



Saudi EFL Learners' Test Anxiety Levels during CBT and PBT

Osama Ahmed Alwazzan

(oso.alwa@gmail.com)

King Saud University, Saudi Arabia

Hesham Suleiman Alyousef

(hesham@ksu.edu.sa)

King Saud University, Saudi Arabia

Abstract

Individuals' attitudes toward computer-based testing (CBT) can either increase or decrease their test anxiety level. As a result, they may be highly resistant to change or they may have different types of behavior. It is, therefore, pertinent to investigate individual students' attitudes toward computers before administering CBT. This study investigated the effectiveness of CBT when compared to traditional paper-based testing (PBT) in terms of reducing test anxiety among Saudi learners of English as a foreign language (EFL). The participants were forty undergraduate male students enrolled in a first-year preparatory EFL course at a Saudi university. The students were divided into two groups, a CBT group and a PBT group, and took the same assessment in the two test formats. The participants completed two questionnaires to examine their anxiety levels before and during each test, and to collect information regarding their prior experience with and attitudes toward either testing format. An analysis of variance comparison was used to compare anxiety levels between the CBT and PBT groups. The findings showed that students had a neutral attitude toward both types of tests. Furthermore, the students' attitudes did not have a significant impact on anxiety levels for either testing format. Rather, anxiety levels experienced in both testing formats were significantly correlated with each other and the attitude factor. Further research is required to delineate factors that may contribute to test anxiety as well as provide further insight into the under-represented student subgroup that participated in the study.

1 Introduction

The recent rapid increase in technological advancement has been altering the field of education. This has resulted in inevitable changes in the way tests and assessments are designed and administered. The integration of technology into education has resulted in a shift away from paper-based testing (PBT) toward computer-based testing (CBT). Regardless of whether these changes are being deliberately or unintentionally integrated into the classroom by instructors, the presence of CBT is changing the landscape of the field of education (Ghaderi, Mogholi, & Soori, 2014).

Furthermore, as mastering the English language accelerates global development by opening doorways to knowledge, commerce, and culture, there is hence a salient need for an in-depth understanding of the implications of administering tests via CBT and PBT for learners of English as a foreign language (EFL). Test anxiety refers to the stress and apprehension students experience when taking tests and exams, and can indicate individual differences among students and how they cope

with stress (Zeidner, 1998). It is important to acknowledge pre-existing individual differences that may have significant impact on levels of test anxiety. For instance, if students' potential achievement on tests is inhibited by factors such as apprehensiveness, anxiety, or fear, then the test results do not accurately represent their learning outcomes. Not only does this result in wasted time and expenses, but it can also delay a student's progress. Therefore, it is necessary to investigate the relationship between the assessment methods of CBT and PBT, as well as the students' backgrounds, attitudes, and perceptions of the two test formats to improve our understanding of the causes of test anxiety. The present study examines the degree and causes of anxiety in English CBT and compares it to the traditional method of testing, PBT. As many Saudi students take computerized tests, it is important to investigate the anxiety levels (if any) they may experience during these tests in order to compare it with that in PBT.

This study examined the anxiety that affects EFL students when undergoing computerized tests and assessments in an English Department at the university level, and it evaluated the shift in the students' attitudes when switching from PBT to CBT. One might argue that student anxiety levels will be very different in normal classroom testing situations from situations where CBT is utilized. Despite the recent dramatic increase in the use of technology in peoples' lives, some students still experience anxiety when undergoing computer-based examinations. The increase in the use of CBT in Saudi Arabia, along with the discomfort with English itself, may create student anxiety during tests. Even if students are comfortable with English outside the classroom, the very nature of testing remains a potential source of student anxiety.

The main aim of this study was to evaluate the differences in test anxiety levels between CBT and PBT among Saudi EFL learners by examining 1) the students' attitudes toward both CBT and PBT and 2) the relationship between the testing formats and levels of anxiety experienced during both tests.

2 Literature review

Over the last two decades, computers have been increasingly used in the classroom. Information technology has become an essential aspect of the educational process, both in the sciences or humanities. Few studies have investigated and compared student anxiety levels related to CBT and PBT, and those that have been conducted have yielded conflicting results (Baig, Al-Zahrani, Al-Jubran, Chaudhry, & Qadri, 2018; Bensoussan, 2012; H. Chen, 2012; K. T.-C. Chen, 2012; Hardre, Crowson, Xie, & Ly, 2007; Huberty, 2009; Kolagari, Modanloo, Rahmati, Sabzi, & Atace, 2018; Shi, 2012; Sulistyaningsih, 2016; Tella & Bashorun, 2012). In the field of EFL research, Ma, Andersson, and Streith (2005) noted that the integration of computer assisted language learning (CALL) decreased student anxiety levels and improved academic performance. Al Shammari (2007) argued that student anxiety stems from both the use of computers and the English language. Excessive negative attitudes toward computers were found to be related to a lack of proficiency in English, which can lead not only to anxiety but also to demotivation and avoidance of CBT. Conti-Ramsden, Durkin, and Walker (2010) and Shi (2012) reported that students with low English proficiency levels are more vulnerable to apprehensiveness and anxiety during CBT.

In fact, there are many factors that may contribute to anxiety in language learning, which in turn leads to anxiety concerning the assessment and/or testing system (Ghaderi, et al., 2014; Liu & Cheng, 2014; Wise & DeMars, 2005). Among these are personal motivation/perception (e.g. the learner's own sense of self-efficacy), personality, characteristics, competitiveness, social interaction, social environment, cultural differences, and amount of exposure to the target language. For example, Al-Amri (2009) provided a succinct account of the comparability of CBT and PBT. While he pointed out the advantages of using CBT, several factors were considered in determining the causes of test anxiety and how to overcome it. Familiarity with computers, attitudes toward computers, testing-mode preference, and test-taking strategies via CBT or PBT can determine student anxiety levels to a large extent. Tella and Bashorun (2012), however, found in their case study that most of the undergraduate Nigerian students preferred CBT over PBT. Parshall, Spray, Kalohn, and Davey (2002) argued that, regardless of the advantages of CBT over PBT, computerized assessment

and/or testing is not necessarily better than PBT and, therefore, there is no guarantee that PBT leads to lower student anxiety levels.

