IOSR Journal Of Humanities And Social Science (IOSR-JHSS)
Volume 23, Issue 8, Ver. 10 (August. 2018) PP 37-45
e-ISSN: 2279-0837, p-ISSN: 2279-0845.
www.iosrjournals.org

Prediction and Evaluation of Quality in Self-Designed Computerized Adaptive Test on ICT for Usability to Assessss2 Students in Rivers State, Nigeria.

Wokoma, Tamuno-Olo Abbott and Dr. Opara, Ijeoma M.
Department of Educational Psychology, Guidance and Counseling,
Faculty of Education, University of Port-Harcourt Rivers State, Nigeria.
Corresponding Author: Wokoma, Tamuno-Olo Abbott

Abstract: In this study, the researchers evaluated quality and usability of self-designed computerized adaptive test (CAT) on ICT to assess senior secondary school two (SS2) students in Rivers state of Nigeria. Based on the investigated variables, researchers posed four research questions and also formulated four corresponding null hypotheses which guided this study. Correlational and evaluation research designs were used in this study. A sample of 122 ICT teachers were drawn from the population of ICT teachers using multistage sampling technique to form the sample of the study based on the three senatorial zones and 23 local government areas of the state. A 5-scale Likert type ICT Teachers Questionnaire (ITQ) containing five sub-sections; Calibrated Item Pool Scale (CIPS), Entry Point Scale (EPS), Item Selection Algorithm Scale (ISAS), Scoring Mechanism Scale (SMS), and, self-designed CAT suitability for use (SD-CAT-SU) was used for data collection. This was developed and validated by the researchersand eachsub-section contained 10 items. The reliability coefficients of CIPS, EPS, ISAS, SMS, and SD-CAT-SU, from Cronbach's Alpha technique for determining internal consistency are 0.81, 0.73, 0.76, 0.77 and 0.81 respectively. Data collected were subjected to linear regression analysis to answer the research questions while analysis of variance (ANOVA) and t-test associated with the regression were used to test the hypotheses at 0.05 level of significance. The results revealed that calibrated item pool, entry point, item selection algorithm, and scoring mechanism of the self-designed CAT on ICT significantly predicts its usage for assessing SS2 students. Based on the outcome of the study, it is recommended that CAT developers should carry out item analyses and calibration on written out test items, select appropriate items for pools of CATs to ensure that good instruments for assessment of traits are designed, and also use good and functioning computers for CAT administration, among others.

Keywords: Self-designed, Computerized Adaptive Test, Algorithm, Evaluation, Mechanism

---

Date of Submission: 20-08-2018 Date of acceptance: 04-09-2018

I. INTRODUCTION

Modern psychometry is currently changing the face of educational and psychological measurement byadopting the most recent technologies that are emerging into its practices. This has enhanced accurate assessment of characteristics that distinguish individuals with use of precisely designed assessment instruments referred to as tests. Onunkwo (2005), Orluwene (2012) and Iweka (2014) defined test as a device or an instrument utilized in detecting trait possessed by an individual. A test is an instrument designed and used for measuring attribute(s) in a person or an object. Tests have been in existence and use for such purposes since creation and are designed and applied for different intentions. There are numerous types of test forms and formats with each type being unique based on its objective. Some types of tests are; criterion-referenced, norm-referenced, diagnostic, placement, predictive, speed, power, teacher made, standardized, formal, formative, summative, and so forth. Each of these types of tests can also be administered using different formats such as; essay, multiple choice, fill-in-the-blank, verbal, pen-on-paper (pop), computer based, etc. In all these characteristics, types and formats of tests, one attribute stands very clear to any personnel involved in testing, and this is its quality to objectively assess the construct it is designed for.

Tests are also designed to administer items tailored to examinees' ability. This mode of test was first developed by AlfredBinet in 1905. Pencil-on-paper (POP) tests were converted into computer versions as the boundary of the new technology extended into education, psychology and other fields. This gave birth to computer-based tests (CBTs). The application of computers in testing further gave rise to computerized adaptive tests (CATs).

