

DIDACTIC APPRAISAL OF COMPUTER-BASED AND PAPER-PENCIL TESTING IN ECONOMICS AT THE 2014 UNIFIED TERTIARY MATRICULATION EXAMINATION IN NIGERIA

Isaac O. Ubi *Ph.D* & Augustine I. Robert

ABSTRACT

The study is an appraisal of Computer-Based Testing (CBT) and Paper-Pencil Testing (PPT) in economics at the 2014 Unified Tertiary Matriculation Examination (UTME). Purposive sampling technique was adopted in selecting a sample of 600 candidates from the list of 2014 examinees in Cross River State Centres who wrote economics. The selection of the sample was based on the two modes of the examination. Accordingly, 300 candidates were selected from each mode. Scores on economics for each candidate was extracted and used for the study. Two hypotheses (one on the level of performance and another on the influence of mode of administration and performance) formulated to guide the study. Results of the study reveal that performance of candidates who wrote the examination using computer was significantly better than those for candidates who wrote the same examination using paper and pencil. It is recommended that JAMB should create more centres for CBT method and move out of the PPT method to enhance candidates' performance.

Key words: Computer-based Testing; Paper-and-pencil Test; Unified Tertiary Matriculation Examination; Cross River State, Nigeria;

Introduction

Test is an important instrument for measuring candidates' performance in any educational institution. It is an instrument or activity used to generate data on a person's ability to perform a specified task. In Nigeria, the Unified Tertiary Matriculation Examination (UTME) is a common entrance examination administered by the Joint Admissions and Matriculation Board (JAMB), to assist Universities, Polytechnics and Colleges of Education, the country over in the selection of candidates for admission. The examination was for Universities alone and it was called University Matriculation Examination (UME) until 2010 when the Federal Government mandated JAMB to include Polytechnics and Colleges of Education in the selection examination (Ubi, 2016). Since that time there have been series of innovations, on the conduct of the

examination, one of which has been the transition from Paper-Pencil Testing (PPT) to Computer-Based Testing (CBT).

Paper-and-pencil testing is a method of assessing candidates' ability in which candidates read questions and respond in writing by indicating the answer or shading an option of choice on an answer script provided by the examination body. At the end of the test session the scripts are collated by the examining body, scored and recorded for the purpose for which the test was administered. On the other hand, Computer-based testing sometimes called e-examination, computerised testing or computer-administered testing is a method of administering a test such that the candidates use computers to assess the items, respond through the computer, and their responses are electronically recorded, assessed, or both. The use of computer system for testing in Nigeria date back more than 10 years, but JAMB began to apply the method only in 2013. In that year and in 2014 the UTME was conducted both using PPT and CBT. Candidates were free to use the option of their choice.

Literature Review

Research focus in the past has been on whether CBTs are equivalent to PPTs when similar test items were given. Gray (2013) carried out a study on computer-based testing and paper-pencil testing on third grade mathematics in the Northwest Missouri State University with a sample of 46 undergraduates. Of the 46 students in the sample, 26 took the paper-pencil test, while 20 took computer-based test. Results of the study showed that, even though candidates for the paper-pencil test performed slightly better than candidates for the computer-based test, the comparison between their mean performance was statistically insignificant. Perhaps the slight difference in performance might be as a result of item types. Some scanty literature support that items that are very verbose may confuse candidates using computer. Mazzeo and Harvey (2008) point out that computer-based test graphics may affect test scores and consequently their equivalence with paper-pencil versions and those tests with reading passages may be more difficult when given on computers.

The study by Ayo, Akinyemi, Adebisi and Ekong (2007) on Nigerian Universities revealed that more of the students found CBT easier than PPT. The study revealed that only 24.2% of the candidates had not been involved in any form of electronic examination before. Of this number, some found it challenging but adjusted with time. Evidence from literature reveal that some research findings point to the fact that most candidates in recent times performed better in CBT than they did in PPT.

