Chapter VIII
Understanding Learner Trait, Test and Computer Anxiety in the Context of Computer-Based Testing

Elena C. Papanastasiou
University of Nicosia, Cyprus

Aimilia Tzanavari
University of Nicosia, Cyprus

Patricia Lowe
University of Kansas-Lawrence, USA

ABSTRACT

Testing is an integral part of the learning process that aims to estimate the learner’s abilities as accurately and efficiently as possible. This estimation frequently is influenced by factors such as the learner’s emotional state and traits. This chapter looks into the area of Computer-based Testing (CBT), visiting the relevant literature on the subject, and then investigates the particular emotional states of learner trait, test and computer anxiety in that context. A study was carried out and revealed that although both trait and test anxiety as variables do not significantly affect learner performance, computer anxiety does. Finally, future research trends in this area are outlined.

INTRODUCTION

The goal of tests in general, as well as of computer-based tests is to be able to estimate the examinee’s abilities as accurately and efficiently as possible, by removing as many extraneous effects (such as the computer itself) as possible. However, in many cases the computer itself might become an
obstacle to the learning and testing process. The medium of computer technology might even be a reason that can explain a portion of the differences in the performance of students on computer administered tests compared to paper-and-pencil tests. For example, many examinees might be unfamiliar with the use of computers for learning purposes, or for taking tests on the computer. More specifically, they might be unfamiliar with the computer-testing mode, or with the fact that they have to read the test items on a computer screen and enter their answers through the keyboard. Therefore, taking tests on the computer might create anxiety for such examinees, and even more so for test anxious examinees (Wise, Roos, Plake & Nebelsick-Gullett, 1994). However, the literature seems to be lacking in this area, since relatively few studies have examined examinee characteristics and their levels of trait, test and computer anxiety in relation to computer administered tests. This is even more so when considering examinees of lower socioeconomic status (SES) that do not have frequent access to computer technology. Such students might not be as familiar with the use of computers, and more specifically with taking tests on a computer, which in turn could interfere with their test performance.

Examinees with high levels of computer or test anxiety might end up responding to computer administered tests in ways that are different compared to the examinees that are familiar with the use of the computer and with taking tests on them. Individual differences like these can provide a framework for explaining significant divergences in test performance and be used to improve the way computer-based tests are designed. Therefore, it is essential for researchers to also understand the examinee characteristics, and on how they can be simulated appropriately in simulation studies, and how that can affect the creation of e-learning content and their corresponding tests (Harwell, Stone, Hsu & Kirisci, 1996; Stocking, Steffen & Eignor, 2001).

This chapter examines the level of test anxiety and computer anxiety that is held by students in Cyprus, in relation to their performance on a computer-based test. More specifically, this chapter attempts to provide answers to the following research questions:

- Can the variables of overall anxiety, computer anxiety, and test anxiety predict a significant portion of an examinee’s score on a computer-based abstract reasoning test?
- What are the variables that affect the levels of test and computer anxiety for undergraduate students in Cyprus?

**BACKGROUND**

**Computer-Based Testing**

There is no doubt that the new trend in assessment worldwide, is that of computer-based and/or web-based testing. Computer-based tests (CBT) could be defined as any type of assessment that is administered through the computer. However, computer-based testing can encompass many forms, depending on how adaptable the test is on the item level (The College Board, 2000). For example, some CBT, which are also called computerized fixed tests, are purely linear (Parshall, Spray, Kalohn & Davey, 2002). These are the tests that most closely resemble paper and pencil tests, since they are fixed form, fixed length, and the test items are organized in advance and placed in a predetermined order. In contrast to computerized fixed tests, computer adaptive tests (CAT) are the computer-based tests that have the maximum degree of adaptivity since they can be adapted for each examinee, based on the amount, difficulty and order in which the items are administered to each examinee. Web-based tests are computer-based tests delivered through the internet or an intranet. The tests themselves reside on the examiner’s web server and examinees can access them through
their web browsers. One of the main strengths of this type is that the tests can be taken anytime, from anywhere.

