An Open Access Journal
ISSN: 1068-3844

Multicultural Education

Research Article

Homepage: www.MC-caddogap.com

STATUS AND EXPERIENCE OF MATHEMATICS TEACHERS' PERCEPTION OF INTEGRATING COMPUTER ADAPTIVE TESTING INTO UNIFIED TERTIARY MATRICULATION EXAMINATION MATHEMATICS

Oluwaseyi Aina Gbolade Opesemowo
Mdutshekelwa Ndlovu
Department of Science and Technology, Faculty of Education, University of Johannesburg, Auckland Park, 2006, South Africa

A B S T R A C T

With the increase of computers in assessment delivery today, adaptive testing has become quite popular, mainly when students must be grouped into pass or fail. The study investigated the status and experience of mathematics teachers' perception of integrating Computer Adaptive Testing (CAT) into Unified Tertiary Matriculation Examination (UTME) mathematics. A descriptive survey design was adopted for the study. The population consisted of all public and private secondary school mathematics teachers in Oyo State, Nigeria. A total of 310 mathematics teachers were purposively selected from three senatorial districts in Oyo State, Nigeria. A self-developed Computer Adaptive Testing Questionnaire showed a content validity index of 0.92 and McDonald omega of 0.76. Data were analyzed using descriptive and ANOVA at 0.05 significant level. Results revealed no statistically significant difference between mathematics teachers' status and their perception of the integration of CAT into UTME mathematics. To ensure an unbiased, valid, and reliable examination, it was determined that CAT should be integrated into UTME mathematics.

A R T I C L E I N F O

Keywords:
Computer Adaptive Testing, Computer Based Test, Unified Tertiary Matriculation Examination, Joint Admission Matriculation Board, Mathematics

Article History:
Received: 12th May 2023
Accepted: 25th Jun 2023
Published: 15th Jul 2023

1. INTRODUCTION

The call for Computerized Adaptive Testing (CAT) could not have come at a better time than now when the world has been confronted with the COVID-19 pandemic, which has affected every aspect of human endeavors. The integration of CAT in large-stake assessments like the Unified Tertiary Matriculation Examination (UTME) is fast gaining momentum as a measure of eliminating or eradicating the challenges of in person assessment caused by the pandemic. In Nigeria the virus broke out on February 27, 2020, when a foreign national tested positive in Lagos, Nigeria. According to the National Centre for Disease Control (NCDC), the pandemic affected over two hundred and sixteen countries worldwide, with seven hundred and sixty-seven million, seven hundred and fifty thousand and eight hundred and fifty-three confirmed cases and six million, nine hundred and forty one thousand and ninety-five deaths. In Nigeria, two hundred and sixty-six thousand, six hundred and seventy-five confirmed cases, with two hundred and fifty-nine thousand, nine hundred and fifty-three discharged cases and three thousand, one hundred and fifty-five confirmed fatalities (WHO, 2023). The NCDC is the nation's public health institution, and it has put in place some public health measures like any other country to ameliorate the further spread of the virus via its website (www.ncdc.gov.ng) and Twitter handle

(@NCDCgov) to ensure the general public has an update about the prevention and management strategies against COVID-19. Globally, the pandemic has had a significant impact on the education system. In the study of Mamun and Ullah (2020), it was stated that since the beginning of April 2020, nearly 90% of the total enrolled students (i.e., 1.5 billion students) from 185 countries had been involved in little or no educational activities because of the outbreak of the COVID-19 pandemic that is ravaging the world leaving the government of nations with no choice than to impose restrictions of movement in other to curtail further spread of the virus. Also, the virus has spread into 58 African countries with over two hundred thousand infections and more than 7 million infected persons worldwide.

The World Health Organisation (WHO) has a heavy burden in proffering solutions to the pandemic and has prescribed policy directions in solving the pandemic. A framework of maintaining high environmental hygiene and the introduction of restrictions like lockdowns and curfews to reduce the spread of the virus was recommended by health expert. In compliance with the COVID-19 guidelines, Nigeria, like other parts of the world, imposed a lockdown that virtually distorted all educational activities of teaching, learning, and assessment. The Unified Tertiary Matriculation Examination (UTME) is among the basic requirements for securing admission into higher institutions in Nigeria. The Joint Admission and Matriculation Board (JAMB) was established in 1978 and saddled with the responsibility of conducting UTME into all forms of tertiary institutions in Nigeria, i.e., a uniform standard matriculation educational assessment and placement for prospective candidates into the nation's universities UTME assessment focuses on four subjects depending on the candidate's course of choice. However, in this study, the Mathematics examination is the primary focus. Although in 2015, JAMB launched the Computer Based Test (CBT), a full-scale assessment of candidates as a response to meet the rising challenges of the paper and pencil test. Based on this paradigm shift, candidates need to keep abreast of this new development and build their computer literacy skills to the optimum level. In today's world of the fourth Industrial Revolution (4IR), which is internet driven and powered by technology, educational tests correlate with education and training as a way to measure candidates' performance level (Georgiadou et al., 2006), and it is one of the core areas of education which cannot be left behind. Similarly, candidates should be acquainted with technology to fit into new modes of assessment in the era of 4IR (Ayanwale et al., 2022; Oladele et al., 2022).

