

Comparative Study of Computer Based Test and Paper Based Test in Assessing Students in General Subject

¹Hafsat Suleiman Jalo, ²Adamu Dahiru, ³Usman Ali Abdullahi

¹Dept. of Computer Engineering Technology, Gombe state polytechnic Bajoga, Gombe State Nigeria.

²Dept. of Computer Engineering Technology, federal polytechnic, Ilaro Ogun state, Nigeria

³Dept. of Computer science, federal college of education (T), Gombe State Nigeria.

hafsatsuleimanjalo@gmail.com

Abstract—This study attempt to compare computer based test and paper based test to gain insight on which of the assessment mode affects the continuity and performance while addressing the challenge of high number learners and scare resources. The study, further tested the effect of gender and modes background knowledge in determining students' performance. An examination assessment is one of the major methods of testing students understanding. It allowed an educational system to be able to grade learners for placement at different levels reward or otherwise. However, they have to be designed to minimize examinees' frustration and to limit the sources of examinee anxiety. This study adopts a distinctive survey using secondary data and a statistical tool was employed for testing and consist of 351 students that participated in both mode of assessment from the school of business FCE (T) Gombe. The p value > (0.05) indicates that the result obtained showed that the students' performance is better on Computer based test than on Paper based test. The study concluded with some recommendations of examination tools which would serve as a guide and enhance performance.

Keywords— CBT, PBT, Examination, performance, Assessment.

I. INTRODUCTION

Assessment is an essential tool in the educational process because it measures the students' understanding. In higher education institutions, the assessment of academic progress takes different formats such as essays, dissertations, examinations, assignments, projects, and presentations. Nevertheless, the most commonly used tool is an examination. As part of e-learning trend, the computer-based tests (CBT) became more prevalent than paper-based tests (PBT) in the domain of educational assessment as changes are made in assessment methodologies reflect practical changes in several methods. The CBT provides several advantages such as proposing a solution to mechanize the assessment process, reducing paper consumption which indirectly reduces greenhouse gases and energy, assisting students to evaluate their strengths and weaknesses and providing quantitative improvements in assessment for academics and tutors. These improvements can be noticed in reducing preparation time and cost, enhancing the test security, analysing the results easily, keeping record for item analysis and reliability of scoring. Moreover, the CBT offer enormous scope for innovations in testing and assessment and measure complex form of knowledge and reasoning which is not possible through traditional PBT methods. Since it requires changes not only in the assessment methodology but also in the curricula methodology in educational institutions,

substituting PBT with CBT is a major shift which raises serious questions regarding the rationale behind this rather radical change. More so, Assessment is an essential tool in the educational process because it measures the students' understanding [1].

In higher education institutions, the assessment of academic progress takes different formats such as essays, dissertations, examinations, assignments, projects, and presentations. Reference [2] had identified more than fifty different techniques used within higher education for assessment purposes; nevertheless, the most commonly used tool is an examination. As part of e-learning trend, the computer-based tests (CBT) became more prevalent than paper-based tests (PBT) in the domain of educational assessment as changes are made in assessment methodologies reflect practical changes in pedagogical methods [3]. Hence, the CBT became dominant in different types of examinations such as standardized tests (e.g., GRE, GMAT, and SAT), accounting professional certifications (e.g., CPA, CIA, and CMA) and college examinations [4]. Nevertheless, the comprehensive implementation of CBT mode has been hindered by questions such as them equivalence of CBT mode to PBT mode [5]. The CBT provides several advantages such as proposing a solution to mechanize the assessment process [6]; reducing paper consumption which indirectly reduces greenhouse gases and energy consumption [7]; assisting students to evaluate

their strengths and weaknesses [8]; and providing quantitative improvements in assessment for academics and tutors [9]. These improvements can be noticed in reducing preparation time and cost, enhancing the test security, analyzing the results easily, keeping record for item analysis and reliability of scoring [10], increasing efficiency [11], and providing instantaneous feedback to students [12]. Moreover, the CBT offer enormous scope for innovations in testing and assessment [13] and measure complex form of knowledge and reasoning which is not possible through traditional methods [14]. On the other end, CBT might have some limitations such as lacking underlining or making notations on computer screen, looking at the computer screen for a long time, and anxiety from changing the exam mode from PBT to CBT [15]. Nonetheless, prior research argued that students still prefer CBT more than PBT as performance assessment because CBT is more promising, credible, objective, fair, interesting, fun, fast and less difficult or stressful [16].