So overall, computer-based tests in general, as well as web-based tests represent the next generation in educational testing because of their many advantages over paper and pencil tests. One of the advantages of computer-based tests in general is that they can be comprised of more creative and interactive item types than regular paper and pencil tests. Tests today no longer have to be confined to pure text items that might include a few two dimensional pictures. Items used for computer-based tests can include colorful high resolution pictures, movies with motion and sound, voice synthesizers, and oral comprehension of spoken language (Parshall, Spray, Kalohn & Davey, 2002; Wainer, 2000), which all make the testing process more interesting and appealing to the students. In addition, when combined with computer-based learning, computer-based tests could also include test items obtained from examples or experiments used during the computer-based learning process. These types of items, when combined with their multimedia properties could make the testing situation a bit more realistic and more similar to the actual learning process. In addition to making the testing situation more realistic, such items have the advantage of eliciting positive attitudes from the students who take such tests on the computer. Consequently, the students end up being more motivated to do their best, which in turn increases the reliability of the student's test scores.

Administering tests on the computer also has the advantage of being able to accommodate students with different special needs. More specifically, because of the advanced multimedia properties of technology today, computers have the capability of easily enabling educators to teach, as well as test students with disabilities such as visual or hearing impairments. For example, CBTs could easily assign longer testing times to students with disabilities, as well as use sound, motion or text to assist these students appropriately, based on their impairments. These types of accommodations could alleviate some of the frustration that is typically faced by students with special needs, which in turn could assist them in demonstrating their actual knowledge more accurately.

Another advantage of computer-based tests is that they can provide direct and immediate feedback to students as well as to teachers. With typical paper and pencil tests, there always tends to be a lag of time between the administration of the test and of its scoring. As a result, most of these tests end up being used only summatively in an attempt to assign grades to students. Educationally this is not a very sound practice since it does not allow the use of formative feedback that could supplement the learning process. With computer-based testing however, test results could be provided to students immediately after they finish responding to their test. In addition, a list of the content areas and objectives that have been met by each student can also be produced by such tests.

Finally, computer-based tests also include some additional advantages that will not be discussed in length. These include the fact that CBTs enhance test security, since test sheets or booklets can no longer be stolen or passed around from student to student. In addition, through CBTs, other collateral types of data can be collected that can provide additional pieces of information for educators and parents. Such information can include response times, which is the amount of time needed to answer each question, and the number of changes that the students made to their answers on the test. Response times are useful because they also tend to provide an indication on the amount of effort that has been placed by each student on each test item. For example items with minimal response times at the beginning of the test might indicate that the student did not give any effort to those items overall. However, items with minimal response times at the end of the test might indicate that a student might not have had enough time to answer the item and has randomly guessed the answer. Examining patterns of answer changing
might also provide information on the quality of certain items (Papanastasiou & Reckase, 2008). For example, items to which many examinees tend to change their answers might be an indicator of a flawed item.

**Computer Adaptive Tests**

In addition to the advantages that are associated with the general use of computer technologies to administer tests, there are some additional advantages that are associated specifically with computer adaptive tests. In order to comprehend the advantages associated with CATs, it is necessary to comprehend the way in which these tests are compiled. Typically, computer adaptive tests start by administering an item of about average difficulty to the examinees. This is done because the ability of most examinees lies in the middle of the ability continuum. Consequently, the majority of the students or examinees are able to respond to such average difficulty items, because they will not be too easy or too difficult for them. After the first item is answered, and based on if it is correct or incorrect, the second item is administered. If for example an examinee A has responded to the first item on a science computer adaptive test correctly, that indicates to the adaptive test that this examinee has above average ability in science. Consequently, the computer will administer a second, more difficult item to that examinee. If this examinee also answers the second item correctly, then the third item administered will be even more difficult that the first two items. The goal is to be able to administer a test whose difficulty matches the achievement level of the examinee. When this is achieved, the ability of the examinee can be estimated with small margins of error.

If another examinee B had responded incorrectly to the first item, that indicates to the adaptive test that the science ability of this student is below average. So in an attempt to find out where the ability level of this examinee really is, the next item that will be administered will be easier than the first one. If examinee B also responds incorrectly to the second item, the third item that will be administered will be even easier. This process will continue to take place until an accurate estimate of the examinee’s science ability is obtained. Therefore, by administering more items that are matched to each student’s ability, the examinees are least likely to despair about their performance, particularly if s/he already suffers from anxiety about the test (Linacre, 2000).

