Comparative Analysis of Senior Secondary School Students' Academic Performance in General Mathematics using Different Testing Modes

Jude Kentom ADENIJI & Dr W. J. UBULOM

1Department of Educational Foundations
Faculty of Technical and Science Education
Rivers State University of Science and Technology
Port Harcourt, Nigeria
E-mail: jkentom@yahoo.com

2Department of Business Education
Faculty of Technical and Science Education
Rivers State University of Science and Technology
Port Harcourt, Nigeria
E-mail: will.ubulom@yahoo.com

ABSTRACT
The study was poised to comparatively analyse the mean performance scores of Senior Secondary Three students in General Mathematics using different modes of testing expected to consequently influence the decision of assessment/evaluation experts. The study consists of the 72 SS3 students of Deeper Life High School, Port Harcourt. Two research questions and hypotheses were generated to guide the study. A fifty-item multiple choice achievement test with options lettered A-D in which only one is the correct answer while others are foils. Arithmetic mean and standard deviation was used for data analysis. The t-test statistics was used to test the two hypotheses generated at 95% confidence level. The hypotheses were accepted as the critical values exceed the calculated values in both cases. It was consequently discovered that a change in mode of testing does not alter the mean performance of SSCE candidates.

Keywords: SSCE candidates, e-learning, e-teaching, Computer-Based Testing

INTRODUCTION
With the growing spate of advancement in technology especially in the world of computers, the need to employ e-education cannot be over emphasized. The introduction of e-learning and e-teaching is fast becoming pronounced in Africa, especially in Nigeria and South Africa. A number of research work has shown the positive impact of the use of electronic gadgets, computers and internet in aiding the teaching-learning process. In fact, the 21st century child is more interested and motivated in learning when it comes to e-education. Children nowadays easily fondle with laptops, ipads, tablets, iphones and android without reading the manual or being tutored. It is also a known fact that the larger percentage of computer/internet users globally are children and youths under 30years. These tyros are fascinated with the myriads of interacting platforms provided by the various social networking sites such as Facebook, Yahoo messenger, Twitter, 2go, Watsapp, Skype, Google+, Twoo, Instagram and many more. This shows that computer-related activities tend to sustain the attention of the 21st century learners more than ever before.
It is of note that the amount of learning and the consequent change in behaviour is largely ascertained through assessment. Hence, the sacrosanctity of testing in education cannot be over emphasized. Every teaching-learning process should be climaxed with at least an achievement test. Although, assessment has been done verbally and through the use of stationeries (all sort of writing materials) from the time immemorial. Specifically, the use of paper, pencil and biro was most popular then. But there has been a paradigm shift from Paper-and-Pencil Testing (PBT) to Computer-Based Testing (CBT) (Zinedine, 2005). In developed countries of the world, computer-based testing has been widely accepted as a system of assessment. In recent times, third world nations have equally seen the need to toll the path of technology in different sectors, education inclusive. Having adopted electronic teaching, it is necessary to employ electronic testing in order to measure the amount of learning outcome that has taken place in the course of instruction. While many African countries seem to be financially incapable of affording the exorbitant cost of electronic education, a good number of institutions and examination bodies in the continent have ventured into e-learning and e-testing. Even though e-education is relatively new and capital intensive, it is generally seen as a 'necessary evil' considering the spate of technological advancement around the globe today.

Computer-Based Testing (CBT) has become widespread in recent years. Some states in U.S. now use an online platform as the primary delivery mode for one or more computer-based tests for accountability purposes. When CBT was emerging in state testing in the early 2000s, Thompson, Thurlow, Quenemoen & Lehr (2002) examined the implications of CBT for students in general. There was not much literature about the use of CBT for large-scale assessments at that time. Thompson et al worked with states to explore what needed to be considered during development for students and how states might address the needs of these students for accommodations in a CBT environment. Since the early 2000s, much has occurred in CBT. CBT seems to have advantages over paper and pencil testing, both for states that run the assessment programs and for the students who participate in them. These advantages are recognized by the U.S. Department of Education, which in one of its major initiatives (Race to the Top Assessment Program), encouraged the development of CBT. There currently is strong interest in CBT and advocates have identified many positive merits of this approach to assessment including: efficient administration, student preference, self-selection options for students, improved writing performance, built-in accommodations, immediate results, efficient item development, increased authenticity, and the potential to shift focus from assessment to instruction (Becker, 2006; Salend, 2009; Thompson et al., 2002). CBT also allows new ways of assessing students that move beyond the traditional multiple choice and constructed response items. For example, innovative assessments are now being developed that enable students to manipulate data and role play. Yet, as states move forward with CBT they are discovering that it is important to consider not only the positive benefits, but also potential negative unintended consequences. These include, for example, the possibility that additional training will be needed for students with disabilities to interact successfully with computers and the challenges of determining the best way to present some accommodations such as screen readers.