It was perceived that computer administered test mode would be better than the POPs in testing and that tests tailored to examiner's ability would also be more precise in determining the strength of the measured construct in such individual, hence the call foruse of CATs for testing.But one question not yet answeredis; "are CATs suitable to all testing situations?" In earnest, CAT has advantages over other testing modes. It is a test mode that is anchored on Item Response Theory (IRT) models. This principle makes it free from some systematic errors as compared to other modes that are anchored on Classical Test Theory (CTT). CAT is a sophisticated test built on some complex algorithm. This enables it to adapt itself to produce optimal items for examinees during test administration. This is the reason CAT is called tailored test because it is design to adjust its level of difficulty to match the ability of an examinee's response to previous test item while it maximize the precision of test scores during test session.

Confounds can alter even the best measuring device. Systematic and random errors are sometimes associated with most measurementprocesses. The quality of any process and its outcome depends on how meticulous and thorough operators of such processes eliminate or guide against such errors. CAT has item pools from which it draws items administered during testing. Item pools are supposed to be designed by subject test experts. The appropriateness of CAT items pool is one factor that determines its quality. The pool design process starts from determining its test objective, through item analytic and calibration processes. The CAT algorithms and other components of the computer being used also determine the effective and efficient functionality during testing. In these factors,the human factor is very prominent in determining the quality of any CAT. He sets the test objective, writes the items, chooses the item format, analyzes and calibrates them, inputs selected items into the pool, sets the CAT parameters, and so forth. Poor CAT quality noticed would likely be attributed to him because he might have failed along his line of dutyto develop a precise measuring CAT.

CATs are not currently used in Nigeria but for the researchers being aware of the drift in trend of educational and psychological testing, and for the purpose of this study, a self-design CAT on ICT for SS2 was developed, by the researchers and used. The items of the self-design CAT were taken from topics on ICT in the SS2 scheme. They were then subjected to item analytic processes to identify and select good items with adequate validity, reliability, homogeneity and friendliness to all examinees. Above all these, test must be objective. Gevshaw (2001) stated that a test is objective, if using the same scoring key. Whoever scores the test will arrive at the same score assuming no clerical errors. Slack (2014) also said that "a test that is objective measures without reference to outside influence". The researchers ensured that the items included in the CAT item pool are having standard psychometric properties.

CATs are developed and used for assessment in both low and high stakes testing. The US military developed and use the CAT-Armed services vocational aptitude battery (CAT-AS VAB) for applicants and personnel assessment. Law school admission tests (LSAT), graduate record examination (GRE) and graduate management admission (GMAT) assessment are CATs used for forms of prediction and placement of test takers. The Japanese language assessment test (J-CAT) is also another high stake CAT used to evaluate proficiency in Japanese language vocabulary, grammar reading and listening (Killoram, 2016).

CATs are also used in other settings such as in the healthcare. The researchersare of the opinion that CATs can be used to assess learners in the classrooms in Nigeria. Papanastasiou (N.D) opined that computer based learning (CBL) can have great potential if used appropriately to enhance the quality of science education. He also said that this quality can be enhanced even further with the use of computer based testing (CBT), and more especially, with computer adaptive testing (CAT).

Trucano (2015) wrote that Georgia, in the records, is identified as the first country in the world to initialize computer adaptive testing for all its school leaving certificate examinations. Georgia is a country that cannot be compared to United States of America nor 'its' contemporaries in terms of development and advancement. This feat by Georgia poses questions such as; can't this be applied in the classroom setting in Nigeria; what knowledge can learn from Georgia's experience? (Trucano, 2015). This inspired the researchers to design a CAT and carry out a predictive study on selected correlates of the CAT to evaluate its usability for testing SS2 students on ICT in Rivers State of Nigeria. Hence the researchers deemed it necessary to evaluate some components/properties of the CAT such as calibrated item pool, starting point (entry level of test), item selection algorithm, and scoring mechanism of the self-adaptive CAT.The following research questions guided the study.

1. To what extent does the calibrated item pool of the self-designed CAT on ICT predict its usage for assessing SS2 students?
2. To what extent does entry point of the self-designed CAT on ICT predict its usage for assessing SS2 students?
3. What is the extent item selection algorithm of the self-designed CAT on ICT predict its usage for assessing SS2 students?