The management of Federal College of Education Gombe has recently implemented the use of CBT to test students' knowledge. The advantages of using computer technology for educational assessment in a global sense have been recognised and these include lower administrative cost, time saving and less demand upon teachers among others. Whilst recognising these system-level advantages it is important to explore the relationship between assessment mode and the behaviour as well as performance of the students being assessed because the assumption of comparability between CBT and PBT without proper investigation within that particular testing context, is inappropriate. Some test takers reported that, it is more difficult to navigate back to rework problems.

The challenge on CBT test designers and administrators is to construct CBT to be fair and reliable and to produce valid test scores. Furthermore, they have to be designed to minimize examinees' frustration and to limit the sources of examinee anxiety. These additional test Design steps are well worth taking, Because of the effective and measurement improvements they offer. CBT implementation should also be constructed to meet the standard requirements such of International Test Commission(ITC) of Technology, Quality, Control, and Security.

It is therefore anticipated that providing specific information about students' performance in CBT in undergraduate courses may enable refinement of the current use of CBT. This is because subjects' attitudes and feelings towards a test is an important factor in test design, since, if it is ignored there may well be adverse effect on test scores. In summary, CBT efficiency in any given context is indicated by the following factors: acceptability by the

students, validity and reliability, compliance with standard guidelines, and comparability with PBT i.e. minimizing shift in test mode effect.

In this paper, we intend to compare computer based test (CBT) and paper based test (PBT) in a course in school of business from the federal college of education (T) Gombe, Nigeria.

II. METHODOLOGY

In this paper the methodology and concept used to carry out this study is discussed in details. This study adopts the quantitative analysis method of research, which comprises and focuses on the research design, location of the study, the number of people required, sampling technique and sample size, research instrument needed to carry out a successful evaluation, validity of the instrument and method used for data collection and analysis.

A. Research Design

The main reason of this study was to investigate the effect of paper-based and computer-based test of assessment and then to investigate how gender, continuity, and the ability to score the various methods automatically, Performance of each on PBT and CBT was within-a desirable variable range. All the participants completed all paper-based and computer-based measures. For this study design, a CBT and PBT exam was conducted and each of the students was scored according to performance, gender, continuity, and the ability to operate the computer.

B. Population

This study or research was done during 1st semester in the faculty of business in FCE(T) Gombe and came to a conclusion that the study or research carried out involved students who wrote both general studies education examination in PBT and CBT and their result was later analysed. The measurement equipment or instruments used to carry out the result analysis with a group of about 351 students. At conclusion, we had two finalized groups which is CBT group and the PBT group with 2924 students allocated (see Table 1).

TABLE I
POPULATION BASE ON DEPARTMENT

Student Department	No of Students
school of business	2924
total	2924

C. Sampling Techniques and Sample Size

A sample is a portion of the population that is used for study. It comprises some members selected from the population. Given the large population of the study, a census cannot be carried out because of time and resource constraint of the study. Therefore, a sample will be taken as a representative

of the population. A sample size, which refers to the number of items to be selected from the population of the study, will be determined using Yamane (1967) formula. The formula is given as:

$$n = \frac{N}{1 + N(e^2)}$$

Where: n = sample size

N = population of the study

e = level of significance (taken at 5% or 0.05)

Using this formula, the sample size is determined as follows:

$$n = \frac{2924}{1 + 2924(0.05^2)}$$

$$n = \frac{2924}{1 + 7.31}$$

$$n = \frac{2924}{8.31}$$

$$n = 351$$

D. Method of Data Collection

In this research/study an exam was conducted in both CBT and PBT and analysed using SPSS statically analysing tool, the secondary data was obtained from the school of business after which the exams in both CBT and PBT have been conducted. The result of each examination is collected into an excel document from the students randomly sampled.

Furthermore, the analysis on the student based on gender and continuity was also collected by accessing each individual sampled student files to identify if there has been a computer learning background at their O' levels. Which was compiled into an excel document as well.

