Are the Keyboards Weightier than the Biros? The Effect of Computer-Based Testing on Students’ Achievement and Anxiety in Mathematics

Adeneye O. A. Awofala
Department of Science and Technology Education, University of Lagos
aawofala@unilag.edu.ng

Abstract:

Are the keyboards weightier than the biros? This question is the focus of the present study in which the investigation of the modes of testing in mathematics (computer-based testing-CBT and paper-based testing-PBT) was carried out using mathematics achievement and mathematics anxiety as dependent measures. Through a pre-test post-test experimental research, the study tested two null hypotheses at 0.05 level of significance using an independent samples t-test. The study lasted for four weeks in two Nigerian senior secondary schools. Results showed that there was no statistically significant difference between the mean mathematics achievement scores of students tested by the computer-based test and the paper-based test. Also, there was no significant difference between the mean mathematics anxiety scores of students tested by the computer-based test and the paper-based test. The implication of these results is that assessing students using the two modes showed no variations in their achievement and anxiety in mathematics. Thus, the mode of testing showed no differential effect on students’ learning outcomes in mathematics. Based on the findings of this study, government at levels in Nigeria need to put in place the enabling environment and infrastructural facilities needed for schools to benefit fully from the advantages inherent in the CBT mode of assessment.

Keywords: Keyboard, Biro, Computer-based testing, Paper-based testing, Achievement, Anxiety, Mathematics
Introduction

All over the world, the 21st century digital renaissance has enabled teachers at all levels of the education system to embark on precise and resourceful assessment of students’ learning outcomes. Recently, the Nigerian tertiary institutions have instituted a policy that integrates online learning and web-based assessment into their postgraduate and undergraduate programmes in conformity with international best practices. The Unified Tertiary Matriculation Examination (UTME) into tertiary institutions in Nigeria conducted by the Joint Admission and Matriculation Board (JAMB) is now being done using the computer starting from 2013 and the Post-UTME conducted by the various tertiary institutions is now computer-based. The 2013 version of UTME was distinctive as prospective tertiary institution students were offered three alternatives- Paper-Based Testing (PBT), Computer-Based Testing (CBT) and Dual-Based Testing (DBT). Out of 1,735,720 prospective tertiary institution students that registered for the UTME, 1,629,102 students representing 93.86% opted for PBT, 15,008 students representing 0.86% elected DBT and 91,610 students representing 5.28% chose CBT (Ogunmakin & Osakuade, 2014). Thus, there was an enormous difference between the number of students that chose PBT and CBT and this is an obvious indication that large proportion of the students avoided the CBT. The reasons were not far-fetched as some of the students indicated their anxiety toward computers use (technophobia) and lack of adequate knowledge in operating computers as responsible for their avoidance of the CBT. They reasoned that their usage of the CBT with these deficiencies could lead to lower scores in the 2013 edition of the UTME than with the PBT. It is earnestly expected that the Nigerian institutions of learning will progressively adopt online or web-based assessment (WBA) or computer-adaptive testing (CAT) or computer-based testing (CBT) mode as part of their assessment systems.

The mounting curiosity in the CBT has engendered alterations in assessment methodologies that showcase real modifications in pedagogical and educational approaches (OECD, 2010). CBT is defined as an assessment mode that makes use of computers in the preparation, administration, and scoring of questions and recording and showcasing of students’ scores in a test or examination. In Nigeria, the CBT features mostly at the tertiary level and the feasibility of adopting the CBT at the pre-tertiary levels is currently an issue of discussion despite the introduction of computer studies as a subject in primary and secondary schools in 2008 and 2011 respectively (Awofala & Sopekan, 2013). The CBT is presently considered a promoter of change, which inevitably brings innovations in education, teaching, assessment, and curricula (Scheuerman & Pereira, 2008). Globally, the CBT is implemented in schools as a mode of assessment for several reasons. Teachers at all levels of the educational system can minimize the time lag in broadcasting students’ scores, accelerate and intensify the efficacy and adeptness of assessment, attain suppleness in time and place, offer instantaneous response and students’ scores are captured straightaway, eliminate scores-entry errors, improve scores management, implement individually customized assessment to students’ instructional needs and make the assessment environment paperless thereby reducing paper usage and cost (Noyes & Garland, 2008; Öz & Özturan, 2018; Cresswell, Schwantner, & Waters, 2015).