4. What is the extent to which the scoring mechanism of the self- designed CAT on ICT predict its usage for assessing SS2 students?

The following null hypotheses were formulated and tested at 0.05 level of significance.

1. The calibrated item pool of self-designed CAT on ICT does not significantly predict its usage for assessing SS2 students.
2. Entry point of self-designed CAT on ICT does not significantly predict its usage for assessing SS2 students.
3. Item selection algorithm of the self-designed CAT on ICT does not significantly predict its usage for assessing SS2 students.
4. Scoring mechanism of the self-designed CAT on ICT does not significantly predict its usage for assessing SS2 students.

II. METHODS

The study adopted both correlational and evaluation design. The population of the study comprised the ICT teachers in the senior secondary schools in the public (government owned) post primary institution in Rivers State. A sample of 122 ICT teachers from the population in the 2017/2018 academic session were drawn from the population (source: Rivers State Senior Secondary Schools Board) through multistage sampling technique based on the three senatorial zones, 23 local Government Areas, and the 268 senior secondary schools in Rivers State. CAT on ICT for SS2 was designed and administered in schools where the teacher's samples were selected. The ICT teachers took part in the administration of the test to SS2 students.

An instrument titled ICT Teachers Questions (ITQ) that contained 5 sub-sections such as; Calibrated Item Pool Scale (CIPS), Entry Point Scale (EPS), Item Selection Algorithm Scale (ISAS), Scoring Mechanism Scale (SMS)and self-designed CAT suitability for use (SD-CAT-SU), which were developed by the researchers, were used for data collection. Each of the scale contained 10 items. The teachers were asked to assessthe self-designed CAT using the ITQ on a five-point Likert scale of "5 for true", "4 for quite true", "3 for moderate", "2 for slightly true", and "1 for not at all" respectively. The total scores for each teacher were obtained by summing the scores of his responses of the items he attended to. The maximum score for each scale is 50 while the minimum score is 10. The instruments were validated by the researchers. The reliability coefficients for each of the 5 sub-sections of the instrument (CIPS, EPS, ISAS, SMS, and SD-CAT-SU) were ascertained through Cronbach's Alpha technique. The coefficient of 0.81, 0.73, 0.76, 0.77 and 0.81 respectively were obtained. Simple regression analysis was used to answer the research questions while ANOVA and t-test associated with the regression analysis were used to test the hypotheses at 0.05 level of significance.

III. RESULTS

The results of the data analyses to answer the research questions and test the hypotheses are shown in the table below.

Research question one: To what extent does the calibrated item pool of the self-designed CAT on ICT predict its usage for assessing SS2 students?In other to answer this research question, simple regression analysis was used.

Table1: **Simple regression analysis of prediction of the calibrated item pool of the self-designed CAT on ICT usage for assessing SS2 students.**

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.880	.775	.773	3.17104	.775	412.524	1	120	.000

a. Predictor: (Constant), Calibrated Item Pool

b. Dependent Variable: Usability Of Self-Designed CAT

In table 1, the predictor and the criterion variables showed strong positive correlation value of 0.880. This table revealed the simple regression analysis of the prediction of calibrated item pool of the self-designed CAT on ICT usage for assessing SS2 students. The simple regression coefficient obtained is 0.880, the value of R^2 is 0.775 while adjusted coefficient of R^2 obtained is 0.773. Based on the R^2 of 0.775, it indicates that the prediction of calibrated item pool explains 77.5% of the variations in the self-designed CAT on ICT usage while the remaining 22.5% may be attributed to other factors outside the scope of this study.

Hypothesis 1: The calibrated item pool of self-designed CAT on ICT does not significantly predict its usage for assessing SS2 students.To ascertain the true position of this hypothesis, analysis of variance (ANOVA) was used.