E. Method of Data Analysis

After identifying the sample an observation of student Paper based Test and Computer based Test was carried out and individual result was extracted for PBT and CBT respectively and tabulated according to performance, gender and computer knowledge. A comparison of the result to compare PBTA and CBTA for general performance, CBTM and CBTF to test gender effect and CBTY and CBTN to test those with ICT knowledge background. Looking at the sampled items T- Test is identified to be best for the comparison because sampled items are up to 351 items and because it can compare two set of samples. Therefore, independent sample t- Test was use as a tool to compare the sampled items. If the p-value is greater than 0.05 the targeted variable is significantly more than the other.

F. Research Instrument

A quantitative research method was practically used in this research. Data was collected after the exams were being conducted and sample according to the sampling size. Each data collected was analysed based on the testing type, gender, and knowledge background respectively.

III. ANALYSIS AND PRESENTATION OF RESULTS

This section presents the data collected from PBT and CBT students results as the analysis and interpretation thereof. The result collected during the student exams were coded into SPSS and analysed using independent sample t- test to obtain the mean, t- value and p value of the analysis respectively. The data collected for this study did not contain high levels of missing data as such no cases were dropped. Record of 351 students who sat for the PBTA (i.e. PBT All and CBTA (i.e. CBT All) were tested.

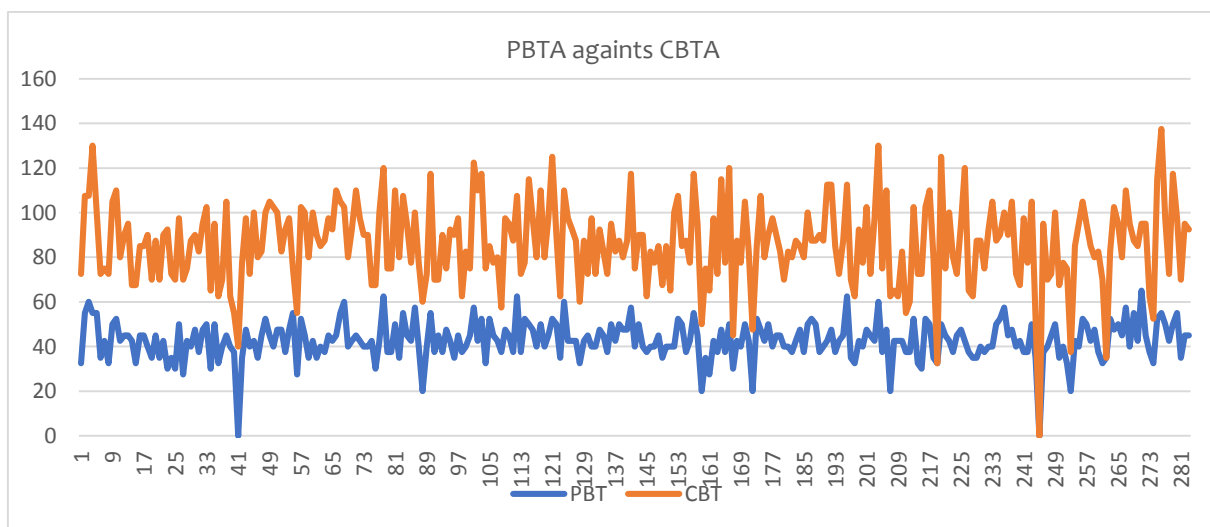


Fig. 1 Graph of PBTA against CBTA

Figure 1 shows the performance of all student for both PBTA and CBTA (PBT of all the student and CBT of all the student). The students' performance in the CBT is generally better than that of their PBT, except for some few students. for example, students with serial numbers 41, 217, 249 score far better with PBT than when using CBT. The possible reason for the better performance using the CBT was due to the

fact that majority of them find it easy to recall and identify correct answers to questions than if ask to supply the answers off heart. The number of student result collected to compare the performance of both the male and female who sat for the CBT Test are used for further analysis in this research.