The study by Dermo and Eyre (2008), George (2011) revealed that candidates' preference on PPT enhances their performance while CBT had a negative effect on their performance. Orlanshilay (2013) studied the impact of computer-based test on senior secondary school candidates in Agege Local Government Area of Lagos, Ngeria and

found that there was a significant difference between PPT and CBT candidates' performance with CBT candidates outperforming the PPT candidates. In an experimental study, this time on undergraduate students of educational psychology, Bodmann and Robinson (2004) observed that the students completed the CBT faster than the PPT with however no significant difference in the scores obtained. Similar studies by Segall, Tony, Doolen and Porter (2005) and Treadwell (2006) showed that handset-based quiz is more efficient than paper-pencil test in terms of speed. That is, candidates completed it in less time with handsets than they needed to complete using paper-pencil. The results however showed no performance effectiveness of candidates using the different modes.

Lukeman and Ogechi (2012) investigated the effect of test mode administration on mathematics performance measures among 120 candidates of the Universities of Nigeria and Ilorin where CBT was already in use. It was revealed that computer knowledge of candidates could influence their performance on CBT either positively or negatively. In consonance with this study, Calarina and Wallace (2002) tried to confirm several key factors in CBT and PPT assessments. Factors used for the study were content familiarity, and competitiveness. Post-test only design with test mode as the only factor was used. The dependent variable was measured on a 100 multiple choice test and on a self-report on a distance learning survey. Four selections of computer fundamental course consisting of 105 candidates were selected as sample. Results of the study showed that CBT delivery impacted more positively on candidates' scores as compared to PPT. the study also found out that the CBT group outperformed the PPT group. Gender, competitiveness, and computer familiarity were not related this performance difference. Evidence to support these findings can be found in the study of Russell and Haney (1997) which showed that candidates perform better on the computer test than paper-pencil test. The reason for the superiority of computer-based over the paper-pencil may be as argued by McDonald (2002) that in presenting a test on computer, a qualitatively different testing experience is created. That is, both examinees and the examiner may see the test situation as being more modern and may show more interest than they would do when the conventional PPT method is used.

Karadeniz (2009) carried out a study on the impact of three modes (paper-pencil based, web-based and mobile-based) of testing on candidates' achievement using a group of 38 candidates for the experiment which lasted for three weeks. Significant differences were found among candidates' scores in the second and third weeks of the experiment but not in the first week of the experiment. The findings of the study also revealed that candidates had positive attitude towards web-based and mobile-based assessment due to ease of use, comprehensive and instant feedback. More so, the study revealed that the most favoured tests were web-based and the least favoured were paper-pencil based. These findings are not surprising. Web-based tests are characterised by prompt release of results and error avoidance, while in most cases (Lukeman & Ogechi, 2012) paper-pencil tests are

manually scored and are error prone. Even when scoring is done using computer, human error still sets in as computers garbage out what is garbage in.

There seem to be a thin line among earlier researchers concerning the impact of CBT and PPT on candidates' performance. Some studies carried out on mode of administration and performance levels reveal varying results, some in support of difference in performance due to mode used and others in support of no significant difference. In addition to the literature reviewed here, the work of Do-Hong (2006) also revealed non-significant difference between CBT and PPT modes of testing on achievement. In another of such studies, Motunrayo (2012) compared students' performance in PPT and CBT of large-scale state wide end-of-course examinations in four subject areas (Mathematics, English, Biology and Physical Science). Results showed that, although some differences were observed, the overall results suggest that these differences were relatively small. For all subject areas, the model fit the data adequately for both PPT and CBT. Although there was some lack of invariance, fit indices showed that all three models of equivalence fit the data adequately and the additional restrictions did not adversely affect model fit for the entire four subject areas of Mathematics, English, Biology and Physical Science. No evidence, according to the study, was found to support the claim that the effects of the number of items tied to a common reading passage in the English test on candidates' performances varied across the administration modes. Studies like Eno (2011); Mazzeo and Harvey (2000) have all reported variations in performance due to mode of test administration.