JAMB as an examination body shifted their assessment from paper and pencil to CBT. CBT is a linear form of test administration where examination questions are deployed via computer, and responses are recorded and examined electronically. Linear and Adaptive CBT are the two types of CBT. Presently, linear CBT is used to assess UTME scores in Nigeria. The UTME is conducted annually for candidates seeking admission into institutions of higher learning. Furthermore, a linear test is a full-length examination in which the computer selects various items for individuals without considering their ability level. Adaptive CBT, or CAT, is a testing procedure that employs on-the-fly techniques to align with candidates' ability levels in enhancing the accuracy of testing (Reckase, 2010; Redecker & Johannessen, 2013). CAT is a technologically-driven method for conducting educational assessments that have been utilized for many years (Oladele, 2021). Chang (2015) opined that CAT had become an increasingly important testing mode in large-scale assessment. CAT is seen to be a more reliable mode of test administration than the usual paper and pencil testing because it is only those items whose level of difficulty matches the ability of the candidates that are delivered to the candidates (Adeyemo & Opesemowo, 2020; Chang, 2015; Chuesathuchon, 2008; Jimoh, 2022; Opesemowo et al., 2023). This will drastically reduce any form of examination malpractice and make students self-reliant and dependable. CAT assesses the capabilities of every candidate accurately while simultaneously solving the problems encountered under the paper and pencil tests because only those items considered appropriate are selected and administrated to the candidates. CAT is based on the item response theory model (IRT) (i.e., 1, 2. 3, & 4 IRT models), calibrated item pool, item selection algorithm, a statistical method for locating candidate on the construct, and rule for stopping the test.

CAT enables the administration of test items from a calibrated bank, with an emphasis on items that offer the most information about students' previous answers to questions (Fries et al., 2009; Fries & Krishnan, 2009; Fries et al., 2014). For Fries et al. (2014), there are three general designs for CAT applications to construct educational assessments. To begin with, CAT can be given to each student on each occasion. It implies that the student will receive the questions at the end of the assessment that they did at the start. The second design would be to use CAT as a benchmark to assign scores to each student and then administer the same items previously selected by CAT at the final observation. If a student improves, the items chosen at baseline could restrict measurement of the improvement by imposing limitations rather than allowing the CAT algorithm (Oladele, 2021; Oladele & Ndlovu, 2021; Oladele et al., 2022) and select items at higher levels of the construct. The third design includes items usually selected and assembled into a static short form that can be administered to all students. In this case, all students fill out the same form throughout the study, customized in advance to be preferentially targeted to the student population level. This implies that throughout the assessment, all students complete the same form of test, with a form that has been customized in advance to be preferentially targeted to the level of the student population. CAT involves the use of technology, and it is imperative to consider how it (CAT) is perceived among the status and years of experience of the mathematics teachers, which are some of the demographic characteristics in this study. Regarding age, Ayanwale et al. (2022) reported that teachers in lower age brackets were perceived to embrace technology in teaching and learning quicker.

Though there has been a dearth of research conducted on CAT, those that have been done (Bulut & Kan, 2012; Ogunjimi et al., 2021; Oladele et al., 2020; Oladele & Ndlovu, 2021; Stone & Davey, 2011; Thompson, 2017) have shown that comparing CAT to paper-based testing is more effective in terms of time management and the student' ability level are considered in adapting the test items. 4IR plays a critical role in the success of the post-industrial era of education assessment. If JAMB can tap into this development, it will provide a better valid and reliable candidates' score and solve the problems associated with paper and pencil examination while combating the scourge of examination malpractices.

Problem Identification

The involvement of some candidates in examination malpractices such as bringing in unauthorized materials, writing on currency notes and identity cards, spying on other candidates in the examination hall, substitution of answer sheets, impersonation, desperate attempts to see questions before the examination, connivance with supervisors and school authorities to cheat, body writing or tattoo in which students especially females write on hidden parts of their bodies (Onyibe et al., 2015) among others bedeviling public examinations made JAMB in 2015 to introduce a full-scale CBT. The introduction of CBT came with numerous challenges (such as epileptic power failure, poor ICT funding, inadequate ICT workforce/skills, and lack of access to the internet, e.t.c.). However, poor ICT infrastructure was identified as a significant problem (Onyibe et al., 2015).

Educational experts should continue searching for the best assessment mode and ensure that reliability is not tempered. The first of the fourth generations of computerized educational assessment is the linear CBT (Redecker & Johannessen, 2013) which is presently deployed for assessments leading to UTME scores used for placement of prospective undergraduates into institutions of higher learning. Nevertheless, this method of educational assessment failed to factor in the candidate's ability with the choice of the item they attempted and inability to generate the candidate's score immediately after the test. A significant advantage of CAT is that it places items and candidates on the same scale, directly matching candidates to the most valuable items. In leveraging CAT, this study hopes to assess the status and experience of mathematics teachers' perception toward integrating CAT into UTME mathematics.

2. THEORETICAL FRAMEWORK: MARTINGALE AND CAT

For this study, the Martingale theory was adopted, which is based on establishing the limitation of the distribution of $\hat{\theta}$ which is critical for furthering CAT research. The theory is a plethora of theoretical tools for obtaining significant sample results when analysing independent random variables, such as the central limitation theorem, the law of large numbers, and inequality. This theory becomes essential because the UTME is a large-scale assessment conducted across the thirty-six (36) states, including the Federal Capital Territory of Nigeria. In CAT, the vast sample size property of $\hat{\theta}$ should be calculated using a series of dependent random variables because the *K*th item is chosen according to the preceding K – I responses X_1, ….., X_{k-1}, and hence X_1, ….., X_k are generally dependent. Nonetheless, such dependence emerges in a mathematically fascinating manner. It is considered that X_1…...,X_k, which represents the observed arrangement at a specific time K, contains the earlier sequence X_1, ….., X_{k-1}. CAT allows for the dynamic administration of items from a calibrated bank (Chang, 2015), with an emphasis on items that provide the most incredible information increment given a subject's prior responses. Items that are frequently selected in a test population can be aggregated to form a static short form tailored to a portion of the strictness range represented by the test population (Leung et al., 2014).