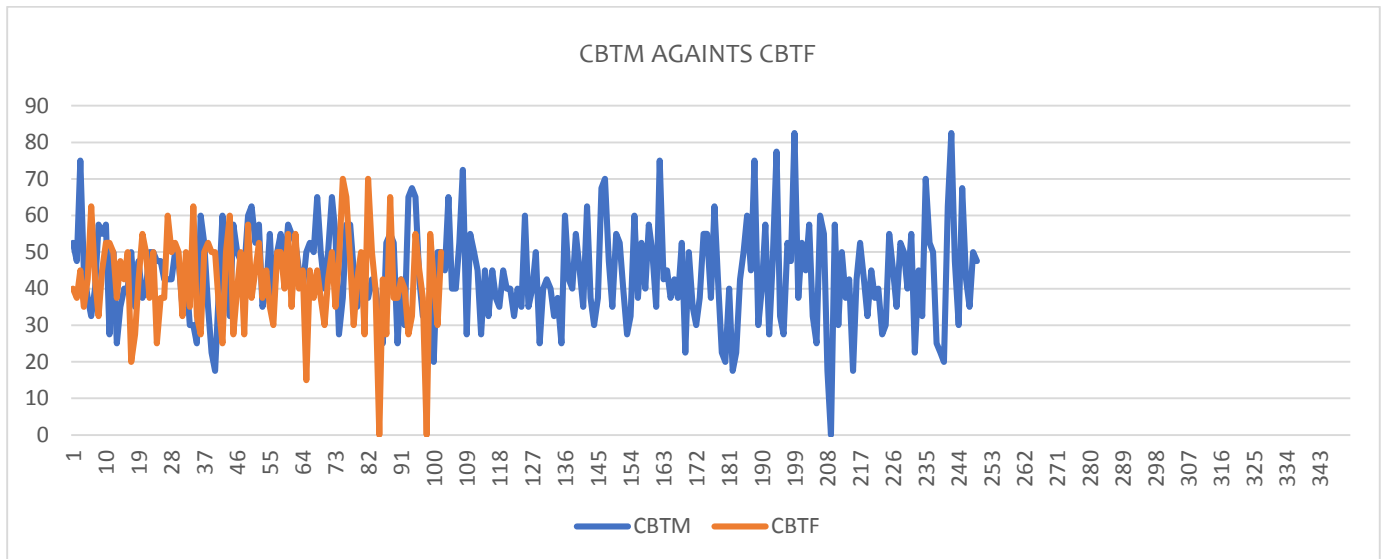


Fig. 2 Graph of CBTM against CBTF

Figure 2 shows the performance of all student for CBTM and CBTF (CBT the male students and CBT the female student). The students' performance for CBTM and CBTF is equal, except for some few students. For example, CBTF students with serial numbers 82 and 100, scored less while CBTM student with serial number 208 scored less also. Despite that the male students outnumbered the female students

The possible reason for the better performance using the CBT was due to the fact that both gender is competitive and both ready to explore new thing. The data collected among the 351 sampled students who had little or no computer educational background at O' Levels. Figure 3 shows the performance of all student for both CBTN and CBTY (CBT of those having no computer background and those with computer background at o' level).