The main advantage of computer adaptive testing however, is based on the fact that it can estimate examinees’ abilities with better accuracy compared to other types of tests (e.g. computerized fixed tests that are not adaptable). Moreover, this accuracy can be achieved by administering fewer items, which in turn reduces the testing time needed for examinees to complete the test and also their fatigue, a factor that can significantly affect an examinee’s test results. Another benefit is that examinees are less likely to be led to unwanted behavior such as guessing, carelessness or response patterns (Linacre, 2000), since the items they receive are matched to their ability, and are therefore neither too easy nor too difficult for them.

**Computers and Anxiety**

In order to be able to properly develop and implement a web-based test, a lot of attention needs to be paid to the educational context within which it will be implemented. In order to ensure that students’ true ability or trait is being measured by a web-based test with a small amount of measurement error, students need to be familiar with the use of computers. Although in most developed countries computers are easily available for most students to use, this is not always the case worldwide. For example, the impact of computers familiarity on achievement in the Republic of Korea, where 98 percent of the students own computers, cannot be directly compared to Indonesia where only 17 percent of students own computers. Nor could one
compare the Islamic Republic of Iran, where only 2 percent of students use computers both at home and at school, with Hong Kong, where 89 percent of students use a computer in both locations (Papanastasiou & Paparistodemou, 2007).

Consequently, if certain students do not know how to use computers and are asked to take a test on a computer, most likely these students’ test scores will actually represent their amount of computer knowledge rather than their true abilities. Such results would be similar to asking very bright students who do not speak English to take a geography test in English. Although the students might know a lot about geography, the language factor will be a barrier to them in demonstrating their actual geography knowledge. The same is the case with computer skills. Students who are not familiar with the use of computers might have to struggle to understand how to enter their test results on the computer, which would prevent them from actually demonstrating what they really know. For example, a study based on data from the PISA database, found that the students in the United States of America who felt least comfortable with using computers to write papers, and who did not have easy access to use a computer at home or in the library, and who rarely used word processing software were more likely to have lower science scores (Papanastasiou, Zembrillas & Vrasidas, 2003).

Consequently, it is possible that computer-based tests could disadvantage students from lower socioeconomic backgrounds since they tend to have limited access to computers. In these cases, it is possible that the computer itself might become an obstacle to the testing process for the students. The medium of computer technology might even be a reason that can explain a portion of the differences in the performance of students on computer administered tests from paper-and-pencil tests. More specifically, many examinees might be unfamiliar with the use of computers or with taking tests on computers, which could hurt their performance on such tests. A study by Stricker and Wilder (2001) found that “computer liking” and “computer confidence” were positively related to the performance of students on a computer adaptive version of the Test of English as a Foreign Language (TOEFL) test. So the students who were most familiar with the use of computers and who liked computers had higher scores on the TOEFL compared to other students with less confidence and more negative attitudes towards computers.

The existence of a negative relationship between test anxiety and test performance is nothing new in the psychometric and educational psychology literature. In a meta-analysis of 562 studies, Hembree (1988) found that an inverse relationship existed between test anxiety and performance on tests, starting from grade 3. Similar results have also been found in more recent research studies. In a study performed by Powers (1999), the worry component of anxiety was negatively related to test performance on the GRE quantitative, analytical and verbal test scores combined. However, the emotionality component of anxiety was not significantly related to the performance on the GRE. The results of a different study performed by Hancock (2001), also found that students with high levels of test anxiety tend to perform more poorly on tests, especially when the tests count towards their evaluation in the classroom.

Other studies claim that computer-based tests cause increased levels of anxiety to students, beyond the regular test anxiety that students typically feel. In a study by Stricker and Wilder (2001), a negative relationship was found between computer anxiety and performance on the computer-based version of the TOEFL. However, Wise, Barnes, Harvey and Plake (1989) found no negative effects of computer anxiety on the performance of students on a computer-based test in a study was much smaller in scale and was not as high stakes as the test used in the Stricker and Wilder (2001) study.