One might justifiably assume that addressing anxiety levels amongst students in an EFL environment is not only associated with the students' experience with the target language, but also directly associated with the testing mode and its effects. Piaw (2012) argued that repeated exposure to PBT leads to an assimilation of testing procedures by students. This sense of familiarity, which involves knowledge of the expectations and requirements associated with PBT, generally leads to a reduction in test anxiety levels. Students are generally accustomed to sitting for traditional tests, with only minimal adjustments or adaptations required when there is a slight difference in the format or structure (Hensley, 2015). Thus, PBT is easier to understand, as there is no need for the students to comprehend any technical terminologies or testing procedures before the test commences, which can add unnecessary stress. Contrary to the general perceptions of students, however, PBT can lead to higher than expected test anxiety levels. Unlike the flexibility of CBT, PBT is administered to all students at the same time and venue. An individual's anxiety level in the presence of others in PBT can be higher than it would be for CBT.

Although PBT is still widely used in Saudi Arabia as an assessment and evaluation medium, there has been an increase in the use of CBT in recent years. Due to technological advancements, CBT has become more common than PBT, because it is a more efficient and convenient method of measurement. CBT aims to produce reliable and fair assessment, which minimizes students' frustration (Chuah, Drasgow, & Roberts, 2006). CBT also overcomes PBT's restriction to non-verbal and paralinguistic modes of communication. Additional symbols replicate the role of non-verbal channels and paralinguistic cues, which suggests that these aspects play an essential role. The use of auditory and visual functions in CBT also enhances the idea of the importance of such cues (Csapó, Ainley, Bennett, Latour, & Law, 2012). There are, however, some concerns being raised regarding CBT, such as access to CBT testing centers, the security of CBT, the reliability of test-delivery systems, the psychometric quality of the tests, and the adequacy of the theoretical models that support its use (Bartram, 2006; Mills, Potenza, Fremer, & Ward, 2005). Kolagari et al. (2018) investigated test anxiety among 39 Iranian nursing students during CBT and PBT. Although the CBT group had higher test anxiety scores than did the PBT, there were no significant differences between the two groups in terms of the mean anxiety score. In the Saudi context, Baig et al. (2018) found a significant difference in the mean anxiety levels for PBT and CBT. Surprisingly, the PBT group had higher test anxiety levels than did the CBT group. The authors attribute this to the health sciences students' familiarity with computers, as there are large computer labs equipped with 200 computers and three-quarters of the students had access to home computers. Baig et al. (2018), however, did not investigate the relationship between the Saudi students' attitudes toward both test formats and their test anxiety level. Another highly influential research factor is student attitudes toward computers, which varies among individuals. Attitudes are important, because they are pervasive. Attitudes can make individuals highly resistant to change and they will be motivated to engage in different types of behavior. We may not always be aware of our attitudes. They can be explicit (controlled & conscious evaluative responses) or implicit (automatic & non-conscious evaluative responses; Deutsch, Herrmann, Frese, & Sandholzer, 2012). Deutsch et al. (2012) found that the attitudes of fourth year medical students towards CBT were overall positive and their anxiety levels decreased. It is, therefore, pertinent to gauge individual student attitudes toward computers before administering CBT. Individuals' attitude can either increase or decrease their test anxiety level. It may be challenging to identify student attitudes, as they can be implicit or explicit. However, according to cognitive dissonance theory, it is also possible for attitudes to be changed by changing behaviors; attitudes are changed as a form of rationalization or self-justification to reduce dissonance (Bandura & Locke, 2003).

Perception is also a key component of self-efficacy. Students with low self-efficacy levels may perceive the procedures involved in CBT to be too complex, even if they are quite straightforward. In this way, low self-efficacy can generate unnecessary and excessive stress, which can in turn lead

to higher levels of test anxiety. On the other hand, students with high self-efficacy levels might engage in self-serving bias, which is the tendency to view oneself favorably when processing self-relevant information. Individuals with high self-efficacy are more likely to perceive themselves as above average and attribute their successes to internal, dispositional causes and their failures to external, situational causes (Duval & Silvia, 2002). Although many studies have demonstrated that different factors contribute to anxiety in both CBT and PBT, it is arguable that students with higher levels of English proficiency and who are more familiar with computers will experience lower levels of anxiety during CBT. However, this theoretical perspective has some limitations. First, very few empirical studies have substantiated this view, and, second, anxiety can be caused by many other factors.

A number of measures for assessment of test anxiety among adolescent students have been developed and commonly used over the last decades, such as Friedman and Bendas-Jacob's (1997) FRIEDBEN Test Anxiety (FTA) scale and Spielberger and Vagg's (1995) Test Anxiety Inventory (TAI), although the constructs underlying these measures are narrowly defined (Lowe & Lee, 2008). Kumke (2008) argues that the development of a test anxiety scale depends on the researcher's conceptualization of test anxiety and its components (or factors). Kumke (2008) defines test anxiety as consisting of cognitive, physiological, and behavioral components. Kumke developed the Adolescent Test Anxiety Scale (ATAS), which is one of the most comprehensive measures of test anxiety among adolescents, because it has high internal consistency and showed evidence for convergent and discriminant validity by its correlations with another measure of test anxiety.

Based on the reviewed literature, the major findings of test anxiety (during CBT and/or PBT) research are summarized as follows:

- Technological advancement has affected test and assessment methods and directed them towards digitalized versions.
- Familiarity with computers and English proficiency levels have a significant relationship with test anxiety; however, confidence, attitude, motivation, sense of self-efficacy, hard work, and personal initiative can also lead to lower levels of test anxiety.
- The integration of CALL decreased student anxiety levels and improved academic performance.
- The attitude that an individual has towards a test can either increase or decrease test anxiety levels during that test.
- Grades, test procedures, environment, length, timing and simplicity of the questions are factors that play an important role in apprehension level during the test.