Martingale's theory is frequently employed when examining a series of recursive estimates with an individual dependency. Recently, practitioners have become aware of the martingale theory for its significant contribution to resolving the CAT problem that arose from educational testing, which no one thought was related. To demonstrate the relationship between CAT and martingale, let X_1,…., X_n represent a series of dependent random variables with joint density $f\,(\chi_1, \dots, \chi_n)$. Let $f\,(\chi_k | \chi_1, \dots, \chi_k - 1)$ be the conditional density of X_k given $X_1, \dots., X_{k-1}$. Then

$$f(\chi_1, \dots, \chi_n) = \prod_{k=1}^{n} f(\chi_k | \chi_1, \dots, \chi_k - 1)$$

To consider the null hypothesis $H_0 : f = f_0$ versus the alternative $H_1 : f = f_1$. Let L_n be the likelihood ratio

$$L_n = \frac{f_1(X_1, \dots\dots, X_n)}{f_0(X_1, \dots\dots X_n)}$$

Then, it can be verified that under H_0, L_n is a martingale with respect to σ – filteration F_n,

$$E(L_n | X_1, \dots\dots\dots, X_{n-1}) \quad (1)$$

Where

$$F_n = \sigma(X_1, \dots., X_n) \quad (2)$$

It is important to note that some probability textbooks present martingale with a gambling example explaining a fair game. In Equation (1), if L_{n-1} represents a gambler's current fortune and $(X_1, \dots\dots, X_{n-1})$ is a σ-field containing all the information about the game after the $((n-1)$th play. The equation states that the expected fortune after the next play is the same as the present fortune, which implies that the fortunes of the gambler and the house are equally weighted. Researchers (Chang, 2015; Chang & Ying, 1999) have shown that martingales exclude the possibility of winning strategies based on game history and thus are models of fair games. However, the study focused on integrating the CAT framework into UTME mathematics items. Based on this, the study seeks to determine the perception of mathematics teachers towards integrating CAT into UTME mathematics items viz-a-viz the demographic variables of the mathematics teachers are statistically different regarding their perception towards integrating CAT into UTME mathematics.

3. CAT FRAMEWORK

Thompson and Weiss (2011) proposed the CAT framework, indicating that it is a powerful method of delivering examinations. Item response theory (IRT) is the foundation for most CATs. In terms of test construction, item analysis, and examinee scoring, IRT is a robust psychometric paradigm.

Table 1: Proposed CAT Framework

Step	Stage	Primary work
1.	Feasibility, applicability, and planning studies	Monto Carlo simulation; business case evaluation
2	Develop item bank content or utilize existing bank	Item writing and review
3.	Pretest and calibrate item bank	Pretesting; item analysis
4.	Determine specifications for final CAT	Post-hoc or hybrid simulations
5.	Publish live CAT	Publishing and distribution; software development

Source: Thompson & Weiss (2011)

The first stage of the CAT framework is the feasibility, applicability, and planning studies. At this stage, items are developed to ascertain whether the CAT approach is feasible for a testing program. Because the CAT algorithm is so conceptually appealing and offers certain advantages, stakeholders might become captivated by the idea and wish to proceed without knowing anything about CAT. This step employs Monto Carlo stimulation and business case evaluation.

The second stage is the development of an item bank. Once the decision to convert to CAT has been made, step to create an item bank will be taken. When possible, this should be done based on empirical evidence. The previous step's simulation studies should be used and may be expanded to provide instructions for the bank; as indicated by Veldkamp and van der Linden (2010), simulations are valuable for this step and are not necessarily limited to use after pilot testing as described by Flaugher (2000).

The third stage is the pretesting, calibrating, and linking of items. At this stage, items must be pretested after they have been designed. This is critical for CAT because items are matched to candidates based on Item Response Theory (IRT) item parameters, which are determined by statistical analysis of real candidates' answers to items.

The fourth stage is the determination of specifications for the final CAT. At this stage, an item bank has been created and calibrated with IRT. However, this is simply the first of five CAT components originally stated. The four components must be determined before the CAT can be published and distributed. At this stage, planning of the item bank should be based on simulation studies rather than random decisions.

The last stage is publishing live CAT. The final CAT can be published once the specifications for all necessary components and any new algorithms have been specified.