Purpose and Hypotheses for the Study

The decision of the Joint Admission and Matriculation Board (JAMB), to shift from the traditional PPT delivery mode of administering the UTME was to overcome the problems of malpractices that were associated with the traditional mode of administering the examination as to ensure that there is independency among candidates so that their true abilities can be assessed. The CBT method was seen as one among methods the board could use to overcome the malpractice challenge. Some other challenges that led to the introduction of the CBT method were high cost of paper production, distribution of materials and logistic problems. Based on this factors JAMB in Nigeria decided to test run the CBT examination in 2013 and 2014. In these two years the examination was conducted using the two modes. These present researchers were poised at carrying out an appraisal of the two modes by comparing performance of candidates from the two modes. The hypotheses stated to guide the study were:

1. Performance of candidate in 2014 UTME economics examination using CBT and PPT mode of administration are not significantly different from the expected.
2. Candidates' performances in 2014 UTME economics examination are significantly dependent on the mode of administration of the examination.

Methodology

The study employed ex-post facto research design. The study area, Cross River State, is a coastal region in the south-south geo-political region of Nigeria. The state is named after the Cross River, which passes through about 70 per cent of the state. Calabar, the capital city of Cross River State, is now the leading tourism city of Nigeria. Visitors from different parts of the world come to the city in large numbers all the year round. The study population was all candidates who took the UTME economics examination through the CBT and PPT mode of administration in 2014 UTME in Cross River State. All of them were taught economics based on the same syllabus JAMB (2014) has given the population of this group of candidates as 3128. Of this number, 1228 took the examination through the CBT administration mode while 1900 took it using PPT mode.

Purposive sampling technique was adopted in selecting a sample of 600 candidates from the list of 2014 examinees who wrote Economics in Cross River State Centres. The selection of the sample was based on the two modes of the examination. Accordingly, 300 candidates were selected from each mode. Data for the study were collected directly from JAMB data base at the national headquarters in Abuja. The data so collected were coded and subjected to analysis using the version 18 of the Statistical Package for Social Sciences (SPSS).

Data Analysis and the Results

The data analysis for the study was done to test the two hypotheses at the .05 alpha level.

Hypothesis One:

Performance of candidate in 2014 UTME economics examination using CBT and PPT mode of administration are not significantly different from the expected.

This hypothesis was tested using population t-test statistical analysis for one sample mean. To carry out the test, the sample mean (X) was compared with the population mean (μ), in this study being the mean for economics performance for all candidates in Economics using CBT and PPT methods. The results are presented in Table 1.

Table 1

Summary of population t-test analysis for performance in CBT and PPT

Table 1

Summary of population t-test analysis for performance in CBT and PPT

Mode of Administration	n	X	μ	SD	t
CBT	300	57.40	53.25	11.53	6.232*
PPT	300	49.10	53.25	7.26	-10.990*

* $p < .05$; $df = 599$; critical $t = 1.96$

Results of analysis as shown in Table 1 show that the absolute values of the calculated t-values of 6.232 and -10.990 are greater than the critical t-value of 1.96 at .05 level of significance using 559 degrees of freedom. A close observation of the group means and the sign and size of the t-values shows that while candidates who took the examination through the CBT mode of administration scored significantly more than the expected, those that took it based on the PPT mode scored significantly less than the expected mean for the entire Cross River State Candidates who took economics in the UTME. Based on these results; the null hypothesis was rejected for candidates who took the examination using each of both administration mode.

Hypothesis two

Candidates' performances in 2014 UTME economics examination are significantly dependent on the mode of administration of the examination

The independent variable in this hypothesis is mode of the examination (classified into CBT and PPT), while the dependent variable is candidates performance. Based on this classification, the hypothesis was tested using independent t-test statistic. Results of the test are presented in Table 2.

Table 2

Summary of Independent t-Test Analysis of the Differences in the Performance of Candidates based on Different Modes of Administration.