It is worthwhile to note that while the test administrators are in general favourably disposed to the implementation of the CBT in schools, the test takers have inconsistent dispositions towards its adoption (Coniam, 2006). This inconsistency can be ascribed to test takers’ familiarity and accessibility to computers and test type itself (Coniam, 2006). Test takers who have access to and are familiar with computers will find the CBT interesting and be keen towards it while those who are less familiar with and have no access to computers will be less keen and show low favourable disposition towards the CBT. In the same vein, test takers will respond meaningfully to the CBT involving multiple-choice items than the CBT involving test takers to type in missing words and phrases (Coniam, 1999). Thus, this low favourable disposition towards the CBT can make test takers to prefer the paper-based testing (PBT) version of the test to the CBT. There are two guiding principles for the two modes of testing (CBT and PBT) as explicated in the literature (Fulcher, 1999). First, both the CBT and the PBT should rank test takers comparably and equivalently in which the basic statistics of the two modes in terms of means and standard deviations should be similar or equivalent. Second, statistical association between the CBT and the PBT should be as high as 0.9 (Jones, Kuijper & Verschoor, 2001) so that preference for a particular mode of testing (e.g CBT) will not jeopardize test takers non-preference for the other (e.g. PBT).

Studies on test delivery modes have been inconsistent. For instance, OECD (2010) recounted that there was no statistically significant main effect of test delivery mode on test performance of test takers (n=5,875) from Korea, Iceland and Denmark. Also, Chua (2012) indicated that there was no significant effect of test
delivery mode (CBT vs. PBT) on psychological test performance after controlling for the effects of testing. Chua and Don (2013) found that testing mode had no statistically significant effect on both psychological test and achievement test scores. More recent study by Öz and Özturan (2018) found that there was no statistically significant effect of test delivery mode on students' test scores of English as a foreign language. Contrastingly, Bunderson, Inouye, and Olsen (1989) recounted that 52% of prior studies indicated disparities amid the CBT and PBT modes, in which 13% obtained greater scores for the CBT while the remaining 39% attained greater scores for the PBT. The equivalence of the CBT and the PBT was reinforced by 48% of the studies with significance differences established in mathematics, sciences, and languages achievement tests (Friedrich & Bjornsson, 2008; DeAngelis, 2000; Choi, Kim, & Boo, 2003; Mazzeo, Druesne, Raffeld, Checketts, & Muhlstein, 1991). In explaining these results Chua and Don (2013) noted that the CBT as an evaluation instrument for psychological and educational testing in institution of learning had a low validity and that the possibility of some extraneous variables (e.g. testing effect) biasing the effects of test delivery mode on test performance was high. A distorting variable that could confound the effect of test delivery mode is testing effect (Yu & Ohlund, 2010), defined as the effect of taking a pretest on the outcome of a posttest. Testing effect works to falsify the treatment effect of the CBT on test performance (Yu & Ohlund, 2010) and so should be eliminated in future studies. Few studies have investigated the effect of the CBT on students' mathematics anxiety. Mathematics anxiety is defined as a performance-based anxiety disorder that involves physiological stimulation, negative cognitions, and avoidance behaviours that lead to an affective drop-in mathematics and mathematics-related activities (Awofala & Odogwu, 2017). Two dimensions of mathematics anxiety currently exist in the literature: cognitive and affective components (Awofala & Akinoso, 2017). The cognitive dimension relates to worry and irrelevant thinking in mathematics while the affective dimension relates to emotionality and bodily indicators. While studies have consistently shown that there exist an inverse relationship between mathematics achievement and mathematics anxiety (Awofala & Akinoso, 2017; Jansen, Louwerse, Straatemeier, Van der Ven, Klinkenberg, & Van der Mass, 2013) evidence suggests that mathematics anxiety stems from negative attitude towards mathematics, uneasiness with testing milieu, and negative sensitivity about technology adoption (Awofala & Akinoso, 2017; Tatar, Zegin & Kagizmanli, 2015). However, the effect of testing modes on students' anxiety and achievement is mixed. Coniam (2006) found that the CBT enhanced students' learning outcomes better than the PBT mode of assessment while Hosseini, Abidin, & Baghdarnia (2014) reported the efficacy of the PBT over the CBT on students’ achievement. Whereas some researchers have found no statistically significant effect of testing modes (PBTvs. CBT) on students’ learning outcomes (Kolagari, Modanloo, Rahmati, Sabzi, & Ataee, 2018; Khoshshima, Hosseini & Hasheni Toroujeni, 2017). Thus, there is a need to ascertain the efficacy of the CBT mode of assessment on students’ mathematics achievement and mathematics anxiety.