3 Methodology

3.1 Research questions

The research questions motivating the present research study are as follows:

1. What are the main characteristics of Saudi EFL students' attitudes and experiences toward both testing formats (PBT and CBT)?
2. Is there a relationship between test mode and test anxiety level among Saudi EFL students?
3. Is there a relationship between student attributes and test anxiety levels among Saudi EFL students?

While research question 1 was answered through descriptive statistics gathered from the attitude questionnaire, the latter two were answered through hypothesis testing. The null hypotheses pertaining to the latter two questions are as follows:

Null Hypothesis One [H₀₁]: Regarding the relationship between test mode and test anxiety, students will experience the same anxiety levels with both testing formats.

Null Hypothesis Two [H₀₂]: Regarding the relationship between student attitudes and test anxiety levels, students' attitudes toward CBT and PBT will not have a significant influence on their anxiety level after each test.

3.2 Participants & instruments

The present study employed a quantitative research methodology, which included descriptive statistics (pre- and post-questionnaires) of the findings related to the Saudi students' anxiety levels during CBT and PBT, as well as an in-depth analysis of the students' attitudes toward both testing modes. Furthermore, the present study investigated the language barrier in relation to anxiety. Thus, 40 male Saudi students in their first year of a Bachelor of Arts English program participated in this study, which was divided into three phases (see Fig. 1). The inclusion of only male students was due to the conservative Saudi traditions where students are gender-segregated in educational institutions. Initially, 52 students participated in the first phase of the research; for different reasons and circumstances, however, 9 participants withdrew during the second and/or the third phase(s). Whereas some withdrew after the first phase of the study, others dropped out after the second phase. It should be noted that every participant had to undergo the three study phases. Thus, missing one phase of this experiment leads to the exclusion of the participant from the study. Thus, only 43 students completed all three stages and their data were analysed. During the analysis process, 3 questionnaires were excluded, because the responses were incomplete, inconsistent, or invalid. Thus, the final sample consisted of a total of 40 students. All the students had had at least one CBT experience and long-term experience with PBT. The participants were beginner- or intermediate-level speakers of English with limited production capabilities. Their prior educational experience with English was also limited, as this was a general EFL education class and most had completed only basic English education prior to university. As the students regularly use computers at home to upload their assignments on the university's portal, it was assumed that none of them had any technophobia. The students were sampled from two classes that were instructed by the same tutor and the same courses were taught to both classes. In order to avoid any subject-expectancy effect, the researchers were not involved in tutoring the two classes. Both PBT and CBT were designed with the assistance of the course instructor and the participants were aware that these tests are part of the course requirements and that the results will affect their overall score.

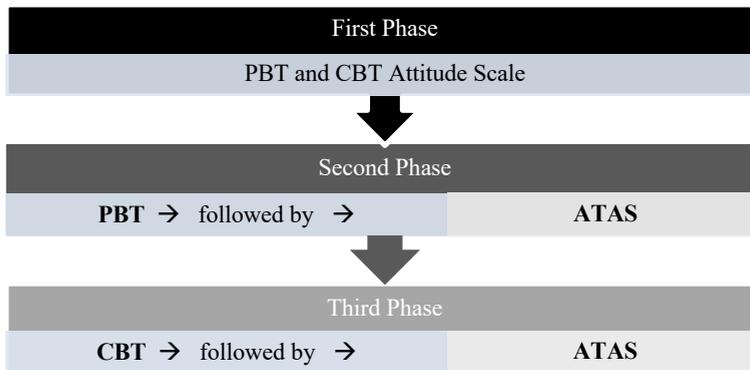


Fig. 1. Phases of the present study

To address the anxiety levels of CBT and PBT among EFL learners, two instruments were used in this study: the PBT and CBT Attitude Scale (Appendix 1) that was specifically developed for the present study and Kumke's (2008) ATAS (Appendix 2). The instruments were piloted by 15 participants to check the comprehensibility of the questions and to point out any possible ambiguities. The two scales were reviewed by two senior linguistics researchers to ensure that they were valid and reliable. Ambiguous questions were modified accordingly.

The PBT and CBT Attitude Scale was administered to all the participants in Phase 1 in order to identify their attitudes toward PBT and CBT. This scale measures which mode a student prefers based on the received grades for each statement. The developed Attitude Scale initially included 55 items that covered almost all aspects related to CBT and PBT. None of the items were adapted from other studies. Accordingly, the items were checked by two senior linguistics researchers for content and reliability measurements and those considered too similar or redundant were removed. For example, the item “I prefer to take the CBT because it develops my typing speed skills” was deleted, because it is synonymous with the first item “the computerized test develops the students’ speed in typing and writing”. Likewise the item “I lose some marks when taking the CBT because I cannot review my answers before submitting the paper” was deleted, because it is similar to statement 6, “I prefer traditional tests because I can review my answers and correct any mistakes.” Consequently, the items were reduced from 55 to only 28 items, taking into account their comprehensiveness and the coverage of the targeted aspects that met the objectives of this study, and consisted of 10 statements favoring PBT, 15 statements favoring CBT, and 3 neutral statements that indicated a student had similar attitudes toward both testing modes. The questionnaire also included some demographic information such as age, university level, and whether they had prior experience with CBT. The participants’ average age was 20.8 years (SD \pm 1.25, range 19–23 years). The 28 items were randomly ordered before being completed by the participants.

The attitudes of students toward both test formats were measured through an attitude factor that was calculated using the following derived formula, where C is the total score for questions that favor CBT, P is the total score for questions favoring PBT, N is the total score for questions 5 and 19, and Q28 is the score for question 28.

$$\text{Attitude Factor} = C - 1.5 * P - N + 2 * Q28$$

$$\text{Preference score CBT} = 75 - 1.5 * 10 - 2 + 10 = +68$$

$$\text{Preference score PBT} = 15 - 1.5 * 50 - 10 + 2 = -68$$

This formula was deduced so that attitude factor values ranged from +68 to -68. A score close to +68 meant that the student preferred CBT, while a score close to -68 indicated a preference for PBT; a score of or close to zero indicated a neutral attitude toward both testing formats. The logical base underlying the derivation of this formula is explained in detail in Appendix 3.