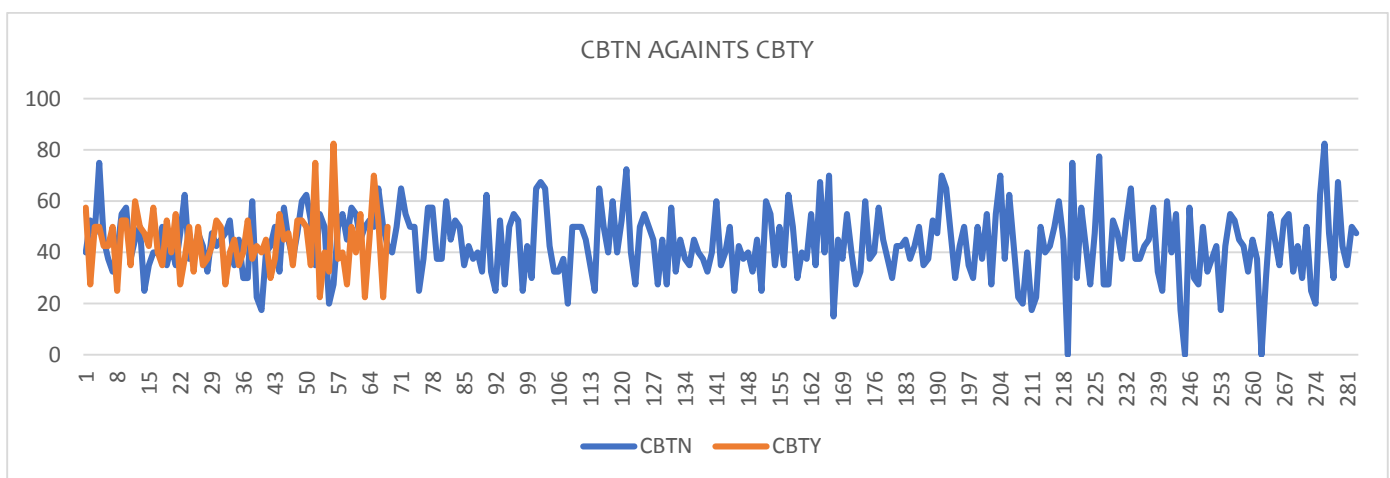


Fig. 3 Graph of CBTN against CBTY

The students' performance in the CBTN and CBTY is basically equal, except for some few students. For example, students with serial numbers 218, 246, 264 from CBTN that scored less. The possible reason for the equal performance was due to the fact there are a lot of computer devices in the hands of individual which enhance familiarity even without computer background knowledge.

A. Hypothesis Test

Hypothesis Analysis -1

The two independent group sampling was used to compare CBTA (CBT of all the students sampled) and PBTA (PBT of all the students sampled) with total population of 351 and obtained a MEAN of (42.9, 43.5) and STD DEVIATION of (12.9, 8.19) for CBTA and PBTA respectively. This implies that the performance during PBT exams was higher than that of the CBT. This may be due to little or less familiarity and knowledge of operations of the computer. Furthermore, the analysis seeks to compare the performance of students who wrote CBT and PBT exams with participant size of 351 students that wrote the PBT exams: 351 students that wrote the CBT exams. There is significant difference for CBT ($M=43.5$, $SD=12.9$) and PBT ($M=42.9$, $SD=8.19$) exams respectively; with $t(700) = 0.770$ and $p(0.441)$. These results suggest that the performance during PBT exams was higher than that of the CBT. This may be due to little or less familiarity and knowledge of operations of the computer.

Hypothesis analysis-2

The two independent group sampling was used to compare CBTM (CBT of all the male students) and CBTF (CBT of all the female students) with total population of 197:86 and obtained a MEAN of (43.9, 42.0) and STD DEVIATION of (13.3, 12.9) for CBTM and CBTF respectively. This implies that the male gender performed better than the female gender. An independent sample t-test was run on the second part of the hypothesis analysis to compare the gender of students who wrote the CBT exams, having participants size of 197 males and 86 females. There was a significant difference for CBT ($M=43.9$, $SD=13.3$) of male gender and CBT ($M=42.0$, $SD=12.9$) of female gender exams respectively; with $t(281) = 1.16$ and $p(0.25)$. These results suggest that the performance of the male gender during the exams has been more encouraging than that of the female gender. This implies that probably the male gender has been more focused and stable and might have more knowledge of using computer than the female gender during the CBT exams.

Hypothesis Analysis -3

Two independent group sampling was used to compare CBTY (CBT of students who had background of computer knowledge in O levels) and CBTN (CBT of students

who had no background of computer knowledge in O levels) with total population of 69:282 and obtained a MEAN of (43.9, 43.4) and STD DEVIATION of (11.6, 13.2) for CBTY and PBTN respectively. This implies that results of the performance of those with computer knowledge or background has out passed those without the computer background at all.

An independent sample t-test was run on the third part of the hypothesis analysis to compare those students who have computer background at O level with participant size of 69 students who had been thought computer basics at the O level :282 students who have not been taught at all. There is statistical significant difference for CBT ($M=43.9$, $SD=13.6$) of those with computer background at O level and CBT ($M=43.4$, $SD=13.2$) of those student without computer background at O level respectively; with $t(349) = 0.29$ and $p(0.77)$. These results suggest that the performance of those with computer knowledge or background and those without showed quite difference, which implies that student with the background knowledge had more advantage.

IV. DISCUSSION

Firstly, Comparing the CBTA to PBTA, the result of the analysis indicates that there is significant difference in the performance of students with questions presented on paper versus questions on the computer considering the p-value ($p > 0.05$). The mean value (PBT > CBT) i.e. (42.9, 43.5), STD DEVIATION (12.9, 8.19), $t(700) = 0.770$ and $p(0.441)$ indicates that student's performance is higher when questions are presented on paper compared to questions presented on computer. Secondly, using gender comparison, the result of the analysis with mean value (43.9, 42.0), the STD DEVIATION of (11.6, 13.2) with $t(281) = 1.16$ and $p(0.25)$, reveals that students use of CBT during learning interaction, compatibility, continue usage intention, ease of use, user satisfaction and usefulness is independent of gender since all the p-values are greater than 0.05. Lastly, the primary data collected on to analyse those students with computer learning background during O level with MEAN of (43.9, 43.4), STD DEVIATION of (11.6, 13.2), $t(349) = 0.29$ and $p(0.77)$, reveals that there was significant impact of the use of CBT on the students to the use of PBT as a means of assessment. In order to enhance the performance of the student, the analysis of the following exam testing applications was done, where the following result was achieved and were recommended them for practices which best guide the student before the main CBT exams is taken by the students. This exam testing application include:

1. Tecnickcom Exam: is an open source exam software which do not require any additional hardware to run. This application contains a large number of modules,

and these modules contain an even larger number of topics and these topics contain an immense amount of questions. This software provides equal opportunities to people with disabilities. It has an option of generating printable OMR sheets which can be filled and scanned later. It supports translation in several languages, even right to left languages like Arabic.

2. Internet Exam: Internet Exam is a non-free web-based examination software that helps you conduct online assessments, examinations, recruitments, tests, and surveys. It enables you to create varied question types and automatically grades the question papers. It provides a simple UI to users, that helps them design question papers without any hassle and can simple work on both iPhone and androids.

IEExam lets examiners create question banks; examiners can continually update and make changes to the question bank. During tests, these question banks can be used to generate question papers easily. User can also customize the question paper format according to their needs, and can assign time, score, and date of their choice. Then generate reports and analysis accordingly. Some of the most advantageous features of iExam software are labelling, student management, grades & examination management. This web-based examination software works as one of the most efficient tools to assess the performance of any student or job candidate.

3. VirtualX: VirtualX is a fully automated, open source exam software which can be used in colleges, companies and e-learning organizations. It supports 12 different types of questions such as multiple-choice questions, fill in the blanks, yes or no, drag and drop and so on. VirtualX online examination system php open source also generates five different types of reports, both graphical and textual, which is beneficial for the examines for further analysis. An option of manual evaluation is also provided for descriptive questions in VirtualX.
4. Eklavya: is a non- free but known online examination software as it provides complete automation and administration of your Examination/ Assessment

Process which provide question bank management, less paperwork, exam result processing, payment management, trusted by many, scalable, no upfront investment, knowledge management.

5. Moodle: is a best open source software for Learning Management System (LMS) coded in PHP and distributed under GNU General Public License. This online examination system php open source has a strong and integrated system which is suitable for educators, administrators and learners. Moodle is one of the most trusted, secure and easy to use software available in the market. This software can also be used to create customizable websites with online courses for educators and trainers to achieve learning goals. Moodle online assessment software can be scaled depending on the number of people using it. So, its suitable for managing a bunch of few students in a coaching institute or hundreds of employees in a large organization.
6. Conduct Exam computer based offline test: is one of the best non- free campus management software that allows educational institutes to conduct exams offline. One of the greatest benefits of Conduct exam computer-based test software is that it does not require the user to have internet connectivity to take an exam. Students can easily download it as many times as they want to take the test from the specific website, can upload a test and later check their result. Computer-based assessment software supports multiple languages so that students can take the test in any language they want.
7. Flexiquiz: is another powerful online test and quiz maker, which helps create tests and assessments with ease. This is a free online exam maker, which also allows you to analyze tests and the performance of students. Tests can be published for selective groups or for everyone, depending upon the test format. This software can serve as the best support to conduct online exam. Having reviewed the above seven (7) examination practice tools, some recommendation on the few the exams tools are presented in table 2.

TABLE XI
LIST OF EXAM APPLICATION

Name of Application	Criteria for Measuring the Examination Tool			
	Ease of use	Availability	Configuration	Deployment
tecnickcom exams	it is simplified	free	can be manipulated according to requirement.	web based
virtualx	it is easy to use	free	can be customize according to required need.	web based
moodle	it is flexible integrated system	free	it offers all-in progress of examinee.	web based
flexiquiz	easy to use	free	it allows auto grading.	web based

IV. CONCLUSIONS

This study had examined the impact of Computer based test assessments in comparison to a paper based assessment in FCE(T) Gombe and overall conclusion is that the Students performed encouragingly well during the Computer based Test assessment compared to a paper based in terms of learning interaction, compatibility, ease of use, user satisfaction and usefulness. However, the continuous learning and administration of the exam practicing tools such as Tecnickcom Exams, Moodle, Internet Exam, VirtualX, Flexiquiz, Eklavya will have effected on their result performance.