Table 2: Analysis of Variance

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	4148.135	1	4148.135	412.524	.000[b]
	Residual	1206.660	120	10.056		
	Total	5354.795	121			

a. Dependent Variable: Usability Of Self-Designed CAT
b. Predictor: (Constant), Calibrated Item Pool

Table 2 indicated that calibrated item pool predicts self-designed CAT on ICT usage significantly. This is revealed in the analysis of variance associated with the regression which had an F-calculated value of 412.524 which was significant at 0.05. The significant level of 0.000 was lower than the chosen level of probability of 0.05, indicating that the null hypothesis is rejected. This implies that calibrated item pool significantly predicted self-designed CAT on ICT usage for assessing SS2 students in Rivers State.

Table 3: Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	13.764	3.325		4.140	.000
	CalibratedItemPool	1.692	.083	.880	20.311	.000

a. Dependent Variable: Usability Of Self-Designed CAT

The "sig" column showed 0.000 p-value. This is less than 0.05, meaning that the calibrated item pool contributed significantly to the model. So in Summary, the values of the regression coefficient and the constant are 1.692 and 13.764 respectively.The independent variable, calibrated item pool, showed a standardized beta coefficient of 0.880. This is the strength of the prediction of the calibrated item poolon usage of the self-designed CAT on ICT for assessing SS2 students. Hence, the hypothesis is not retained.

Research question 2: To what extent does entry point of the self-designed CAT on ICT predict its usage for assessing SS2 students?In order to answer the research question, simple regression analysis was used.

Table4: Simple regression analysis of prediction of the entry point of the self-designed CAT on ICT usage for assessing SS2 students.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.448	.201	.194	5.97203	.201	30.141	1	120	.000

a. Predictor: (Constant), Entry Point
b. Dependent Variable: Usability Of Self-Designed CAT

In table 4 entry point and usability of self-designed CAT are positively related but this is a weak one. The regression coefficient (R) showed a value of 0.448. R^2 value of 0.201 and adjusted R^2 of 0.194. Based on the R^2 value of 0.201, it reveals that the prediction of entry point only explain 20.1% of the variations in the self-designed CAT on ICT usage while 79.9% may be attributed to others factors outside the scope of this study.

Hypothesis 2: Entry point of self-designed CAT on ICT does not significantly predicts its usage for assessing SS2 students. To evaluate the statistical significance, ANOVA generated from simple regression was used.

Table 5:Analysis of Variance

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	1074.981	1	1074.981	30.141	.000[b]
	Residual	4279.814	120	35.665		
	Total	5354.795	121			

a. Dependent Variable: Usability of Self-Designed CAT
b. Predictor: (Constant), Entry Point

The test showed a variety of values from the regression sum of square of 1074.981, to the residual sum of square 0f 4279.814. From this table, the obtained F is 30.141 at aP-value of 0.000. This is less than 0.05 alpha levels, meaning that entry point of self-designed CAT on ICT significantly predicts its usage for assessing SS2 students. The statistics in coefficients table also support this fact.

Table 6:Coefficient

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	43.688	6.825		6.401	.000
	Entry Point	.937	.171	.448	5.490	.000

a. Dependent Variable: Usability of Self-Designed CAT

The p-value of 0.000 confirmed that entry point of the CAT significantly contributes to the model. The unstandardized coefficients of the constant and entry point are 43.688 and 0.937. The standardized coefficient (Beta), which is the magnitude ofprediction of entry point, is 0.448. This beta weight is significant because it has a corresponding p-value of 0.000. The null hypothesis is therefore rejected.

Research question 3: What is the extent item selection algorithm of the self-designed CAT on ICT predict its usage for assessing SS2 students? The simple regression statistics was used to answer the research question.

Table7:Simple regression analysis of prediction of item selection algorithm of the self-designed CAT on ICT usage for assessing SS2 students.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.549	.301	.296	5.58336	.301	51.771	1	120	.000

a. Predictor: (Constant), Item Selection Algorithm
b. Dependent Variable: Usability Of Self-Designed CAT

In table 7, the predictor and the dependent variables showed moderate positive correlation value of 0.549. This table also revealed the prediction of the item selection algorithm of the self-designed CAT on ICT usage for assessing SS2 students. The simple regression coefficient obtained is 0.549, the value of R^2 is 0.301 while adjusted coefficient of R^2 obtained is 0.296. Based on the R^2 of 0.301, it indicates that the prediction of item selection algorithm explains 30.1% of the variations in the self-designed CAT on ICT usage while the remaining 69.9% may be due to other factors.