Once the participants filled out the pre-questionnaire, they completed the PBT in Phase 2 which consisted of 20 questions: 10 multiple-choice questions, 5 closed-ended questions that simply required a “yes” or “no” answer, and 5 open-ended questions that required longer answers. All the questions revolved around topics they have already covered during their language lessons. Both the PBT and the CBT are similar in language points, difficulty level, questions types, and the time required to complete the test. However, the content is different, since each test format included the topics students have covered during the lessons. Once all the participants have completed the PBT, they were asked to answer the ATAS.

Phase 3 was conducted a month later, and all the students were asked to complete the CBT. This test was designed with the students' instructor using Typeform (<https://www.typeform.com/>), an interactive web-based platform for designing and preparing tests, surveys, and different types of forms. Several definitions and explanations of the concept of CBT were reviewed, and a number of examples were examined. The conditions for conducting CBT such as test environment, time required, ease, type and duration of questions were taken into consideration. As the study aimed to measure anxiety levels before and during the test, it was not necessary to have a counter-balanced measure design between the PBT and CBT. The anxiety questionnaire was distributed immediately after the test in order to retrieve responses that truly reflect the students’ anxiety state during the test. Thus, once all the participants had completed the CBT, they were directed to a link to answer the ATAS that was also entered into Typeform. The CBT survey aimed to identify the students’ anxiety levels after the CBT. After surveying a number of questionnaires that measure test anxiety (Friedman & Bendas-Jacob, 1997; Sapp, 2013; Spielberger & Vagg, 1995), we decided to adopt Kumke’s (2008) measure of test anxiety among adolescents, ATAS. Although ATAS was developed to measure test anxiety in a sample of 313 high school students, we decided to employ it in our study of university students, since they also belong to the adolescents’ group. As stated earlier,

ATAS is one of the most comprehensive measures designed for this purpose, and it has high internal consistency. The scale also showed evidence for convergent and discriminant validity by its correlations with another measure of test anxiety and a general measure of anxiety. It is practical, meets the requirements of the present study, and can be applied to both test types. Kumke initially designed a 70-item scale, which appears to be a valid and reliable measure. After the expert reviews of this scale, Kumke disposed of 45 items based upon low item redundancy, item-total correlations, and disagreement among the expert reviewers as to what factor the item was measuring. The remaining 25 items were chosen based on the reliability statistics, item content, and the hypothesized underlying factor structure. These items included 16 cognitive aspects and 9 physiological aspects of test anxiety. The final version of ATAS had over 85% agreement among expert reviewers and a Cronbach's alpha of .95 (see Table 1). Given the fact that ATAS was developed as part of a Ph.D. dissertation, this is the first study to use this instrument, to the best of our knowledge.

Both the PBT and CBT Attitude Scale and the ATAS items included a 5-point Likert scale (i.e. strongly agree, agree, neutral, disagree, and strongly disagree). Furthermore, they were translated into Arabic since the students were beginner- or intermediate-level EFL speakers. The translation was conducted using a forward and back translation process to ensure the functional equivalence of the Arabic terms in the translated version of the questionnaire. The aim of the two tests was to gather more meaningful, reliable data to help identify factors contributing to or reducing student anxiety in the context of EFL learning. Since some Saudi universities and colleges are still using the PBT, the inclusion of CBT will help identify the parameters of comparability to better understand different levels of anxiety that occur among different students. Anxious students, especially those who had taken only a limited number of computerized tests prior to the survey or those who always exhibit some discomfort with assessment regardless of its mode, will feel more comfortable expressing their attitudes in the pre-questionnaire.

Statistical analysis of all data was done using the Statistical Package for the Social Sciences (SPSS) software. The data obtained from the questionnaire were treated as ordinal data, that is, median and range were used to summarize ordinal variables, frequencies, and percentages for categorical nominal data. A chi-square (χ^2) goodness-of-fit test was used to test if the distribution of the demographic variables in our sample was similar to the population distribution of that variable. Reliability of the various constructs was tested through Cronbach's alpha coefficient with values above 0.8 considered acceptably reliable. Table 1 summarizes the reliability tests conducted for the two questionnaires. All questionnaires had a Cronbach's alpha coefficient of 0.767 or more, which indicated that the questionnaire items were reliable.

Table 1. Cronbach's alpha coefficients for the questionnaires

Questionnaire	Number of items	Cronbach's alpha coefficients
Attitude (PBT)	10	0.767
Attitude (CBT)	15	0.856
ATAS (PBT)	25	0.950
ATAS (CBT)	25	0.956

Spearman's correlation was used for hypotheses testing and multicollinearity through which a variable can be linearly predicted from others. In addition, a univariate correlation analysis, Kruskal-Wallis ANOVA, and Mann-Whitney U test were used to test some of the hypotheses, as required. All tests were 2-tailed, and p-values of >0.05 were considered statistically insignificant, <0.05 statistically significant, and <0.01 highly statistically significant.

4 Findings

This section presents and discusses the findings of the scale that measured students' attitudes toward PBT and CBT, based on the results of the PBT and CBT ATAS analyses.

4.1 Students' attitudes toward PBT and CBT

The attitude questionnaire was conducted at the beginning of the study to identify the Saudi students' background with, attitudes toward, and perceptions of PBT and CBT. As previously stated, an attitude factor was calculated for each student. Table 2 outlines the results of the attitude questionnaire.

Table 2. Summary statistics for the attitude questionnaire

Construct	Mean [†]	SD	No. of students with Attitude Factor
PBT	53.68	9.74	< -10 = 7
CBT	33.08	6.41	> 10 = 13
Neutral	5.68	2.08	between -10 and 10 = 20

SD=Standard Deviation

[†] Total score of each construct averaged across respondents

Most students expressed neutral attitudes toward both CBT and PBT (50% of students had an attitude factor of around zero). On the other hand, 32.5% and 17.5% of the students preferred CBT and PBT, respectively.