Currently, there is a growing interest in the implementation of computer-based testing in Nigeria as educational institutions move towards paperless classrooms. However, this study intended to add to the literature the experience regarding the enactment and assessment of the CBT in Nigeria. Thus, this study investigated the effect of test delivery mode (CBT vs. PBT) on senior secondary school students' achievement and anxiety in mathematics.

**Research Hypotheses**

**Ho1:** There is no significant difference between the mean mathematics achievement scores of students tested by the computer-based test and the paper-based test.

**Ho2:** There is no significant difference between the mean mathematics anxiety scores of students tested by the computer-based test and the paper-based test.

**Methods**

**Research Design**

This study adopted a pre-test, post-test true experimental research design in which participants were randomly assigned to the groups in order to accurately segregate and quash any irritant or confusing variables.
Participants

The respondents were made up of 60 students (30 males and 30 females) from two Nigerian senior secondary school. The sample was predominantly Yoruba- an ethnic nationality found in the South-West Nigeria. The participants’ average age was 16.5 years (SD=0.88) and were all registered for a mathematics course. The respondents showed similar acquaintance with information technologies as reflected in the self-reported survey regarding computer self-efficacy they were requested to complete. The respondents were arbitrarily allocated into two clusters: 30 participants to the CBT and 30 participants to the PBT. The CBT group utilized computers while the PBT group utilized paper and pencil to provide answers to twenty (20) personal evaluation tasks handed down to them within a period of four weeks.

Instruments

Two assessment instruments were used for data collection in this study: Mathematics Achievement Test (MAT) and Mathematics Anxiety Questionnaire (MAQ). The MAT was developed by the researcher and consisted of 30 multiple choice questions covering number and numeration, algebraic processes, trigonometry, probability and statistics, and calculus of the senior secondary school year three mathematics curriculum. The MAT has been validated and enunciated in a table of specification (Table 1) covering the lower-order cognitive domain of knowledge, comprehension and application. The MAT captured conceptual knowledge and were similar for the CBT and PBT in which the interface was the same between the two modes of testing. The average difficulty index of the MAT was 45% and all items showed adequate discrimination indices of being greater than 0.3 (Akinsola & Awofala, 2009). The reliability coefficient of the MAT using Kuder-Richardson formula 20 was 0.88. The MAQ was used to assess participants’ mathematics anxiety and was adopted from Wigfield and Meece (1988) and had been validated for Nigerian use (Awofala & Akinoso, 2017). The MAQ (Cronbach α=0.89) (Awofala & Akinoso, 2017) consisting of 11 items on a modified 5-points Likert type scale (0-undecided, 1-strongly disagree, 2-disagree, 3-agree, and 4-strongly agree) has two dimensions worry (Cronbach α=0.80) and emotionality (Cronbach α=0.90) (Awofala & Akinoso, 2017).