In a single study that examined the effects of computer adaptive testing on anxiety, the results
that were found indicated that the test anxiety levels of examinees who took computer adaptive tests were lower than their counterparts who took a paper-and-pencil test (Powers, 2001). According to the author however, the design of the study was not able to demonstrate whether this difference was due to the mode of testing, or whether there were other preexisting differences between the two groups of students.

Overall, based on the extant literature, it can be concluded that issues of computer experience, computer anxiety, test anxiety, and state and trait anxiety are all variables that add additional sources of error in students’ test scores on computer-based tests. What is also likely is that these sources of error might be even larger for students who are not as familiar with the use of computer technology. Therefore, it is imperative that researchers and educators should determine what extraneous variables possibly mask or interfere with the performance of students on computer and web-based tests. Knowledge of these factors is essential to be able to prepare students for the specific testing situation. This would ensure that measurement error in students’ test scores would be minimized, which in turn would lead to more accurate test results.

METHODS

For the purpose of this study, the students who were enrolled in a research methods course at the University of Cyprus were asked to participate in this study. The majority of the sample consisted of sophomore students, who were all training to become elementary school teachers in Cyprus. The average age of the students in the sample was 19.96 years, with a standard deviation of 2.42 years. Of the 381 students who were in the sample, 15.2% were male and 84.8% were female. The disproportionate amount of males to females in the sample reflects the fact that elementary school teaching is a predominantly female oriented field in the Greek culture, as in many other Western cultures.

The students that participated in the study were asked to respond to a background survey, a computer anxiety measure, an abstract reasoning test as well as to two anxiety surveys. All of these measures and surveys were web-based. The software used for the administration of the computer administered portion of this study was CATGlobal. In addition, the students were also asked to respond to the Adult Manifest Anxiety Scale-College Version (AMAS-C; Reynolds, Richmond, & Lowe, 2003a) that was administered on paper. On average, the students spent a total of one hour in this study.

Computer Anxiety Measure

The computer anxiety measure, which was created by the researchers, is a 7-item self-report measure used to assess the student’s anxiety in relation to the use of computers. The items included in this measure assess the student’s self perceived anxiety and unfamiliarity and fear of computers. Cronbach’s alpha reliability estimate for this measure was 0.83.

Adult Manifest Anxiety Scale-College Version (AMAS-C)

The Adult Manifest Anxiety Scale-College Version (AMAS-C; Reynolds, Richmond, & Lowe, 2003a) is a 49-item self-report measure used to assess chronic, manifest anxiety in the college student population. The AMAS-C consists of four anxiety subscales (Worry/Oversensitivity, Physiological Anxiety, Social Concerns/Stress, and Test Anxiety) and a Lie scale. The Lie scale is a measure of social desirability. In addition to the four anxiety subscales and Lie scale, the AMAS-C has a Total Anxiety scale. The Total Anxiety scale provides a global measure of chronic, manifest anxiety (Lowe, Papanastasiou, DeRuyck & Reynolds, 2005). Raters rate their responses to the
49 items on a dichotomous scale, using a yes/no format. A “yes” response to an item is indicative of how a student generally thinks, feels, or acts, whereas a “no” response to an item is not indicative of how an individual generally thinks, feels, or acts. Scores are derived from a student’s “yes” responses, with higher scores suggesting higher levels of anxiety (Reynolds et al., 2003b).

The AMAS-C has adequate psychometric properties. Internal consistency reliability estimates for the Total Anxiety scale, four anxiety subscales, and Lie scale scores ranged from .72 to .95. Temporal stability of the AMAS Total Anxiety scale and four anxiety subscales scores ranged from .76 to .87 over a 1-week test-retest interval. Evidence supporting the construct validity of the AMAS-C test scores has been reported (see Reynolds et al., 2003b). The convergent validity of the AMAS-C test scores has also been examined. Lowe, Peyton, and Reynolds (2007) found moderate correlation coefficients between the AMAS-C Total Anxiety scale scores and four anxiety subscale scores and the Trait scale scores of the State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, Vagg, & Jacobs, 1977). These findings provide support for the convergent validity of the AMAS-C test scores.