4.2 The relationship of test mode and anxiety level

The average total anxiety score after PBT was 59.08, as reported by the students who responded to our survey, compared to an average total of 56.80 after CBT, as shown in Table 3. Statistical analysis showed that this difference was insignificant ($p = 0.145$). This means that students reported similar anxiety levels after taking either of the two test formats. Similar results were obtained when comparing the mean and median anxiety scores for PBT and CBT.

Table 3. Anxiety score statistics for CBT and PBT

Anxiety Score	Questionnaire	Mean	SD	p -value
Total	PBT	59.08	18.92	0.145 [†]
	CBT	56.80	23.24	
Mean	PBT	2.35	0.75	0.119 [†]
	CBT	2.27	0.93	
Median	PBT	2.20	0.97	0.122 [†]
	CBT	2.05	1.2	

[†] Mann-Whitney U test

Thus, the first null hypothesis [H_01] pertaining to the relationship between test mode and test anxiety was accepted since the students experienced similar anxiety levels with both testing formats.

4.3 Correlation between student attitudes and anxiety levels

The findings of the correlation analysis between student attitude and anxiety level revealed that there was a significant positive correlation between the attitude factor and anxiety level after PBT, as shown in Table 4. The Spearman's correlation for the attitude factor and the anxiety score for PBT was 0.33 (p -value 0.037), while that of CBT was 0.29 (p -value 0.041). This means that the higher the preference of a student for PBT, the higher the anxiety level he experienced during both test formats. Interestingly, there was a significant positive correlation between anxiety level after PBT and anxiety level after CBT (Spearman's $r = 0.668$, p -value = 0.0001), which means that students who experienced high levels of anxiety during PBT were very likely to experience high levels of anxiety during CBT as well.

Table 4. Correlation analysis of students' attitudes and anxiety levels

		Attitude Factor	Anxiety Score PBT	Anxiety Score CBT
Attitude Factor	Pearson Correlation		0.336*	0.290
	p-value		0.037	0.041
Anxiety Score PBT	Pearson Correlation	0.336*		0.668**
	p-value	0.037		0.0001
Anxiety Score CBT	Pearson Correlation	0.290	.668**	
	p-value	0.041	.000	

* .20-.39= "weak"

** .60-.79= "strong"

The second null hypothesis [H₀₂] related to the relationship between student attitudes and test anxiety levels was rejected, since the students' attitudes toward CBT and PBT had significant influence on their anxiety level after PBT or CBT.

5 Discussion

Overall, it was noted that the valid response rate to the survey was satisfactory. Only 3 out of 43 responses were excluded from our analysis, because they were invalid, incomplete, or inconsistent. This is considered a good indicator of the comprehensibility of the survey questions and the ease with which they could be completed. The findings of the attitude factor of the questionnaire indicate that a student had a neutral attitude if he had a score of or near zero. A cut-off value of -10 to +10 was considered to indicate a neutral attitude. However, this cut-off value is empirical and may require validation to ensure that the scores reflect actual attitudes. According to this factor, most students expressed a neutral attitude toward both PBT and CBT, with the remaining students more in favor of CBT than PBT. We note that a sensitivity analysis was performed to elucidate the effect of varying the cut-off value on the overall result. It turned out that narrowing the cut-off range to half (-5 to +5) or widening to (-15 to 15) does not change the overall result that most respondents had a neutral attitude towards both PBT and CBT.

The statistical analysis showed that the test mode did not affect the anxiety level experienced by the students after taking either of the two test formats. This means that the noted anxiety level was independent of the test mode and could be attributed to other factors not investigated in our study. However, there was a significant direct relationship between attitude factor and anxiety level in both test modes. This could be interpreted as students who had high anxiety levels during exams in general tended to choose the traditional approach to exams, as they might have felt that it is safer to avoid changes in test conditions, considering that most students had much more prior experience with PBT than CBT. More important and interesting evidence to support this finding, which is the strong positive correlation between anxiety levels experienced after PBT and CBT, indicates that the students who were highly anxious during PBT were likely to also be anxious during CBT, and vice versa. All these findings suggest that test mode has nothing to do with test anxiety, and thus there are other factors that contribute to test anxiety. Such factors were investigated by Al-Amri (2009), who reported that language proficiency, gender, and motivational factors were the most important predictors of test anxiety and, by extension, test performance.

The findings of the present study are consistent with those reported by Parshall, Spray, Kalohn, and Davey (2002), who contended that no difference in anxiety level between PBT and CBT could be proven. Along similar lines, Toth (2010) argued that English proficiency, not test mode, is what plays the most important role in the development/removal of anxiety. Toth (2010), whose study took place in the Hungarian context, argued that test anxiety was instead related to individual, internal factors (i.e. cognitive, affective, personality).

However, these findings contrast with those reported by Rahimi and Yadollahi (2011), who studied the relationship between overall success in EFL learning, computer anxiety, and computer ownership. The researchers reported that there was a negative correlation between computer anxiety and success in EFL learning. This means that, according to the authors, the more familiar the student is with computers, the more likely he or she is to succeed in learning English. This inconsistency between our findings and Rahimi and Yadollahi's (2011) findings could be attributed to the difference in endpoints measured. While our study focused on anxiety following both types of tests, Rahimi and Yadollahi's study focused on students' exam performance represented by their official scores. These two endpoints are not necessarily correlated, and therefore, the two studies may not be comparable.

6 Conclusion

This study examined students' attitudes toward PBT and CBT and the relationship between test formats and levels of student anxiety. The present study is the first to produce a new concept for measuring students' attitudes toward PBT and CBT, that is, the attitude factor. This factor is an innovative, logically plausible factor used to determine if the students prefer PBT or CBT based on their responses to various questions in the validated attitude questionnaire. The findings of this study demonstrated that the test mode did not significantly affect student anxiety levels and that the anxiety levels for both testing formats were significantly correlated. An important finding was that the students' anxiety levels after PBT and CBT were positive; that is, their preferences did not affect their anxiety levels after taking either of the two test formats. This means that the noted anxiety level was independent of the test mode and could be attributed to other factors not investigated in our study. Finally, the results of the study showed that most students did not have any preference for PBT or CBT; that is, they had neutral attitudes. However, of those who preferred a specific format, most students demonstrated a preference for CBT. The analysis of the ATAS, which was conducted after the PBT, revealed that there was no statistically significant difference in the students' anxiety levels after the PBT, compared to the CBT. While both CBT and PBT can be actively used in the classroom as testing methods or assessment tools, students need to feel comfortable in their learning environment without feeling threatened by being assessed or evaluated. This will further help minimize anxiety levels among them.