Table 1. Mathematics Achievement Test (MAT) Item Specification

<table>
<thead>
<tr>
<th>Content</th>
<th>Knowledge</th>
<th>Cognitive Levels</th>
<th>Application</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number and numeration</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Algebraic processes</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Trigonometry</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Probability and statistics</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Calculus</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>30</td>
</tr>
</tbody>
</table>

Procedure

At the commencement of the investigation, both the MAT and the MAQ were administered on the participants as pretest and through simple randomization the participants were assigned to the CBT (n=30) or the PBT (n=30) groups. During the testing, the experimental group was trained on the assessment administration delivered via the computer local network. In administering the CBT, participants registered with the website using already created passwords and usernames. After registering with the website, each participant in the CBT mode went through the test instructions and started the test with a countdown timer located at the screen top showing the residual time. During the same time the PBT mode took the test. The MAT items for the CBT and the PBT were the same in area of content, difficulty of questions, number of questions, and time allotted for dealing with the questions. For the CBT mode, the question ordering and options were presented arbitrarily to the participants to eliminate examination malpractice.

Data Analysis

The data collected were first reviewed for normality distribution using Kolmogorov-Smirnov test and the test showed normality of data in respect of the achievement and the anxiety (p>0.05). Raw data were coded into SPSS version 22 and independent samples t-test was used for the analysis of the data in respect of achievement and anxiety to test the null hypotheses at 0.05 level of significance. Prior to analyses of posttreatment data, the two modes of testing groups were comparable in the area of prior mathematics
achievement and prior mathematics anxiety. An independent samples t-test showed no significant difference on mathematics achievement between the CBT (Mean=10.70, SD=1.34, Std. Error Mean (SEM) =0.25) and the PBT (Mean=11.17, SD=1.76, SEM=0.32) modes prior to treatment (t_{58}=1.15, p=0.25). Also, there was no significant difference on prior mathematics anxiety (t_{58}= 0.10, p=0.92) between the CBT (Mean=2.23, SD=1.28, SEM=0.23) and the PBT (Mean=2.20, SD=1.42, SEM=0.26) modes of testing.

**Results**

**Null Hypothesis One:** There is no significant difference between the mean mathematics achievement scores of students tested by the computer-based test and the paper-based test.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>SEM</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBT</td>
<td>30</td>
<td>15.80</td>
<td>2.17</td>
<td>0.40</td>
<td>58</td>
<td>0.06</td>
<td>.95</td>
</tr>
<tr>
<td>PBT</td>
<td>30</td>
<td>15.83</td>
<td>2.17</td>
<td>0.40</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 above showed the independent samples t-test analysis of the mathematics achievement based on testing mode. The results showed that the CBT group pooled a mean score (Mean=15.80, SD=2.17) which was slightly lower than the mean score of the PBT group (Mean=15.83, SD=2.17) and this slight difference in mean score was statistically not significant (t_{58}=0.06, p=0.95) at 0.05 level of significance. Thus, we do not reject the hypothesis which states that there is no significant difference between the mean mathematics achievement scores of students tested by the computer-based test and the paper-based test.

**Null Hypothesis Two:** There is no significant difference between the mean mathematics anxiety scores of students tested by the computer-based test and the paper-based test.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>SEM</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBT</td>
<td>30</td>
<td>2.83</td>
<td>1.09</td>
<td>0.19</td>
<td>58</td>
<td>1.15</td>
<td>.26</td>
</tr>
<tr>
<td>PBT</td>
<td>30</td>
<td>3.13</td>
<td>0.94</td>
<td>0.17</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3 above showed the independent samples t-test analysis of the mathematics anxiety based on testing mode. The results showed that the CBT group pooled a mean score (Mean=2.83, SD=1.09) which was slightly lower than the mean score of the PBT group (Mean=3.13, SD=0.94) and this slight difference in mean score was statistically not significant (t_{58}=1.15, p=0.26) at 0.05 level of significance. Thus, we do not reject the hypothesis which states that there is no significant difference between the mean mathematics anxiety scores of students tested by the computer-based test and the paper-based test.