4. METHODOLOGY

The study used a descriptive survey research method to investigate the perceptions of mathematics regarding integrating CAT into UTME mathematics in Oyo State, Nigeria. The population for the study consisted of mathematics teachers in Oyo State, Nigeria. The population for the study comprised mathematics teachers from public and private secondary schools, with a purposive sampling technique used to select 310 participants (118 female and 192 male) from three senatorial districts in Oyo State. The data obtained from the questionnaire were subjected to descriptive and ANOVA analysis at a significant level of 0.05. Their ages range from 20 to 56 years, with a mean of 36.59. The instrument has two sessions. The first session contains demographic variables (i.e., gender, school ownership, age, and years of experience) of the mathematics teachers. Session two includes questions about integrating the CAT framework into the UTME mathematics examination. On a four-point Likert scale, respondents indicated their extent level based on statements ranging from '4=Very large extent' to '3= Large extent' to '2 = Some extent' and '1=Not at all'. Based on literature reviews and 25 pools of items were developed. The validity and reliability of these items were assessed based on the content. After reviewing their suggestions, three computer science experts selected 20 items based on clarity, language use, relevance, and readability of the construct underpinning the study. The surviving items had a content validity index (CVI) proposed by Lawshe (1975) and Baghestani et al. (2019), the most widely reported approach for content validity (Ayanwale, 2023; Grant & Davis, 1997; Zamanzadeh et al., 2015). After the rating by the five panelists in terms of "essentiality and non-essentiality," the returned instrument revealed an index of 0.92, and MacDonald Omega reliability (Chowdhary et al., 2020) implemented in the user-friendly package of Jamovi gave an index of 0.76. In Jamovi software version 2.3.28 (Jamovi project, 2021), descriptive statistics and one-way analysis of variance (ANOVA) were applied to the analyzed data. Open-source spreadsheet software Jamovi is a third-generation statistical spreadsheet built with R (R Core Team, 2021). A further step was taken to recode negatively worded items before analyzing them.

5. RESULT

A preliminary investigation was performed to verify some underlying assumptions before conducting the primary analysis for the study. The mathematics teachers' responses were tested using the normality and homogeneity of variance assumption. In addition, Shapiro Wilk's test revealed no statistically significant in the dataset, with kurtosis and skewness values falling within the trench old of -2.58 to +2.58 (Byrne, 2016; Hair et al., 2010; Opesemowo et al., 2022). An insignificant value was observed with $p > 0.05$ using Levene's test of homogeneity of variance. Therefore, the respondents in the study were not statistically different, meaning they have similar characteristics as mathematics teachers regardless of their demographic variables, i.e., regardless of status and years of experience, mathematics teachers exhibit similar traits. The subsequent analysis phase can be carried out since the two major assumptions have been fulfilled. Analysis of Variance (ANOVA) was used to assess mathematics teachers' perception toward integrating CAT into the UTME mathematics test. The demographic variables of mathematics teachers were also evaluated using ANOVA. This was performed on the item level, and the overall scale determined how their perceptions differed according to their demographic variables. The results are presented in Table 1.

Table 1: one-way ANOVA for the individual item across the status of mathematics teachers' perception in integrating CAT into UTME mathematics

	Status	N	Mean	SD	F	p
ITEM 1	Principal	10	3.50	0.53	0.38	0.77
	Vice-Principal	20	3.55	0.89		
	Head of Department	36	3.31	0.92		
	Teacher	244	3.43	0.74		
ITEM 2	Principal	10	3.60	0.52	1.80	0.17
	Vice-Principal	20	3.40	0.75		
	Head of Department	36	3.36	0.68		
	Teacher	244	3.61	0.61		
ITEM 3	Principal	10	3.00	0.94	0.98	0.42
	Vice-Principal	20	3.50	0.83		
	Head of Department	36	3.19	0.79		
	Teacher	244	3.19	0.76		
ITEM 4	Principal	10	3.30	1.06	0.21	0.89
	Vice-Principal	20	3.45	0.60		
	Head of Department	36	3.42	0.69		
	Teacher	244	3.49	0.66		
ITEM 5	Principal	10	3.80	0.63	2.86	0.05
	Vice-Principal	20	3.35	0.67		
	Head of Department	36	3.28	0.88		
	Teacher	244	3.19	0.93		
ITEM6	Principal	10	2.20	1.32	0.93	0.44
	Vice-Principal	20	2.50	1.10		
	Head of Department	36	2.31	1.04		
	Teacher	244	2.12	1.05		
ITEM7	Principal	10	2.30	1.34	0.77	0.52
	Vice-Principal	20	1.95	1.23		
	Head of Department	36	2.08	0.94		
	Teacher	244	1.87	1.01		
ITEM 8	Principal	10	3.50	0.97	1.00	0.41
	Vice-Principal	20	3.40	0.60		
	Head of Department	36	3.31	0.75		
	Teacher	244	3.19	0.94		
ITEM 9	Principal	10	3.30	1.06	0.36	0.78
	Vice-Principal	20	2.95	0.83		
	Head of Department	36	3.14	0.68		
	Teacher	244	3.11	0.82		
ITEM 10	Principal	10	3.10	1.29	1.98	0.14
	Vice-Principal	20	3.45	0.83		
	Head of Department	36	2.81	1.09		
	Teacher	244	3.15	0.91		
ITEM11	Principal	10	3.00	1.05	0.54	0.66
	Vice-Principal	20	2.90	0.97		
	Head of Department	36	2.64	0.99		