Mode of Administration	n	X	SD	t
CBT	300	57.40	11.52	10.549*
PPT	300	49.10	7.26	

*p < .05; df = 598; critical t = 1.962

Results as presented in Table 2 show that the calculated t-value of 10.549 is greater than the critical t-value of 1.962 at .05 level of significance using 598 degrees of freedom. This means that candidates' performances in 2014 UTME economics do not significantly differ based on mode of examination. The results also indicate that candidates who wrote the entrance examination with the CBT mode significantly outperformed their counterparts who wrote the entrance with the PPT mode. Based on these results, the null hypothesis was rejected.

Discussion of Findings

Findings of the study on whether or not candidates' performance in Computer-Based Testing (CBT) and Paper-Pencil Testing (PPT) were significantly positive indicate that it was CBT candidates' performance that was significantly positive while PPT candidates' performance was significantly negative. These findings indicate that candidates who used the CBT mode to write the Economics entrance examination in Cross River State performed above average when compared with the entire population of candidates who wrote Economics in the state using either of the modes. Similarly, the findings indicate that candidates who used PPT mode to write the Economics entrance examination in Cross River State performed below average when compared with the entire population of candidates who wrote Economics in the state using either of the modes.

These findings point to the fact that scores which, emanated from candidates who used computer were relatively high, while those that emanated from candidates who used the traditional paper-pencil were relatively low. The findings, in a way are consistent with

those of earlier researchers like Karadeniz (2009), Lukeman and Ogechi (2012) who found that candidates had positive attitude towards web-based and mobile-based assessment due to ease of use, comprehensive and instant feedback. More so, the two independent studies revealed that the most favoured tests were web-based and the least favoured were paper-pencil based. There is a seeming contention among people that candidates' scores are better when they write multiple choice examinations with computer. The reason for this seeming overriding effect of computer mode on candidates' scores may attract future research interest.

The study findings showed that there was a significant difference in performance between candidates who used CBT mode for the examination and those who used PPT mode. The CBT mode was found to be better than the PPT mode in candidates' performance in the 2014 UTME Economics examination. This finding is consistent with those of Bodmann and Robinson (2004), Demo and Fyre (2008), George (2011), Orlanshilay (2013) who independently found that grades for students who used CBT mode for their examinations were better than those for students who used PPT mode for the same examinations. Bodmann and Robinson (2004) observed that students completed the CBT faster than the PPT with however no significant difference in the scores obtained while similar studies by Segall, Tony, Doolen and Porter (2005) and Treadwell (2006) showed that handset-based quiz is more efficient than paper-pencil test in terms of speed. That is, candidates completed it in less time with handsets than they needed to complete using paper-pencil. Calarina and Wallace (2002) have been reported to have tried to confirm several key factors in CBT and PPT assessments and found that CBT delivery impacted more positively on candidates' scores as compared to PPT with CBT group outperforming the PPT group.

The finding of this present study has, however, failed to replicate that of Gray (2013) who found that, even though candidates for the paper-pencil test performed slightly better than candidates for the computer-based test, the comparison between their mean performance was statistically insignificant. The finding of the present study has also failed to be in consonance with that of Do-Hong (2006), which has revealed non-significant difference between CBT and PPT modes of testing on achievement.

Conclusion and Recommendation

The Nigerian experience in entrance examinations for tertiary institutions dates back to the beginning of higher education in the country. Before now, as reported in Ubi, Joshua and Umoinyang (2012), candidates for admissions into tertiary institutions had freedom to write as many entrance examinations as they could financially cope. This caused a lot of problems as candidates secured so many admissions at a time. The new UTME which replaced UME has, according to critiques and even JAMB, has been plagued with so many problems, two of which are cost of administration and malpractices. For these

problems, JAMB test-ran the CBT as a parallel mode to PPT in the administration of the UTME. This present scholarship was poised at comparing the two modes in the bid to discover their efficacies. Results of the study have revealed that the CBT mode is more efficacious in performance at the UTME. It is therefore recommended that the Joint Admissions and Matriculation Board should fully adopt the CBT method and stop the traditional PPT method. This, the board can do by first providing all its centres with serviceable computers and then uninterrupted internet services. Future researchers should attempt to investigate some demographic variables in the examination like gender, location, age, school type and school ownership, to verify if such variables can interact with the mode of delivery to determine performance levels.