Classical test theory (CTT) and Item Response Theory (IRT) are largely concerned with the same problems but are different bodies of theory and entail different methods. Although the two paradigms are generally consistent and complementary, there are a number of points of difference:

- IRT makes stronger assumptions than CTT and in many cases provides correspondingly stronger findings; primarily, characterizations of error. Of course, these results only hold when the assumptions of the IRT models are actually met.
- Although CTT results have allowed important practical results, the model-based nature of IRT affords many advantages over analogous CTT findings.
- CTT test scoring procedures have the advantage of being simple to compute (and to explain) whereas IRT scoring generally requires relatively complex estimation procedures.
- IRT provides several improvements in scaling items and people. The specifics depend upon the IRT model, but most models scale the difficulty of items and the ability of people on the same metric. Thus the difficulty of an item and the ability of a person can be meaningfully compared.
Another improvement provided by IRT is that the parameters of IRT models are generally not sample- or test-dependent whereas true-score is defined in CTT in the context of a specific test. Thus IRT provides significantly greater flexibility in situations where different samples or test forms are used. These IRT findings are foundational for computerized adaptive testing.

**Research Questions**

The following questions guided the study.

1. How does e-testing on the achievement of Computer-Based Tested Mathematics candidates in comparison with the achievement of their Paper-Pencil Tested counterparts?
2. How does e-testing impact on the achievement of Computer-Based Tested Mathematics candidates in comparison with the achievement of their Dual-Based Tested counterparts?

**Hypotheses**

The following null hypotheses were postulated and tested in the study.

1. There is no significant difference between the achievement of Computer-Based Tested Mathematics candidates and the achievement of Paper-Pencil Tested Mathematics candidates.
2. There is no significant difference between the achievement of Computer-Based Tested Mathematics candidates and the achievement of Dual-Based Tested Mathematics candidates.

**METHOD**

The research design adopted for the study was the descriptive research design. The design examines the difference between the performances of Paper-Pencil Tested (PPT), Dual-Based Tested and Computer-Based Tested (CBT) SS3 Mathematics students in Deeper Life High School, Port Harcourt. The population for this study comprises the 72 SS3 students of Deeper Life High School who participated in the six weeks 2015 Summer School. Owing to the small size of the population, the researchers used the whole population as the sample for the study.

The instrument for the study was compiled using standardized questions adapted from the West African Examination Council General Mathematics question papers between 2011 and 2014. The instrument contains 50 multiple choice objective questions with 1 mark awarded to each item to make a total of 50 marks. An item contains a stem and four different options lettered A, B, C, and D in which only one is the correct answer while the others are termed by evaluating experts as decoys (distracters). The reliability of the test was ascertained using the Pearson’s Product Moment Correlation Coefficient and Spearman Brown's Correction statistics to give \( \gamma \approx 0.9 \). It has two versions which were the printed/hardcopy and the softcopy for PPT and CBT respectively (Appendix B). The CBT version was programmed on the candidates’ personal computers for the exercise. Mean and standard deviation were the descriptive statistics used in answering the research questions while t-test statistics was used to test the hypotheses at 0.05 alpha level.

**RESULTS**

**Research Question 1**

*How does e-testing on the achievement of Computer-Based Tested Mathematics candidates in comparison with the achievement of their Paper-Pencil Tested counterparts?*

**Table 1: Mean and Standard Deviation Computation of the achievement of Computer-Based Tested Mathematics candidates in comparison with the achievement of their Paper-Pencil Tested counterparts \((N = 72)\)**

<table>
<thead>
<tr>
<th>Modes</th>
<th>No of candidates</th>
<th>Total Score</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPT</td>
<td>24</td>
<td>899</td>
<td>37.5</td>
<td>5.8</td>
</tr>
<tr>
<td>DBT</td>
<td>24</td>
<td>921</td>
<td>38.4</td>
<td>6.1</td>
</tr>
<tr>
<td>CBT</td>
<td>24</td>
<td>919</td>
<td>38.3</td>
<td>7.8</td>
</tr>
</tbody>
</table>

Table 1 shows the analysis of the data using simple mean and standard deviation with the inclusion of the total score for each group; namely, PPT, DBT and CBT. The table indicates the mean and standard deviation scores for the 72 candidates that sat for the examination using paper-pencil test (PPT), dual based test (DBT) and computer based test (CBT) on 50-Mark maximum scores were 37.5 (5.8), 38.4 (6.1) and 38.3 (7.8) respectively. This is an indication that the candidates performed excellently well in their paper-pencil test (PPT), dual based test (DBT) and computer based test (CBT).