	Teacher	244	2.70	0.91		
ITEM12	Principal	10	3.10	0.88	0.27	0.85
	Vice-Principal	20	3.20	0.77		
	Head of Department	36	3.03	0.88		
	Teacher	244	3.16	0.83	1.78	0.17
ITEM13	Principal	10	2.90	1.29		
	Vice-Principal	20	3.50	0.95		
	Head of Department	36	3.25	0.91		
	Teacher	244	3.54	0.77		
ITEM14	Principal	10	3.70	0.67	2.32	0.10
	Vice-Principal	20	3.60	0.50		
	Head of Department	36	3.17	0.85		
	Teacher	244	3.50	0.68		
ITEM15	Principal	10	3.50	0.53	5.13	0.01
	Vice-Principal	20	3.00	0.79		
	Head of Department	36	2.69	0.75		
	Teacher	244	2.88	0.82		
ITEM16	Principal	10	2.80	1.14	0.23	0.87
	Vice-Principal	20	2.90	0.85		
	Head of Department	36	3.06	0.83		
	Teacher	244	2.98	0.84		
ITEM17	Principal	10	3.40	0.70	1.81	0.17
	Vice-Principal	20	3.05	0.51		
	Head of Department	36	2.92	0.81		
	Teacher	244	2.91	0.84		
ITEM18	Principal	10	2.60	1.17	0.45	0.72
	Vice-Principal	20	2.65	0.93		
	Head of Department	36	2.58	0.94		
	Teacher	244	2.45	0.98		
ITEM19	Principal	10	2.70	1.16	1.16	0.34
	Vice-Principal	20	3.05	0.89		
	Head of Department	36	2.72	1.00		
	Teacher	244	3.02	0.88		
ITEM20	Principal	10	3.20	0.92	0.97	0.42
	Vice-Principal	20	2.65	1.09		
	Head of Department	36	2.83	0.85		
	Teacher	244	2.73	0.90		

Table 1 presents the results of the one-way ANOVA conducted to analyze the perception of mathematics teachers regarding the integration of CAT into UTME mathematics based on their status. Among the 20 items measured, only item 15 showed a statistically significant difference. The ANOVA test yielded an F-value of 5.13 with degrees of freedom (3, 306) and a p-value of 0.01, which is less than the significance level of 0.05. This indicates that the rank or position of mathematics teachers can influence their perception of integrating CAT into UTME mathematics. It suggests that teachers' roles as primary drivers of academic innovations can influence their acceptance and adoption of CAT in the context of UTME mathematics.

On the other hand, the remaining 19 items, including items 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, and 20, did not show any statistically significant difference among the various ranks of the mathematics teachers. The ANOVA tests for these items yielded F-values ranging from 0.05 to 2.32, with corresponding p-values ranging from 0.05 to 0.89, all of which are greater than the significance level of 0.05. This implies that the ranks of mathematics teachers did not significantly impact their perception of integrating CAT into UTME mathematics for these particular items. In summary, the statistical analysis of the data suggests that while the rank or position of mathematics teachers may influence their perception of CAT integration in UTME mathematics for item 15, it does not significantly impact their perception of the remaining 19 items. However, ANOVA was conducted on the status of mathematics teachers' perception toward integrating CAT into UTME mathematics. Further analysis was performed using ANOVA analysis on the overall perception of mathematics teachers' status in integrating CAT into UTME mathematics in Table 2.

Table 2: One-way ANOVA on the perception of mathematics teachers across status in integrating CAT into UTME mathematics

	Status	**N**	**Mean**	**SD**	**df**	F	**p**
CAT Integration	Principal	10	59.80	11.82	3	0.67	0.57
	Vice-Principal	20	58.95	8.50	306		
	Head of Department	36	56.36	8.13			
	Teacher	244	57.23	8.58			

Table 2 displays the mean and standard deviation of the CAT integration across the different status of mathematics teachers. It showed that those in the principal status had the highest mean score with N = 10, $\bar{x}$ = 59.80, SD = 11.82, following were those in the category of vice-principal with N = 20, $\bar{x}$ = 58.95, SD = 8.50, followed by teacher with N = 244, $\bar{x}$ = 57.23, SD = 8.58, and the least was those in the status of Head of Department with N = 36, $\bar{x}$ = 56.36, SD = 8.13. However, one-way ANOVA was utilised to determine the statistical difference. It was discovered that there was no statistically significant difference between mathematics teachers at different status ($F (3, 306)) = 0.67$, $p > 0.05$, and their perception of integrating CAT into UTME mathematics. By implication, the null hypothesis was accepted. The study went further to consider the mathematics experience of the teachers presented in Table 3.

Table 3: One-way ANOVA for an individual item across years of experience of mathematics teachers' perception in integrating CAT into UTME mathematics

	Experience	**N**	**Mean**	**SD**	**F**	**P**
ITEM1	1.00	153.00	3.42	0.70	1.28	0.28
	2.00	91.00	3.35	0.82		
	3.00	53.00	3.49	0.89		
	4.00	13.00	3.77	0.44		
ITEM2	1.00	153.00	3.62	0.56	3.95	0.01
	2.00	91.00	3.43	0.70		
	3.00	53.00	3.57	0.72		
	4.00	13.00	4.00	0.00		
ITEM3	1.00	153.00	3.12	0.80	3.85	0.01
	2.00	91.00	3.14	0.82		
	3.00	53.00	3.45	0.61		
	4.00	13.00	3.62	0.51		
ITEM4	1.00	153.00	3.47	0.64	0.71	0.55
	2.00	91.00	3.42	0.76		
	3.00	53.00	3.51	0.67		
	4.00	13.00	3.69	0.48		
ITEM5	1.00	153.00	3.27	0.89	1.00	0.39
	2.00	91.00	3.25	0.89		
	3.00	53.00	3.17	0.96		
	4.00	13.00	2.85	0.99		

