME..... AGE .. DATE ..... SEX ..... (M) (F)

DIRECTIONS: A number of statements which people have used to describe themselves are given below. Read each statement and then blacken the appropriate circle to the right of the statement to indicate how you *generally* feel. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe how you generally feel.

	almost never	some-times	often	almost always
1. I feel confident and relaxed while taking tests	(1)	(2)	(3)	(4)
2. While taking examinations I have an uneasy, upset feeling	(1)	(2)	(3)	(4)
3. Thinking about my grade in a course interferes with my work on tests	(1)	(2)	(3)	(4)
4. I freeze up on important exams	(1)	(2)	(3)	(4)
5. During exams I find myself thinking about whether I'll ever get through school	(1)	(2)	(3)	(4)
6. The harder I work at taking a test, the more confused I get	(1)	(2)	(3)	(4)
7. Thoughts of doing poorly interferes with my concentration on tests	(1)	(2)	(3)	(4)
8. I feel very jittery when taking an important test	(1)	(2)	(3)	(4)
9. Even when I'm well prepared for a test, I feel very nervous about it	(1)	(2)	(3)	(4)
10. I start feeling very uneasy just before getting a test paper back	(1)	(2)	(3)	(4)
11. During tests I feel very tense	(1)	(2)	(3)	(4)
12. I wish examinations did not bother me so much	(1)	(2)	(3)	(4)
13. During important tests I am so tense that my stomach gets upset	(1)	(2)	(3)	(4)
14. I seem to defeat myself while working on important tests	(1)	(2)	(3)	(4)
15. I feel very panicky when I take an important test	(1)	(2)	(3)	(4)
16. I worry a great deal before taking an important examination	(1)	(2)	(3)	(4)
17. During tests I find myself thinking about the consequences of failing	(1)	(2)	(3)	(4)
18. I feel my heart beating very fast during important tests	(1)	(2)	(3)	(4)
19. After an exam is over I try to step worrying about it, but just I can't	(1)	(2)	(3)	(4)
20. During examinations I get so nervous that I forget facts I really know	(1)	(2)	(3)	(4)

APPENDIX B

SINAV TUTUMU ENVANTERİ

İSİM .... YAŞ .. TARİH ..... CİNSİYET ... (E) (K)

YÖNERGE: Aşağıda, insanların kendilerini tanımlamak için kullandıkları bir dizi ifade sıralanmıştır. Bunların herbirini okuyun ve *genel olarak* nasıl hissettiğinizi gösteren ifadenin sağındaki boşluklardan uygun olanın içini karalayın. Burada doğru ya da yanlış yanıt yoktur. İfadelerin hiçbiri üzerinde fazla zaman harcamayın, ancak yazılı ve sözlü sınavlarda genel olarak nasıl hissettiğinizi gösteren yanıtı işaretleyin.

	hemen hiçbir zaman	bazen	çoğu zaman	hemen her zaman
1. Sınav sırasında kendimi güvenli ve rahat hissederim	(1)	(2)	(3)	(4)
2. O dersten alacağım notu düşünmek, sınav sırasındaki başarımı olumsuz yönde etkiler	(1)	(2)	(3)	(4)
3. Önemli sınavlarda donup kalırım	(1)	(2)	(3)	(4)
4. Sınavlar sırasında, birgün okulu bitirip bitiremeyeceğimi düşünmekten kendimi alamam	(1)	(2)	(3)	(4)
5. Bir sınav sırasında, ne kadar çok uğraşırsam kafam o kadar çok karışır	(1)	(2)	(3)	(4)
6. Sınavlarda kendimi huzursuz ve rahatsız hissederim	(1)	(2)	(3)	(4)
7. Önemli bir sınav sırasında kendimi çok sinirli hissederim	(1)	(2)	(3)	(4)
8. Başarısız olma düşünceleri, dikkatimi sınav üzerinde toplamama engel olur	(1)	(2)	(3)	(4)
9. Bir sınava çok iyi hazırlandığım zamanlar bile kendimi oldukça sinirli hissederim	(1)	(2)	(3)	(4)
10. Önemli sınavlarda sinirlerim öylesine gerilir ki midem bulanır	(1)	(2)	(3)	(4)
11. Bir sınav kağıdını geri almadan hemen önce çok huzursuz olurum	(1)	(2)	(3)	(4)
12. Önemli sınavlarda kendimi adeta yenilgiye iterim	(1)	(2)	(3)	(4)
13. Sınavlar sırasında kendimi çok gergin hissederim	(1)	(2)	(3)	(4)
14. Önemli bir sınav sırasında paniğe kapılırım	(1)	(2)	(3)	(4)
15. Sınavların beni bu kadar rahatsız etmemesini isterdim	(1)	(2)	(3)	(4)
16. Önemli bir sınava girmeden önce çok endişelenirim (kurarım)	(1)	(2)	(3)	(4)
17. Sınavlar sırasında, başarısız olmanın sonuçlarını düşünmekten kendimi alamam	(1)	(2)	(3)	(4)
18. Önemli sınavlarda kalbimin çok hızlı attığını hissederim	(1)	(2)	(3)	(4)
19. Sınav sona erdikten sonra endişelenmemeye (kurmamaya) çalışırım, fakat yapamam	(1)	(2)	(3)	(4)
20. Sınavlar sırasında öylesine sinirli olurum ki aslında bildiğim şeyleri bile unuturum	(1)	(2)	(3)	(4)

APPENDIX C

Formulas for:

(1) item-total correlation

$$r_{yi} = \frac{\Sigma y_i - \frac{(\Sigma_y)(\Sigma_i)}{N}}{\sqrt{(\Sigma_y^2 - \frac{(\Sigma_y)^2}{N})(\Sigma_i^2 - \frac{(\Sigma_i)^2}{N})}}$$

where y = total score
i = item score

(2) item-remainder correlation (corrected form of item-total correlation)

$$r_{1(y-1)} = \frac{r_{y1}\ \sigma_y - \sigma_1}{\sqrt{\sigma_1^2 + \sigma_y^2 - 2\sigma_1\sigma_y\ r_{y1}}}$$

where r_{y1} = correlation of item 1 with total scores y
σ_y = standard deviation of total scores
σ_1 = standard deviation of item 1

$r_1(y-1)$ = correlation of item 1 with sum of scores on all items exclusive of item 1
(Nunnally, 1967, pp.261-263).