The difference of 0.8 between the average scores of PPT candidates and CBT candidates is so negligible to conclude that computerised testing can foster a better performance among Mathematics students.

**Research Question 2**

How does e-testing impact on the achievement of Computer-Based Tested Mathematics candidates in comparison with the achievement of their Dual-Based Tested counterparts?

Table 1 also shows the e-testing impact on the achievement of Computer-Based Tested Mathematics candidates in comparison with the achievement of their Dual-Based Tested counterparts.

The difference of 0.1 between the average scores of DBT and CBT candidates is negligibly infinitesimal.

**Hypothesis 1**

There is no significant difference between the achievement of Computer-Based Tested Mathematics candidates and the achievement of Paper-Pencil Tested Mathematics candidates.

**Table 2:** t-test Computation of the difference between the achievement of Computer-Based Tested Mathematics candidates and the achievement of Paper-Pencil Tested Mathematics candidates

<table>
<thead>
<tr>
<th>Type of Test</th>
<th>N</th>
<th>SD</th>
<th>Mean</th>
<th>Df</th>
<th>t-cal.</th>
<th>t-crit.</th>
<th>Level of Sig.</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPT</td>
<td>24</td>
<td>5.8</td>
<td>37.5</td>
<td></td>
<td>0.4032</td>
<td>1.684</td>
<td>0.05</td>
<td>Accept</td>
</tr>
<tr>
<td>CBT</td>
<td>24</td>
<td>7.8</td>
<td>38.3</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the table above, the t-test statistics of significance was used for analysis at 0.05 level of significance. The result shows that the null hypothesis of no significant difference stands accepted since the calculated value of 0.4032 is less than the critical value of 1.6840. This implies that students' achievement remains the same when tested using either Paper-Pencil Test or Computer-Based Test.

**Hypothesis 2**

There is no significant difference between the achievement of Computer-Based Tested Mathematics candidates and the achievement of Dual-Based Tested Mathematics candidates.

**Table 3:** t-test Computation of the difference between the achievement of Computer-Based Tested Mathematics candidates and the achievement of Dual-Based Tested Mathematics candidates

<table>
<thead>
<tr>
<th>Type of Test</th>
<th>N</th>
<th>SD</th>
<th>Mean</th>
<th>Df</th>
<th>t-cal.</th>
<th>t-crit.</th>
<th>Level of Sig.</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBT</td>
<td>24</td>
<td>6.1</td>
<td>38.4</td>
<td>46</td>
<td>0.0495</td>
<td>1.684</td>
<td>0.05</td>
<td>Accept</td>
</tr>
<tr>
<td>CBT</td>
<td>24</td>
<td>7.8</td>
<td>38.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the table above, the t-test statistics of significance was used for analysis at 0.05 level of significance. The result shows that the null hypothesis of no significant difference stands accepted since the calculated value of 0.0495 is less than the critical value of 1.6840. This implies that students' achievement remains the same when tested using either Dual-Based Test or Computer-Based Test.

**CONCLUSION**

It is crystal clear that the achievement of Mathematics candidates is not influenced by the medium of testing. This reveals that students do not score higher or less even when the testing is computerised. Empirical deduction made from this study is not just interesting but informative about the need to test the 21st century child using his/her medium of interest to get the best. The sample used for the study excelled almost equally when tested using the computer system (both DBT and CBT). As noted earlier, e-learning cannot be said to be holistic in the black continent without the assessment being done electronically.
RECOMMENDATIONS
The following recommendations were made:
1. Government should provide an enabling environment for Computerised testing in government schools
2. Mission and private schools should intensify their drive towards e-learning and assess students electronically.
3. Public Examination bodies (JAMB, WAEC, NECO and NABTEB) and assessment/evaluation experts in the country are implored to key into the best practice of e-testing.

REFERENCES
McGroroy, K., & Sampson, H. (2010, June 7). Glitches Delay Florida Test Score Results. Miami Herald. Published online.