ITEM6	1.00	153.00	2.08	1.03	1.12	0.34
	2.00	91.00	2.29	1.06		
	3.00	53.00	2.28	1.18		
	4.00	13.00	1.92	1.04		
ITEM7	1.00	153.00	1.79	0.95	2.22	0.09
	2.00	91.00	2.02	1.06		
	3.00	53.00	2.15	1.17		
	4.00	13.00	1.69	0.95		
ITEM8	1.00	153.00	3.21	0.89	0.08	0.97
	2.00	91.00	3.25	0.90		
	3.00	53.00	3.25	0.94		
	4.00	13.00	3.15	0.99		
ITEM9	1.00	153.00	3.16	0.75	0.80	0.49
	2.00	91.00	3.08	0.91		
	3.00	53.00	3.00	0.83		
	4.00	13.00	3.31	0.63		
ITEM10	1.00	153.00	3.10	0.93	2.00	0.11
	2.00	91.00	3.00	1.04		
	3.00	53.00	3.38	0.81		
	4.00	13.00	3.31	0.85		
ITEM11	1.00	153.00	2.71	0.86	1.80	0.15
	2.00	91.00	2.65	1.02		
	3.00	53.00	2.94	0.97		
	4.00	13.00	2.38	0.77		
ITEM12	1.00	153.00	3.12	0.83	2.84	0.04
	2.00	91.00	3.07	0.87		
	3.00	53.00	3.19	0.81		
	4.00	13.00	3.77	0.44		
ITEM13	1.00	153.00	3.46	0.84	2.34	0.07
	2.00	91.00	3.37	0.88		
	3.00	53.00	3.62	0.71		
	4.00	13.00	3.92	0.28		
ITEM14	1.00	153.00	3.48	0.72	0.56	0.64
	2.00	91.00	3.48	0.67		
	3.00	53.00	3.42	0.72		
	4.00	13.00	3.69	0.48		
ITEM15	1.00	153.00	2.86	0.80	0.55	0.65
	2.00	91.00	2.85	0.84		
	3.00	53.00	3.00	0.81		
	4.00	13.00	3.00	0.71		
ITEM16	1.00	153.00	2.95	0.83	1.12	0.34
	2.00	91.00	3.11	0.81		
	3.00	53.00	2.87	1.00		
	4.00	13.00	2.92	0.49		
ITEM17	1.00	153.00	2.82	0.85	2.50	0.06

	2.00	91.00	2.99	0.80		
	3.00	53.00	3.15	0.77		
	4.00	13.00	3.00	0.41		
ITEM18	1.00	153.00	2.35	1.00	2.60	0.06
	2.00	91.00	2.54	0.93		
	3.00	53.00	2.70	0.95		
	4.00	13.00	2.85	0.80		
ITEM19	1.00	153.00	3.05	0.94	1.56	0.20
	2.00	91.00	2.82	0.89		
	3.00	53.00	2.98	0.91		
	4.00	13.00	3.23	0.44		
ITEM20	1.00	153.00	2.71	0.92	0.36	0.78
	2.00	91.00	2.82	0.90		
	3.00	53.00	2.72	0.91		
	4.00	13.00	2.85	0.90		

Table 3 demonstrated the one-way ANOVA for individual items of perception in integrating CAT into UTME mathematics based on the teachers' years of experience. It depicts that only item 2 with F (3, 306) = 3.95, $p = 0.01 < 0.05$, item 3 with F (3, 306) = 3.85, $p = 0.01 > 0.05$, and item 12 with F (3, 306) = 2.84, $p = 0.04 < 0.05$ respectively, were statistically significant. This indicates that the mathematics teachers' years of experience will have an impact on the integration of CAT into UTME mathematics because teachers play a critical role in the success of any academic innovation since they have direct contact with the students while the remaining 17 items including items 1, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 16, 17, 18, 19, and 20 indicate there was no statistically significant difference among the various years of experience of the mathematics teachers. The ANOVA tests for these items revealed F-values ranging from 0.05 to 2.60, with a p-value ranging from 0.05 to 0.97, all of which were greater than 0.05 significance level. This implies that the years of experience of mathematics teachers did not significantly impact their perception of integrating CAT into UTME mathematics for those items. Consequently, ANOVA was conducted on the entire perception of mathematics teachers toward integrating CAT into UTME mathematics. Further analysis was performed using ANOVA analysis on the overall perception of mathematics teachers' years of experience in integrating CAT into UTME mathematics in Table 4.

Table 4: One-way ANOVA on the overall perception of mathematics teachers across years of experience in integrating CAT

	Experience	N	Mean	SD
CAT Integration	0 -10 Years	153.00	56.72	8.36
	11 – 20 Years	91.00	57.11	9.23
	21 – 30 Years	53.00	58.85	8.63
	31 – 40 Years	13.00	59.69	6.91
		F (3, 306) = 1.15		P = 0.32

Table 4 displays the mean and standard deviation of the CAT integration across the various years of experience of the mathematics teachers. It was observed that teachers who indicated that they had spent 31 – 40 years had the highest mean score with N = 13, $\bar{x} =$ 59.69, SD = 6.91, following were those who have spent 21 – 30 years teaching mathematics with N = 53, $\bar{x}$ = 58.85, SD = 8.63, followed by teachers who have spent 11 – 20 years with N = 91, $\bar{x}$ = 57.11, SD = 9.23, and the least were those who have spent 0 – 10 years with N = 153, $\bar{x}$ = 56.72, SD = 8.36. However, one-way ANOVA was used to determine the statistical difference. It was noticed that there was no statistically significant difference between mathematics teachers' years of experience with (F (3, 306)) = 0.67, $p > 0.05$, and their perception of the integration of CAT into UTME mathematics. Therefore, the null hypothesis was accepted.