Based on the study results, EFL/ESL course instructors should embrace the tendency to use CBT in their courses. However, since the present study included only male students, those teachers with female students ought to be cognizant of findings in other studies, namely, that females have been found to experience higher levels of test anxiety (See e.g. Baig et al., 2018). The findings of the present study cannot be generalized to all students in all subject areas, since they are limited to EFL students of the same age, gender, and proficiency level, and the study was restricted to one academic subject area. Therefore, further studies should encompass diverse student populations, different genders, and other subject areas. Furthermore, future studies need to examine other potential causes of student anxiety such as language proficiency, learner self-perceptions, familiarity with computers, previous exposure to English, and classroom environment, which might increase or reduce students' anxiety and apprehensiveness in the testing or assessment process.

The present study has a number of limitations. First, a convenience sample of 40 participants is not sufficient to generalize about the experiences of most students. Furthermore, although the study is related to EFL learners, one might argue that the outcome would be different, if the participants had been from other disciplines, such as IT, mathematics, psychology, or sociology. In addition, gender is another variable that may have affected the findings of the study, as gender has been shown to be an important predictor of computer anxiety (Deryakulu & Çalışkan, 2012). The results, however, can be generalized with caution to all the Saudi students in the preparatory stage at the university. This study represents a model for other researchers who can reproduce its procedure to apply to other population subgroups. Moreover, the age range of the surveyed sample was limited to 19–23 years. This age range, however, represents the actual age range of the target population that this

study aimed to investigate. Therefore, the age factor may not limit the generalizability of the findings. Finally, a further limitation appears in the choice of closed-ended Likert scale responses, rather than open-ended responses. This was a potential difficulty for those participants who were willing to complete or add some points to the survey.

Acknowledgments

The authors are indebted to the two anonymous reviewers for their insightful and helpful comments. The authors also wish to express their gratitude to both the Deanship of Scientific Research at King Saud University and the Research Centre at the Faculty of Arts for funding the present study. They also thank RSSU for their technical support.

References

- Al-Amri, S. S. (2009). *Computer-based testing vs paper-based testing: establishing the comparability of reading tests through the evolution of a new comparability model in a Saudi EFL context* (Unpublished PhD dissertation). The University of Essex, Colchester, UK. Retrieved from <http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.496244>
- Al Shammari, M. H. (2007). *Saudi English as a foreign language learners' attitudes toward computer-assisted language learning* (Unpublished EdD dissertation). West Virginia University, Morgantown, USA. Retrieved from www.proquest.com
- Baig, W. A., Al-Zahrani, E. M., Al-Jubran, K. M., Chaudhry, T., & Qadri, A. A. (2018). Evaluation of test anxiety levels among preparatory year students of PSMCHS during computer-based versus paper-and-pen examination. *Int J Med Res Health Sci*, 7(1), 48–52.
- Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, 88(1), 87–99. doi: <http://dx.doi.org/10.1037/0021-9010.88.1.87>
- Bartram, D. (2006). Testing on the Internet: Issues, challenges and opportunities in the field of occupational assessment. In D. Bartram & R. K. Hambleton (Eds.), *Computer-based testing and the Internet: Issues and advances* (pp. 13–37). Chichester: Wiley.
- Bensoussan, M. (2012). Alleviating test anxiety for students of advanced reading comprehension. *RELC Journal*, 43(2), 203–216. doi: <http://dx.doi.org/10.1177/0033688212449511>
- Chen, H. (2012). The moderating effects of item order arranged by difficulty on the relationship between test anxiety and test performance. *Creative Education*, 3(3), 328–333. doi: <http://dx.doi.org/10.4236/ce.2012.33052>
- Chen, K. T.-C. (2012). Elementary EFL teachers' computer phobia and computer self-efficacy in Taiwan. *TOJET: The Turkish Online Journal of Educational Technology*, 11(2), 100–107.
- Chuah, S. C., Drasgow, F., & Roberts, B. W. (2006). Personality assessment: Does the medium matter? No. *Journal of Research in Personality*, 40(4), 359–376. doi: <https://doi.org/10.1016/j.jrp.2005.01.006>
- Conti-Ramsden, G., Durkin, K., & Walker, A. J. (2010). Computer anxiety: A comparison of adolescents with and without a history of specific language impairment (SLI). *Computers & Education*, 54(1), 136–145. doi: <https://doi.org/10.1016/j.compedu.2009.07.015>
- Csapó, B., Ainley, J., Bennett, R. E., Latour, T., & Law, N. (2012). Technological issues for computer-based assessment. In P. Griffin, B. McGaw & E. Care (Eds.), *Assessment and teaching of 21st century skills* (pp. 143–230). Dordrecht: Springer.
- Deryakulu, D., & Çalışkan, E. (2012). A twin study of computer anxiety in Turkish adolescents. *Cyberpsychology, Behavior, and Social Networking*, 15(4), 212–218. doi: <https://doi.org/10.1089/cyber.2011.0499>
- Deutsch, T., Herrmann, K., Frese, T., & Sandholzer, H. (2012). Implementing computer-based assessment: A web-based mock examination changes attitudes. *Computers & Education Journal*, 58(4), 1068–1075. doi: <https://doi.org/10.1016/j.compedu.2011.11.013>
- Duval, T. S., & Silvia, P. J. (2002). Self-awareness, probability of improvement, and the self-serving bias. *Journal of Personality and Social Psychology*, 82(1), 49–61. doi: <http://dx.doi.org/10.1037/0022-3514.82.1.49>
- Friedman, I. A., & Bendas-Jacob, O. (1997). Measuring perceived test anxiety in adolescents: A self-report scale. *Educational and Psychological Measurement*, 57(6), 1035–1046