REFERENCES

- [1] J. Brinke, V. Bruggen, J. Hermans, H. Burgers, J. Giesbers, B. Koper, R. (2007). Modeling Assessment for Re-use of Traditional and New Types of Assessment. *Computers in Human Behavior*, 23(6), 2721-2741. <https://doi.org/10.1016/j.chb.2006.08.009>.
- [2] G. Sim, P. Holifield, & M. Brown,. (2004). Implementation of Computer Assisted Assessment: Lessons from the Literature. *Association for Learning Technology Journal*, 12(3), 215-229.
- [3] OECD. (2010). PISA Computer-based assessment of student skills in science. Retrieved Jun. 15, 2017, from <http://www.oecd.org/publishing/corrigenda>.
- [4] R. Bull, 1999; G. Conole & D. Warburton, 2005) A review of computer-assisted assessment ALT-J, *Research in Learning Technology* Vol. 13, No. 1, March 2005, pp. 17–31.
- [5] D. Akdemir & A. Oguz, 2008; Effects of computer-based educational achievement test on test performance and test takers' motivation September 2013 *Computers in Human Behavior* 29(5):1889–1895 DOI:10.1016/j.chb.2013.03.008.
- [6] E. Triantafyllou, E. Georgiadou, & A. Economides,. (2008). CAT-MD: Computerized adaptive testing on mobile devices. *International Journal of Web-Based Learning and Teaching Technologies (extended versions of the best papers presented at m-ICTE2006 Conference)*, Vol. 3, No. 1, pp. 13-20. IGI Global. ISSN: 1548-1093, E-ISSN: 1548-1107.
- [7] M. DeRosa, 2007. The medium matters: Mining the long-promised merit of group interaction in creative idea generation tasks in a meta-analysis of the electronic group brainstorming literature *May 2007 Computers in Human Behavior* 23(3):1549-1581.
- [8] A. Kaklauskas et al., 2010 Multiple criteria assessment of alternatives for built and human environment renovation. Pages 257-266 | Received 07 Oct 2009, Accepted 22 Jan 2010.
- [9] C. Singleton, (1997). Computer-based assessment of reading. In Beech, J.R. & Singleton, C.H. (Eds.), *Psychological assessment of reading*. London: Routledge.
- [10] Gvozdenko & Chambers, 2007 Applying Computerised Testing & Certainty Based Assessment to Reveal More about Student Learning: From Theory to Practice. *Source: International Journal of Learning*. 2007, Vol. 13 Issue 12, p205-216. 12p. 2 Color Photographs, 1 Chart, 13 Graphs. Sanni, A. A., & Mohammad, M. F. (2015). Computer based testing (CBT): An assessment of student perception of JAMB UTME in Nigeria. *Computing, Information Systems, Development Informatics & Allied Research Journal*, 6(2), 13-28.
- [11] C. Halcomb, D. Chatfield, B. Stewart, M. Stokes, Cruse, B. & Weimer, J. (1989). A computer-based instructional management system for general psychology. *Teaching of Psychology*, 16(3), 148-151.
- [12] A. Bugbee Jr, (1996). The equivalence of paper-and-pencil and computer-based testing. *Journal of Research on Computing in Education*, 28(3), 282-299 Bull, J. & McKenna, C. (2003). *A Blueprint for Computer-Assisted Assessment*. New York: Routledge.
- [13] Bennett, R. (1998). *Reinventing assessment: speculations on the future of large-scale educational testing*. Princeton, NJ: Educational Testing Service, Policy Information Center.
- [14] S. Bodmann. & D. Robinson. (2004). Speed and performance differences among computer-based and paper-pencil tests. *Journal of Educational Computing Research*, 31(1), 51-60.
- [15] D. Butler. (2003). *The Technology Source archives at the University of North Carolina*, January /February. Retrieved on June 26, 2017 from http://technologysource.org/article/impact_of_computer_based_testing_on_student_attitudes_and_behavior.
- [16] K. Sambell, A Sambell, & G. Sexton. (1999). Student perceptions of the learning benefits of computer-assisted assessment: a case study in electronic engineering. In Brown, S., Race. P. & Bull. J. (Eds.), *Computer assisted assessment in higher education*. London: Kogan Page.