Hypothesis 3: Item selection algorithm of the self-designed CAT on ICT does not significantly predict its usage for assessing SS2 students. To determine the significant prediction of item selection algorithm of self-designed CAT on ICT usage, analysis of variance associated with the regression was used.

Table 8: **Analysis of variance**

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	1613.920	1	1613.920	51.771	.000[b]
	Residual	3740.875	120	31.174		
	Total	5354.795	121			

a. Dependent Variable: Usability ofSelf-DesignedCAT
b. Predictor: (Constant), ItemSelectionAlgorithm

Table 8 indicated that item selection algorithm also predicts self-designed CAT on ICT usage significantly. This is shown in the analysis of variance associated with the regression which had an F-calculated value of 51.771 which was significant at 0.05. The significant level of 0.000 was also lower than the chosen level of probability of 0.05, indicating that the null hypothesis is also not retained. This implies also that item selection algorithm significantly predicted self-designed CAT on ICT usage for assessing SS2 students in Rivers State.

Table 9: Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	47.241	4.725		9.999	.000
	ItemSelectionAlgorithm	.850	.118	.549	7.195	.000

a. Dependent Variable: Usability ofSelf-DesignedCAT

The "sig" column in table 9 showed 0.000 p-value. This is less than 0.05, meaning that the item selection algorithm contributed significantly to the CAT usage. So in Summary, the values of the regression coefficient and the constant are 0.850 and 47.241 respectively. The independent variable, item selection algorithm, showed a standardized beta coefficient of 0.549. This is the strength of the prediction of the item selection algorithmon usage of the self-designed CAT on ICT for assessing SS2 students. Hence, the hypothesis is not retained.

Research question 4: What is the extent to which the scoring mechanism of the self- designed CAT on ICT predicts its usage for assessing SS2 students? In order to answer the research question, simple regression analysis was also used.

Table10: Simple regression analysis of prediction of the scoring mechanism of the self-designed CAT on ICT usage for assessing SS2 students.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.838[a]	.702	.699	3.64719	.702	282.555	1	120	.000

a. Predictor: (Constant), Scoring Mechanism
b. Dependent Variable: Usability Of Self-Designed CAT

In table 10, the predictor and the criterion variables showed very strong positive correlation value of 0.838. This table also revealed the simple regression analysis of the prediction of scoring mechanism of the self-designed CAT on ICT usage for assessing SS2 students. The simple regression coefficient obtained is 0.838 the value of R^2 is 0.702 while adjusted coefficient of R^2 obtained is 0.699. Based on the R^2 of 0.702, it indicates that the prediction of scoring mechanism explains 70.2% of the variations in the self-designed CAT on ICT usage while the remaining 29.8% may be attributed to some other factors outside this study scope..

Hypothesis 4: Scoring mechanism of the self-designed CAT on ICT does not predict its usage for assessing SS2 students. To determine the significant prediction of scoring mechanism of self-designed CAT on ICT usage, analysis of variance associated with the regression was used.

Table 11: **Analysis of Variance**

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	3758.554	1	3758.554	282.555	.000[b]
	Residual	1596.241	120	13.302		
	Total	5354.795	121			

a. Dependent Variable: Usability Of Self-Designed CAT
b. Predictor: (Constant), Scoring Mechanism

In this table F-statistic (282.56) had a p-value of 0.000. This is less than 0.05. These values are indicating that scoring mechanism of the self-designed CAT on ICT predicts its usage for assessing SS2 students.

Table 12: **Coefficients**

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	22.567	3.494		6.458	.000
	Scoring Mechanism	1.465	.087	.838	16.809	.000

a. Dependent Variable: Usability Of Self-Designed CAT

The p-value of 0.000 in table 12 confirmed that scoring mechanism of the CAT significantly contributes to the model. The unstandardized coefficients of the constant andscoring mechanism are 22.567 and 1.465. The standardized coefficient (Beta), which is the magnitude of prediction of scoring mechanism, is 0.838. This beta weight is significant because it has a corresponding p-value of 0.000. The null hypothesis is also rejected.