References

- Ayo, C. K., Akinyemi, I. O., Adebisi, A. A. & Ekong, U. O. (2007). The prospects of examination implementation in Nigeria. *Turkish Online Journal of Distance Education*, 8(4), 125-134.
- Bodmann, S. M., & Robinson, D. H. (2004). Speed and performance differences among computer based and paper-pencil tests. *Journal of Educational Computing Research*, 31(1), 51-60.
- Calarina, E., & Wallace, N. I. (2002). Key factors in computer-based testing versus paper-pencil testing assessment. *Journal of Research in Computer Education*, 28(3), 282-299.
- Gray, C. (2013). *Computer-based testing compared to paper-pencil testing in third grade Mathematics*. College of Education and Human Services, Maryville, MO 64468.
- Dermo, J., & Eyre, S. (2008). Secure, reliable and effective institution-wide e-assessment: Paving the ways for new technologies. In F. Khandia (Ed.), *Proceedings of International CAA Conference*(95-105). Loughborough: University of Loughborough.
- Do-Hong K. (2006). A comparison of student performance on paper- and-pencil and computer-based testing in four subject areas. Carolina: University of South Carolina.
- Eno, L. P. (2011). Comparing the reading performance of high –achievement adolescents: Computer-based testing versus paper-pencil testing. Seton Hall University Dissertations and Theses (EDTs). Paper 1764
- George, Y. (2011). Candidates' perception of computer-based assessment in the University of Ilorin, Nigeria. Retrieved from www.scribd.com/doc/71979150 3rd March, 2014.

- Karadeniz, S. (2009). The impacts of paper, web and mobile based assessment on candidates' achievement and perceptions. *Scientific Research and Essay*, 4(10), 984-991.
- Lukeman, S. A., & Ogechi, N. G. (2012). Test mode administration effect of Mathematics performance measures: A dissertation for the degree of Doctor of Philosophy, University of Nebraska, Lincoln.
- Mazzeo, J. & Harvey A. L. (2008). *The equivalence of scores from automated and conventional educational and psychological tests: College Entrance Examination Board*. New York: College Board Press.
- Mazzeo, J., & Harvey, A. L. (2000). *The equivalence of scores from automated and conventional education and psychological tests: A review of the literature*. Princeton NJ: Educational Testing Service.
- McDonald, A.S. (2002). The impact of individual differences on the equivalence of computer-based and paper-pencil educational assessment. *Computers and Education*, 39(4), 299-312
- Motubrayo, A. (2012). Planned adoption of computer-based test for UTME. Punch Nigeria Online Extra. Retrieved 12th January, 2015.
- Orlanshilay, F. J. (2013). The impact of computer-based test on senior secondary school candidates in Agege Local Government Area of Lagos, Nigeria. Retrieved from www.scribd.com 13th June, 2016.
- Rusell, M., & Haney, B. (1997). Testing writing on computers: An experiment comparing candidates' performance on test conducted via computer and via paper-pencil. *Education Policy Analysis Archives*, 5(3), 1-19.
- Segall, N., Toni, L., Doolen, J. & Porter, D. (2005). A usability comparison of PDA-based quizzes and paper-pencil quizzes. *Computers and Education*, 45(4), 417-432.
- Treadwell, k. (2006). The usability of personal digital assistants (PDAs) for assessment of practical performance. *Medical Education*, 40(9), 855-861.
- Ubi, I. O., Joshua, M. T. & Umoinyang, I. E. (2012). Assessment of dimensionality of Mathematics tests of University Matriculation Examination in Nigeria: Implications for regional development. *Journal of Educational Assessment in Africa*, 7, 122-130.