Test Anxiety Inventory (TAI)

The Test Anxiety Inventory (TAI; Spielberger, 1980a) is a 20-item self-report measure designed to assess test anxiety in high school and college students. The TAI consists of a Total scale and two subscales (Worry and Emotionality). Respondents rate their responses to the 20 different statements on a 4-point Likert scale.

The TAI has good psychometric properties. Coefficient alphas of .92 and above were found for the TAI Total scale scores. Internal consistency reliability estimates for the Worry and Emotionality subscales scores were .88 and .90, respectively. Test-retest reliability coefficients for the TAI Total scale scores ranged from .80 to .81 over a two- to four-week test-retest interval. Evidence supporting the construct validity of the TAI test scores has been found (see Spielberger, 1980b).

Abstract Reasoning Test

The only timed section of the study was the abstract reasoning test (Embretson, 1998), in which the students had 35 minutes to respond to 30 multiple-choice abstract reasoning items. The items were presented in a 3x3 matrix format, with the bottom left corner intentionally left blank. A series of 8 response options were provided and the students had to choose a single correct answer for each matrix. An example of such a problem is presented in Figure 1. In this example, the student would have to examine the 3x3 matrix that is presented on the left of Figure 1, and try to determine which figure should be included where the question mark is located. The eight response options are located in the left of the figure.

After the test was completed, the total score of each student was converted to t-scores for the analyses that were performed for this study.

RESULTS

A primary purpose of this study was to try to explain a portion of the variance of student’s scores on a web-based test. Possible variables

Figure 1. Abstract reasoning item example
that were hypothesized to influence the performance of the students on computer-based tests were issues of test anxiety and computer anxiety. These analyses were performed through a series of regression models.

Predicting Computer-Based Testing Test Scores

The first regression model that was performed, examined whether the performance of the students on the computer-based test of abstract reasoning could be explained in part by their trait anxiety, by their test anxiety, as well as by their levels of computer anxiety. As presented in Table 1, the overall model was statistically significant \( (F_{3,204}=8.205, p=0.000) \), which explained 10.9% of the variance of the student’s abstract reasoning score. However, of the three variables that were entered into the regression, only the variable of computer anxiety was statistically significant in predicting the student’s test score \( (\beta=-0.241, t=-4.185, p=0.000) \). So students with lower levels of computer anxiety performed better on the CAT than students with higher levels of computer anxiety. The student’s trait and test anxiety though, did not significantly predict the dependent variable of the student’s total score on abstract reasoning.

Test Anxiety

The second regression that was performed tried to determine the variables that could predict a student’s test anxiety. The overall regression whose results are presented in Table 2 was statistically significant \( (F_{3,286}=19.413, p=0.000) \), which explained 16.9% of the variance of the student’s test anxiety level. This analysis included three independent variables. Two of those were demographic (age and gender), while the third variable was that of the student’s computer anxiety. The variables of gender and computer anxiety were significant in predicting test anxiety in students, while age was not. More specifically, the results of this analysis showed that females tended to have higher levels of test anxiety than males \( (\beta=-0.723, t=-4.557, p=0.00) \), and that students with a higher level of computer anxiety also had a higher level of test anxiety \( (\beta=0.031, t=5.802, p=0.00) \).

Computer Anxiety

There was 31.8% of the sample that indicated that they were afraid of computers, 30.5% who indicated that they usually feel nervous when they use computers, and 27.3% who found it difficult to use a computer. In terms of computer anxiety, 6.5% of the students responded that they have a lot of computer anxiety, 60.8% responded that they have some anxiety, while only 32.7% of the sample indicated that they have no computer anxiety.