- Ghaderi, M., Mogholi, M., & Soori, A. (2014). Comparing between computer based tests and paper-and-pencil based tests. *International Journal of Education & Literacy Studies*, 2(4), 36–38. doi: <http://dx.doi.org/10.7575/aiac.ijels.v.2n.4p.36>
- Hardre, P. L., Crowson, H. M., Xie, K., & Ly, C. (2007). Testing differential effects of computer-based, web-based and paper-based administration of questionnaire research instruments. *British Journal of Educational Technology*, 38(1), 5 – 22. doi: <https://doi.org/10.1111/j.1467-8535.2006.00591.x>
- Hensley, K. K. (2015). *Examining the effects of paper-based and computer-based modes of assessment on mathematics curriculum-based measurement* (Unpublished PhD dissertation). The University of Iowa, Iowa City, USA.
- Huberty, T. J. (2009). Test and performance anxiety. *Principal Leadership Journal*, 10(1), 12–16 .
- Kolagari, S., Modanloo, M., Rahmati, R., Sabzi, Z., & Ataee, A. J. (2018). The effect of computer-based tests on nursing students' test anxiety: A quasi-experimental study. *Acta Informatica Medica*, 26(2), 115. doi: <http://dx.doi.org/10.5455/aim.2018.26.115-118>
- Kumke, P. J. (2008). *The Adolescent Test Anxiety Scale: Initial scale development and construct validation*. (Unpublished Ph.D. dissertation). University of Wisconsin, Madison, USA. Retrieved from <https://proquest.com> (UMI Number: 3314217)
- Liu, H.-J., & Cheng, S.-H. (2014). Assessing language anxiety in EFL Students with varying degrees of motivation. *Electronic Journal of Foreign Language Teaching*, 11(2), 285–299 .
- Lowe, P. A., & Lee, S. W. (2008). Factor structure of the Test Anxiety Inventory for Children and Adolescents (TAICA) scores across gender among students in elementary and secondary school settings. *Journal of Psychoeducational Assessment*, 26(3), 231–246. doi: <https://doi.org/10.1177/0734282907303773>
- Ma, W. W.-K., Andersson, R., & Streith, K.-O. (2005). Examining user acceptance of computer technology: An empirical study of student teachers. *Journal of Computer Assisted Learning*, 21(6), 387–395. doi: <https://doi.org/10.1111/j.1365-2729.2005.00145.x>
- Mills, C. N., Potenza, M. T., Fremer, J. J., & Ward, W. C. (2005). *Computer-based testing: Building the foundation for future assessments*. New York & London: Routledge.
- Parshall, C. G., Spray, J. A., Kalohn, J. C., & Davey, T. (2002). *Practical considerations in computer-based testing*. New York: Springer Science & Business Media.
- Piaw, C. Y. (2012). Replacing paper-based testing with computer-based testing in assessment: Are we doing wrong? *Procedia - Social and Behavioral Sciences*, 64, 655–664. doi: <https://doi.org/10.1016/j.sbspro.2012.11.077>
- Rahimi, M., & Yadollahi, S. (2011). Success in learning English as a foreign language as a predictor of computer anxiety. *Procedia Computer Science*, 3, 175–182. doi: <https://doi.org/10.1016/j.procs.2010.12.030>
- Sapp, M. (2013). *Test anxiety: Applied research, assessment, and treatment interventions*. New York, NY: University Press of America.
- Shi, F. (2012). Exploring students 'anxiety in computer-based oral English test. *Journal of Language Teaching & Research*, 3(3), 446–451. doi: <https://doi.org/10.4304/jltr.3.3.446-451>
- Spielberger, C. D., & Vagg, P. R. (1995). *Test anxiety: Theory, assessment, and treatment*. Washington, DC: Taylor & Francis.
- Sulistyaningsih, E. (2016). *Student anxiety facing computer based test (CBT) system of national examination* . Paper presented at the 3rd. International Conference on Research, Implementation and Education of Mathematics & Science, Yogyakarta, Indonesia.
- Tella, A., & Bashorun, M. (2012). Attitude of undergraduate students towards computer-based test (CBT): a case study of the University of Ilorin, Nigeria. *International Journal of Information and Communication Technology Education (IJICTE)*, 8(2), 33–45. doi: <http://dx.doi.org/10.4018/jicte.2012040103>
- Toth, Z. (2010). *Foreign language anxiety and the advanced language learner: A study of Hungarian students of English as a foreign language*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment*, 10(1), 1–17. doi: https://doi.org/10.1207/s15326977ea1001_1
- Zeidner, M. (1998). *Test anxiety: The state of the art*. New York: Kluwer Academic Publishers.

Appendices

Appendix 1

PBT and CBT Attitude Scale (28-Item English Version)