6. DISCUSSIONS

This study involves the integration of the CAT framework into UTME mathematics. The integration of CAT into any examination like UTME mathematics is to improve the quality of items, ensuring proper and unbiased administration of items. CAT is beneficial for teachers in instructional planning and educational assessment, enabling teachers to design questions for each student. It would provide greater accuracy and dependability of educational assessment about each student's understanding and allow teachers to identify better areas where students need additional clarification (Chang, 2015; Chiu, 2021). The study went further to consider the mathematics teachers' perception toward integrating CAT into UTME mathematics viz-a-viz the demographic structures (including status and years of experience) of the mathematics teachers towards integrating CAT into UTME mathematics.

Regarding the status of mathematics teachers' perception towards integrating CAT into UTME mathematics, the study noticed that mathematics teachers at the principal status had a higher mean score than their rest counterparts. This means that the principals in high school who are mathematics teachers had a better perception of integrating CAT into UTME mathematics than follow counterparts. The study further reviewed that there was no statistically significant difference between mathematics teachers' years of experience and their perception of integrating CAT into UTME mathematics. This result opposed the study of Ayanwale (2023), who claimed a statistically significant difference between teachers' years of experience and their adoption of relevant 4IR skill sets. Adaptive measurement was a more sophisticated and efficient way of delivering examinations than non-adaptive ones (Bulut & Kan, 2012). Developed nations like the United States of America are leveraging on CAT for accurate ability placement. It is now imperative for African countries like Nigeria to adopt CAT in high-stakes examinations to improve the conduct of educational assessments. Ayanwale et al. (2022) report was also buttressed in the submissions of (Giordano, 2007; Gorder, 2008; Hernández-Ramos, 2005; Wong & Li, 2008), who presented that the use of 4IR devices in classroom instruction is influenced by teaching experience. In addition, more experienced teachers are more likely than less experienced teachers to use 4IR devices to improve the teaching and learning of mathematics education and better students' performance in mathematics examinations. The findings of this study were supported by Oladele et al. (2023), who reported that teachers with more experience use technology in the classroom more occasionally than those with lesser experience. According to Russell et al. (2003), it was stated that teachers with less experience but high technology proficiency did not incorporate technology into their lessons. A teacher with little experience teaching mathematics may focus on using technology rather than integrating it into the classroom. Researchers (Ayanwale et al., 2022; Baek et al., 2008; Niederhauser & Stoddart, 2001) revealed that experienced teachers are less likely to integrate CAT into classroom assessment.

7. CONCLUSION AND RECOMMENDATIONS

The study focused on status and experience mathematics teachers’ perception in integrating the CAT framework into UTME mathematics to improve the quality of UTME mathematics items. We clinched that the CAT framework should be integrated into UTME mathematics in other to have an unbiased assessment for students. The study then recommended that JAMB, an institution of Nigeria's government agent responsible for conducting UTME, integrate CAT into UTME mathematics, enhancing mathematics education. In addition, experts (e.g., Psychometricians) should be consulted on how to use CAT to develop and administer UTME mathematics. This study was restricted to data retrieved from mathematics teachers from the three senatorial districts in Oyo State, Nigeria, both in private and public high schools. This constrained the generalizability of the findings since the scope of the study was geographically limited to only one subject teacher, i.e., mathematics teachers in high schools, and further studies can be conducted using a more comprehensive sample size to have a more general inference.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Adeyemo, E. O., & Opesemowo, O. (2020). Differential Test let Functioning (DTLF) in Senior School Certificate Mathematics Examination Using Multilevel Measurement Modelling. Sumerianz Journal of Education, Linguistics and Literature, 3(11), 249-253.
2. Ayanwale, M. A. (2023). Can Experience Determine the Adoption of Industrial Revolution 4.0 Skills in 21st Century Mathematics Education? Research in Social Sciences and Technology, 8(1), 74-91.
3. Ayanwale, M. A., Ndlovu, M., & Oladele, J. I. (2022). Mathematics Education and the Fourth Industrial Revolution: Are the High School Mathematics Teachers Ready? In B. Chirinda, K. Luneta, & A. Uworwabayeho

(Eds.), Mathematics Education in Africa: The Fourth Industrial Revolution (pp. 77-96). Springer International Publishing. https://doi.org/10.1007/978-3-031-13927-7_5