When asked about their familiarity with the use of computers, 1.2% of the sample indicated that they were very unfamiliar and 13.2% were unfamiliar with the use of computers.

\[
\begin{array}{|c|c|c|c|}
\hline
\text{Unstandardized Coefficients} & \text{Standardized Coefficients} & \text{t} & \text{Sig.} \\
\hline
\text{(Constant)} & 63.507 & 2.896 & 21.927 & 0.000 \\
\text{Trait anxiety (AMAS-C)} & 0.478 & 0.620 & 0.057 & 0.771 & 0.442 \\
\text{Test anxiety (TAI)} & -0.866 & 0.637 & -0.105 & -1.359 & 0.176 \\
\text{Computer anxiety} & -0.241 & 0.058 & -0.293 & -4.185 & 0.000 \\
\hline
\end{array}
\]

Table 1. Predicting student’s scores on a computer-based tests
six percent of the sample indicated that they were familiar and only 9.2% responded that they were very familiar with the use of computers. When asked about how comfortable they felt with the use of computers, 7.7% indicated that they were not at all comfortable, 34.8% were somewhat comfortable and only 12.0% were very comfortable with the use of computers.

In terms of taking tests on the computer, the students from Cyprus tended to have higher levels of anxiety. More specifically, only 36.5% of the students had no anxiety, while 53.5% had some anxiety and 10.0% had a lot of anxiety. It should be noted however, that these responses are in light of the fact that this test was not high stakes.

The third regression that was performed examined whether students’ computer anxiety was influenced by the variables of owning a computer, number of years of using computers, hours per week of using computers, and gender. This regression model was also statistically significant ($F_{4,313} = 39.523, p=0.000$), and it explained 33.8% of the variance of the student’s computer anxiety. All but one of the variables used in this regression were statistically significant. More specifically, students who owned a computer at their home had lower levels of computer anxiety ($\beta=4.712, t=2.906, p=0.004$). Students who had been using computers for many years also had lower levels of computer anxiety ($\beta=-1.552, t=-7.452, p=0.000$). Finally, students who used the computer for many hours per week also had lower levels of computer anxiety ($\beta=-0.577, t=6.154, p=0.000$). However, gender was not significant in predicting the computer anxiety score of the sample. These results are presented in Table 3.

### CONCLUSION

The results of the study presented in this chapter show that trait anxiety and test anxiety are not variables that can significantly affect performance of students on computer administered tests. However, computer anxiety is a variable that significantly affects the performance of students.
In turn, this anxiety is affected by whether the student has a computer available to use at home, by the number of years each student has been using computers, as well as by the number of hours of computer use per week. The results of this study indicate that since the exposure that students have to computers can be manipulated by intentionally increasing their use of computers in their educational settings, the possible disadvantages associated with computer-based tests (such as the student’s computer anxiety) can be reduced dramatically. This is especially important when taking into account students with a lower socioeconomic status that might be disadvantaged from such technology if they cannot afford to own a computer at home. Therefore, it is important to ensure that all students have equal access to such technologies before trying to implement computer-based tests on a broad basis.

An encouraging result of this study is that gender does not play a role on students’ computer anxiety although males typically experienced lower levels of test anxiety compared to females. Although the gender composition of the present sample appears to reflect that of most educational courses (Onwuegbuzie, Bailey & Daley, 1999), the fact that participants were predominantly female is a limitation of the present study (see also, Onwuegbuzie, 2000). Certainly, the inclusion of more male students would facilitate the generalizability of the findings. The generalizability of the findings would also be strengthened with the inclusion of high school students in the sample, in addition to graduate students at the University level.

In sum, a conclusion that is very clear from this study is that the attempt to reach the goal of removing as many extraneous effects of testing from computer-based tests has not been completely successful yet. Although computer-based and adaptive tests might be able to estimate the examinee’s abilities with high levels of accuracy and efficiency, they have not managed to meet this objective yet. It is possible that in the future, when the use of computers becomes even more widespread than it is today, many such problems will be alleviated. However, until that time comes, we need to be more considerate to those students who are not privileged enough to have access to computer technology. In addition, future research should also determine whether the anxiety levels of examinees on computer adaptive tests are comparable to those expressed on traditional computer-based tests or not, and if not, try to examine how they can be dealt with.

ACKNOWLEDGMENT

We would like to sincerely thank CATGlobal for permitting us to use their software and services through the Computer Adaptive Technologies (CAT) Software System research grant. We would also like to acknowledge the invaluable help of John Stahl at Promissor for his assistance to us while we were using the CATGlobal software throughout this process.

REFERENCES