DIRECTIONS:						
This questionnaire lists a series of sentences that tries to understand your background and attitude toward paper-based tests and computer-based tests. Remember, there is no right or wrong answer. The most important thing is that you answer each item truthfully. Thank you for your cooperation.						
So, kindly follow these instructions:						
1- Read the questionnaire items carefully.						
2- Circle the number that reflects your point of view.						
3- Check the accuracy of the information provided and answer every item.						
Age:						
University level Have you ever completed a computer-based test? YES () NO ()						
#	Items	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1.	The computerized test develops the students' speed in typing and writing ^C .	1	2	3	4	5
2.	The student is accustomed to the traditional paper-based method from a young age ^P .	1	2	3	4	5
3.	Students write faster during traditional paper-based tests ^P .	1	2	3	4	5
4.	Students must be familiar with computer basics before taking a computerized test ^P .	1	2	3	4	5
5.	There is no difference in my performance whether I take a computerized test or a traditional test ^N .	1	2	3	4	5
6.	I prefer traditional tests because I can review my answers and correct any mistakes ^P .	1	2	3	4	5
7.	Reviewing students' test answers is more accurate if a computer is used in the marking process ^C .	1	2	3	4	5
8.	The tutor spends a lot of time correcting student papers in traditional paper tests ^C .	1	2	3	4	5
9.	Students get their grades quicker in a computerized testing system ^C .	1	2	3	4	5
10.	Students wait longer to get their scores if traditional paper tests are used ^C .	1	2	3	4	5
11.	The use of a computerized testing system develops students' learning skills ^C .	1	2	3	4	5
12.	I prefer to take traditional paper tests because the question sheets are separate from the answer sheets ^P .	1	2	3	4	5
13.	Traditional paper tests allow me to move between questions easily ^P .	1	2	3	4	5
14.	I make spelling mistakes when I take traditional paper-based tests ^C .	1	2	3	4	5
15.	It is easier to read questions on a computerized test than a paper-based test ^C .	1	2	3	4	5
16.	It requires a lot of training to take a computerized test ^P .	1	2	3	4	5
17.	Taking a computerized test is easier than taking a paper-based test ^C .	1	2	3	4	5
18.	Taking computerized tests improves my computer skills ^C .	1	2	3	4	5
19.	In my point of view, there isn't any difference between computerized tests and traditional tests ^N .	1	2	3	4	5
20.	Writing in traditional tests is easier than in computerized tests ^P .	1	2	3	4	5
21.	For me, computerized tests are more flexible than traditional tests ^C .	1	2	3	4	5
22.	Computerized tests combine the measurement of information and the development of computer skills ^C .	1	2	3	4	5
23.	I can organize my time better during traditional tests ^P .	1	2	3	4	5
24.	Computerized tests allow me to move between test questions easily ^C .	1	2	3	4	5
25.	The environment of the computer lab is better than the environment of the traditional classroom to take a test ^C .	1	2	3	4	5
26.	Traditional tests help me to improve my handwriting ^P .	1	2	3	4	5
27.	I might lose marks on a traditional test if the tutor cannot read my handwriting ^C .	1	2	3	4	5
28.	My way of studying for the exam differs depending on the test type (computerized or traditional) ^N .	1	2	3	4	5

^C Favoring= CBT ^P Favoring= PBT ^N =Neutral

Appendix 2

Adolescent Test Anxiety Scale, Adapted from Kumke (2008)

DIRECTIONS:						
This questionnaire lists a series of sentences that describe how you thought, acted, felt, and/or reacted to the test you have just finished. Remember, there is no right or wrong answer! The most important thing is that you answer each item truthfully. Thank you for your cooperation.						
So, kindly follow these instructions:						
1- Read the questionnaire items carefully.						
2- Please circle the number that best describes your state during the test.						
Age:						
University level						
#	Items	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1.	I doubted my abilities to do well on this test.	1	2	3	4	5
2.	My mind went blank when took this test.	1	2	3	4	5
3.	My palms sweated when I took this test.	1	2	3	4	5
4.	My heart raced when I took this test.	1	2	3	4	5
5.	Worrying about doing poorly interfered with my performance on the test.	1	2	3	4	5
6.	I got very worried just before sitting for the test.	1	2	3	4	5
7.	I could not sleep the night before the test.	1	2	3	4	5
8.	I found the test to be very difficult.	1	2	3	4	5
9.	I could not focus when I began the test.	1	2	3	4	5
10.	I sweated a lot when I took this test.	1	2	3	4	5
11.	I was worried that everyone else was doing better than me.	1	2	3	4	5
12.	My hands were shaking during the test.	1	2	3	4	5
13.	I could not concentrate during the test.	1	2	3	4	5
14.	I felt like I was going to pass out when taking the test.	1	2	3	4	5
15.	I thought I will fail the test when it was over.	1	2	3	4	5
16.	I felt like I will fail the test when taking it.	1	2	3	4	5
17.	I got a headache while taking the test.	1	2	3	4	5
18.	I lost confidence when I realized that I got one (or more) item(s) wrong.	1	2	3	4	5
19.	I did not feel I was doing well on the test.	1	2	3	4	5
20.	I could not stop worrying about the test when it was over.	1	2	3	4	5
21.	I have had a hard time breathing when I was taking the test.	1	2	3	4	5
22.	I felt like I forget everything when I was taking the test.	1	2	3	4	5
23.	I became very uneasy when the teacher announced the start of the test,.	1	2	3	4	5
24.	My muscles tensed up during the test.	1	2	3	4	5
25.	I got so nervous on the test that I felt I have made careless errors.	1	2	3	4	5

Appendix 3

The logical construct underlying the derivation of the attitude factor formula

The attitude factor formula was derived based on the following logical statement:

Attitude score = Preference Score of CBT – Preference Score of PBT.

As all the questions within the questionnaire have been validated, this means that the sum of scores of each construct reflects the preference of a respondent towards that test mode. Therefore, if a student prefers CBT, for instance, then his scores in the CBT construct will be high while in the PBT construct will be low. This

means that the difference between CBT and PBT scores will indicate both the direction and magnitude of preference. This logically means if the questionnaire is validated then the above statement is correct, which is true in our case. However, the above statement assumes that the number questions in both constructs are equal which is not true. The CBT construct has got 1.5 times more questions than the PBT (15 for CBT versus 10 for PBT). This necessitates weighting of the equation term by multiplying the PBT score by 1.5. The equation becomes:

$$\text{Attitude score} = \text{Preference Score of CBT} - 1.5 * \text{Preference Score of PBT}$$

The last two terms of the final equation were derived as the following. The questionnaire includes 3 questions that indicate neutral attitude. As these questions' scores should not count with the CBT or the PBT constructs, the overall contribution of these three questions towards the above equation should be zero (if the respondent is actually neutral), which means they should be added to the equation in a way that they cancel each other: sum of scores of 2 questions – 2 * one question score. The equation then becomes:

$$\text{Attitude score} = \text{Preference Score of CBT} - 1.5 * \text{Preference Score of PBT} - N + 2 * Q28$$

In this way, the attitude score summarizes the responses to all questions of the questionnaire in a single number that is high (close to +68) if the respondent prefers CBT, low (close to -68) if the respondent prefers PBT and gets closer to zero as the respondent has more neutral attitude.