**Discussion**

Two findings emerged in the present study. First, the present study showed that there was no significant main effect of testing mode on senior secondary school students’ achievement in mathematics although the CBT group recorded slightly lower mean score than the PBT group in mathematics achievement. This result was in agreement with the findings of Kolagari et al. (2018) and Khoshsima et al. (2017) which showed that there was no significant difference between the mean achievement scores of students tested by the computer-based test and the paper-based test. However, the result disagreed with the findings of Nikou and Economides (2016) and Chua and Don (2013) who in their separate studies reported significant effect of testing mode on students’ learning achievement and concluded that the CBT should be deployed as auspicious substitute to the PBT assessment techniques. Still disagreeing with present study result, Hosseini et al. (2014) reported the efficacy of the PBT over the CBT on students’ achievement. The present study result is anticipated because the two testing modes were similar in item content, length, and sequence. Second, there was no significant main effect of testing mode on senior secondary school students’ mathematics anxiety although the CBT group recorded slightly lower anxiety mean score than the PBT group. This result conformed to the findings of Kolagari et al. (2018) which found that there was no significant main effect of testing mode on student anxiety. This result is expected in that there was no difference in the testing situations or examinee physiognomies of the school used for the study. The higher mathematics
anxiety recorded by the two testing modes conformed to the prevalence of mathematics anxiety in the Nigerian population (Awofala & Odogwu, 2017; Awofala & Akinoso, 2017; Awofala & Awolola, 2011). This study has shown that both the CBT and the PBT modes of assessment similarly impacted students’ achievement and anxiety in mathematics thereby removing the differential effect of testing mode on students’ learning outcomes. This outcome negated previous findings in which the CBT enhanced the motivation and performance of students over the PBT (Niko et al., 2016).

Conclusion

The present study has provided empirical proof that there is no differential effect of testing modes (CBT vs. PBT) on senior secondary school students’ learning outcomes in mathematics (mathematics anxiety and mathematics achievement). The implication of this is that both modes of testing are good for students’ assessment in mathematics as the two modes showed no variations in either students’ achievement in mathematics or mathematics anxiety. Thus, it is empirically sound to state that the CBT mode does not show better alternative to the PBT mode in mathematics assessment. However, more empirical investigations with larger samples of different cultural backgrounds regarding the effectiveness or otherwise of the CBT mode are needed in this age of information explosion as schools move toward paperless community. In conclusion, an adept comprehension of the effect of testing on students’ learning outcomes in mathematics could lead to the promotion of better-quality design of learning situations and procedures in mathematics with finally healthier schooling results for mathematics teaching and learning. It should be noted that online assessment such as the CBT remains a supple and effective way of carrying out credible and trustworthy examination with instant response devoid of human errors and idiosyncrasies. However, the credibility of the CBT mode of assessment may be in jeopardy if effort is not made to make available constant power supply, ensure constant network connection, and provide enabling environment at the CBT centres for proper security of questions to avoid leakages which may affect the test soundness and reliability. Government at levels in Nigeria need to put in place the enabling environment and infrastructural facilities needed for schools to benefit fully from the advantages inherent in the CBT mode assessment. In conclusion, the keyboards are not weightier than the biros in the assessment of mathematics achievement and mathematics anxiety of senior secondary school year three students in Nigeria.

Acknowledgments

The author would like to thank the senior secondary school students and teachers who willingly gave themselves for the research.
References


**About the Author**

**Adeneye O. A. Awofala.** The author is a Senior Lecturer in the Department of Science and Technology Education, Faculty of Education, University of Lagos. He holds a PhD in Mathematics Education, MSc in Mathematics specializing in Functional Analysis, MED in Mathematics Education and BEd in Teacher Education/Mathematics/Economics, all from the University of Ibadan, Nigeria. He teaches undergraduate and postgraduate courses in the Faculty of Education, University of Lagos.