IV. DISCUSSION

The result of this study showed that calibrated item pool of the self-designed CAT on ICT predicted usage of this instrument for assessing SS2 students. The calibrated item pool explained for 77.3% of the total variation of the usability of the instrument. The F-and the P- values obtained also revealed how the calibrated pool predicted the usage of the self-design CAT on ICT for assessing SS2 students significantly. The beta coefficient of the predictor and criterion variables significantly supported this outcome as the value (0.880) revealed that the two variables are strongly related. This implies that properly calibrated item pool is very important for any computerized adaptive testing. In a similar study, Chuesathuchon (2008) constructed a mathematics CAT and used it to investigate PrathomSuksa 6 students in Thailand. His study aimed to calibrate test items, investigate validity and reliability of the test data and set up the items in an item bank for use. The outcome of the study indicated that; the item bank of equations for the PrathomSuksa 6 students contained 98 items that fitted the measurement model and consisted of nine aspects, ordered from -1.27 logits (very easy) to +1.57 logits (very hard). It also showed that test lengths, testing times and mathematical competencies were significantly different among four groups of stopping criteria studied (Chuesathuchon, 2008).

The extent to which entry point (also known as starting point) of the self-designed CAT on ICT predicts usage of the CAT for assessing SS2 students is weak. From the analysis, a value of 0.201 was obtained, meaning that entry point from the teachers' responses on the instrument explained for 20.1% of the CAT usage for accessing SS2 students on ICT. The correlation coefficient value obtained between the two valuables (0.448) revealed weak relationship. Although the relationship that exists between the predictor and the predicted variables is weak, it is positive and significant. This is further revealed in the values of the "sig" columns of ANOVA of table 5 and coefficients of table6 (which gave 0.000).This implies that entry point is a weak predictor of this CAT usage for assessing SS2 students. Empirical works related to this investigation were not found butWise and Kingsbury(2000) stated that entry point of CAT is the first item presented to examinees during test sessions. This item being presented in an adaptive testing is enhanced by the information that is made available about the examinees and the items in the test pool. (Wise & Kingsbury, 2000).Differences in such information lead to different entry points presented to examinees during test sessions. It is a characteristic correlate of item selection in CAT.Its predictive ability of this CAT usage for assessment is low. It showed moderate positive correlationand significance.

The result of this study also showed that item selection algorithm of the self-designed CAT significantly predicted usage of the CAT of the assessing SS2 students. The correlation coefficient ($r = 0.549$) and the variance explained ($R^2=0.301$) in the model summary statistics are indications of the significant prediction. From this observation, the variance explained by the predictor (item selection algorithm) is 30.1% and this has moderate positive relationship with the dependent variable (CAT usage). The F-value obtained (51.771) in the ANOVA (table 8) showed a p-value of 0.000. This implies that the null hypothesis is rejected because the item selection algorithm significantly predicted the self-design CAT usage for assessing SS2 students. In a related study, Leung, Chang and Hau (2003) compared three content balancing methods in CAT. These were constrained CAT, modified constrained CAT, and modified multi-normal model, under different text length conditions and target maximum exposure rate. Their result showed that there is no effect of content balancing method in measurement efficiency and item pool utilization. The result also revealed that the modified multi-normal model showed consistent over exposure of few items. Basically,item selection algorithm

in CATs have component parts such as item selection criterion, item exposure control, test content balancing, test length, etc.Each of these components has sub-components.

From the responses of the study subjects and analysis of data, it was revealed that there is not only positive correlation (R=0.838) between scoring mechanism and self-designed CAT usage, the predictor variable also explained for 0.702 (70.2%) of total variance in the dependent variable. This is very large. The F-valve (282.555) in table 15 and the regression coefficients in table 16 all showed P-valve of 0.000.This implies that scoring mechanism of the self-designed CAT predicts its usage for assessing SS2 students on ICT. Economides and Roupas (2007) conducted an evaluation study on computer adaptive systems. In their study they investigated the state of CAT systems and identified their strengths and weaknesses. Ten CAT systems were studied using evaluation framework of fifteen domains categorized into educational, technical and economical dimensions. Their result revealed that majority of CAT systems gave priority to score reliability, security and maintainability (Economides &Roupas, 2007).

V. CONCLUSION

The following conclusions were drawn from the result of this study.

1. Calibrated item pool of the self-designed CAT on ICT significantly predicts its usage for assessing SS2 students.
2. Entry point of the self-designed CAT on ICT significantly predicts its usage for assessing SS2 students.
3. Item selection algorithm of the self-designed CAT on ICT significantly predicts its usage for assessing SS2 students.
4. Scoring mechanism of the self-designed CAT on ICTsignificantly predicts its usage for assessing SS2 students.

VI. RECOMMENDATIONS

Based on the findings of this study, the following were recommended.

1. CAT developers should carry out item analysis and calibration on written out test items and also select the appropriate ones for item pools of CATs to ensure that good instruments for assessment of traits are designed.
2. CAT developers should trial-test instrument being developed at different stages of development to estimate the psychometric properties of the CAT, and persons within the same bracket of examinees the test is intended for should be used during trial testing.
3. Modern computer systems with well programmed and functioning algorithm should be used for test administration.
4. CATs intended for assessment purpose should meet the specification that is capable of scoring examinees and providing their maximum information after test sessions,

REFERENCE

[1]. Chuesathuchon, C (2008). Computerized Adaptive Testing in Mathematics. Faculty of Education and Arts. Edith Cowan University.Retrieved from

[2]. http://www.ro.ecu.edu.au/theses/1591 on 10th May, 2018.

[3]. Economides, A. A., &Roupes, C (2007) Evaluation of Computer Adaptive Testing Systems.

[4]. International Journal of Web-Based Learning and Teaching Technologies, 2(1), 70-87, 2007.

[5]. Gevshaw, D.A., (2001) Quality of a good test.Retrievedfromvirgil.azwestern.edu>- dag.Goodtest on 15th July, 2018.

[6]. Iweka, F.O.E. (2014). Comprehensive Guide to Test Constitution and Administration.Rivers State, Nigeria.Chifas Nigeria.

[7]. Killoram, D. (2016). Studying GMAT Critical Reasoning Consider Using LSATLR Questions.

[8]. LSAT and law School Admission Blog.Retrieved from http://blog.powerscore.com>isat>studon 3rd July, 2018.

[9]. Leung, C, Chang, H., &Hau, H. (2003) Computerized Adaptive Testing; A Comparison of Three Content Balancing Methods.Journal of Technology, Learning and Assessment. 2(5), 3-16.Retrieved fromwww.jtla.orgon 14th June, 2018.

[10]. Onunkwo, G.I.N. (2005). Tests Preparation, Administration, Communication.Onitsha, AnambraState, Nigeria: Vigo Publishers International.

[11]. Orluwene, G.W. (2012). Introduction to Test Theory and Development Process.Port Harcourt, Nigeria: Chris – Ron Integrated Services.

[12]. Papanastasiou, E. (N.D.). Computer-Adaptive testing in Science Education.University of Kansas and University of Cyprus.Department of Education.Retrieved from Citeseerx.ist.psu.edu>viewdoc>download on 25th April, 2018.

[13]. Slack, K.J. (2014).Targeting Objectivity Reliability and Validity.Retrievedfrom http://www.thewisdomthatworks.com>ton 5th July, 2018.
[14]. Trucano, M (2015). The introduction of Large Scale Computer Adaptive Testing in Georgia blogs.Retrieved fromworldbank.org>edutech>health on 6th July, 2018.
[15]. Wise, S. L. & Kingsbury G. G. (2000) Practical issues in developing and Maintaining a Computerized |Adaptive Testing Program.Psicologica 21(1), 135-155.

Wokoma, Tamuno-Olo Abbott.“ Prediction and Evaluation of Quality in Self-Designed Computerized Adaptive Test on ICT for Usability to Assessss2 Students in Rivers State, Nigeria..” IOSR Journal Of Humanities And Social Science (IOSR-JHSS). vol. 23 no. 08, 2018, pp. 37-45.