4. Baek, Y., Jung, J., & Kim, B. (2008). What makes teachers use technology in the classroom? Exploring the factors affecting facilitation of technology with a Korean sample. Computers & Education, 50(1), 224-234. https://doi.org/10.1016/j.compedu.2006.05.002
5. Baghestani, A. R., Ahmadi, F., Tanha, A., & Meshkat, M. (2019). Bayesian Critical Values for Lawshe's Content Validity Ratio. Measurement and Evaluation in Counseling and Development, 52(1), 69-73. https://doi.org/10.1080/07481756.2017.1308227
6. Bulut, O., & Kan, A. (2012). Application of computerized adaptive testing to entrance examination for graduate studies in Turkey. Eurasian Journal of Educational Research, 49, 61-80.
7. Byrne, B. M. (2016). Structural equation modeling with AMOS: Basic concepts, applications, and programming. routledge.
8. Chang, H.-H. (2015). Psychometrics Behind Computerized Adaptive Testing. Psychometrika, 80(1), 1-20. https://doi.org/10.1007/s11336-014-9401-5
9. Chang, H.-H., & Ying, Z. (1999). a-Stratified Multistage Computerized Adaptive Testing. Applied Psychological Measurement, 23(3), 211-222. https://doi.org/10.1177/01466219922031338
10. Chiu, T. K. F. (2021). A Holistic Approach to the Design of Artificial Intelligence (AI) Education for K-12 Schools. TechTrends, 65(5), 796-807. https://doi.org/10.1007/s11528-021-00637-1
11. Chowdhary, N., Kaurav, R. P. S., & Sharma, S. (2020). Segmenting the domestic rural tourists in India. Tourism Review International, 24(1), 23-36. https://doi.org/154427220X15791346544761
12. Chuesathuchon, C. (2008). Computerized adaptive testing in mathematics for primary schools in Thailand Edith Cowan University]. https://ro.ecu.edu.au/theses/1591
13. Flaugher, R. (2000). Item Pools. In Wainer, H. (Ed.) Computerized adaptive testing: A Primer. Mahwah, NJ: Erlbaum.
14. Fries, J. F., Cella, D., Rose, M., Krishnan, E., & Bruce, B. (2009). Progress in Assessing Physical Function in Arthritis: PROMIS Short Forms and Computerized Adaptive Testing. The Journal of Rheumatology, 36(9), 2061-2066. https://doi.org/10.3899/jrheum.090358
15. Fries, J. F., & Krishnan, E. (2009). What constitutes progress in assessing patient outcomes? Journal of Clinical Epidemiology, 62(8), 779-780. https://doi.org/10.1016/j.jclinepi.2008.12.001
16. Fries, J. F., Witter, J., Rose, M., Cella, D., Khanna, D., & Morgan-DeWitt, E. (2014). Item Response Theory, Computerized Adaptive Testing, and PROMIS: Assessment of Physical Function. The Journal of Rheumatology, 41(1), 153-158. https://doi.org/10.3899/jrheum.130813
17. Georgiadou, E., Triantafillou, E., & Economides, A. A. (2006). Evaluation parameters for computer-adaptive testing. British Journal of Educational Technology, 37(2), 261-278. https://doi.org/10.1111/j.1467-8535.2005.00525.x
18. Giordano, V. A. (2007). A Professional Development Model to Promote Internet Integration into P-12 Teachers' Practice: A Mixed Methods Study. Computers in the Schools, 24(3-4), 111-123. https://doi.org/10.1300/J025v24n03_08
19. Gorder, L. M. (2008). A Study of Teacher Perceptions of Instructional Technology Integration in the Classroom [Article]. Delta Pi Epsilon Journal, 50(2), 63-76. https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=eue&AN=508059595&site=ehost-live&scope=site&custid=s6390179
20. Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. Research in Nursing & Health, 20(3), 269-274. https://doi.org/https://doi.org/10.1002/(SICI)1098-240X(199706)20:3<269::AID-NUR9>3.0.CO;2-G
21. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate data analysis (7th ed.). Upper Saddle River, New Jersey: Pearson Educational International.
22. Hernández-Ramos, P. (2005). If Not Here, Where? Understanding Teachers' Use Of Technology In Silicon Valley Schools. Journal of Research on Technology in Education, 38(1), 39-64. https://doi.org/10.1080/15391523.2005.10782449
23. Jimoh, K., Opesemowo, O. A., & Faremi, Y. A. (2022). Psychometric Analysis of Senior Secondary School Certificate Examination (SSCE) 2017 NECO English Language Multiple Choice Test Items in Kwara State Using Item Response Theory. Journal of Applied Research and Multidisciplinary Studies, 3(2). https://doi.org/10.32350/jarms.32.01
24. Lawshe, C. H. (1975). A quantitative approach to content validity. Personnel psychology, 28(4), 563-575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x
25. Leung, Y.-Y., Png, M.-E., Conaghan, P., & Tennant, A. (2014). A Systematic Literature Review on the Application of Rasch Analysis in Musculoskeletal Disease — A Special Interest Group Report of OMERACT

11. The Journal of Rheumatology, 41(1), 159-164. https://doi.org/10.3899/jrheum.130814
26. Mamun, M. A., & Ullah, I. (2020). COVID-19 suicides in Pakistan, dying off not COVID-19 fear but poverty? – The forthcoming economic challenges for a developing country. Brain, Behavior, and Immunity, 87. https://doi.org/10.1016/j.bbi.2020.05.028
27. Niederhauser, D. S., & Stoddart, T. (2001). Teachers' instructional perspectives and use of educational software. Teaching and Teacher Education, 17(1), 15-31. https://doi.org/https://doi.org/10.1016/S0742-051X(00)00036-6
28. Ogunjimi, M. O., Ayanwale, M. A., Oladele, J. I., Daramola, D. S., Jimoh, M. I., & Owolabi, H. O. (2021). Simulated evidence of computer adaptive test length: Implications for high stakes assessment in Nigeria.
29. Oladele, J. I. (2021). Computer-Adaptive-Testing Performance for Postgraduate Certification in Education as Innovative Assessment. International Journal of Innovation, Creativity and Change, 15(9), 89-101.
30. Oladele, J. I., Ayanwale, M. A., & Ndlovu, M. (2023). Technology Adoption for STEM Education in Higher Education: Students' Experience from Selected Sub-Saharan African Countries. Pertanika Journal of Science & Technology, 31(1), 237-256. https://doi.org/10.47836/pjst.31.1.15
31. Oladele, J. I., Ayanwale, M. A., & Owolabi, H. O. (2020). Paradigm Shifts in Computer Adaptive Testing in Nigeria in Terms of Simulated Evidence. Journal of Social Sciences, 63(13), 9-20.
32. Oladele, J. I., & Ndlovu, M. (2021). A review of standardised assessment development procedure and algorithms for computer adaptive testing: Applications and relevance for fourth industrial revolution. International Journal of Learning, Teaching and Educational Research, 20(5), 1-17.
33. Oladele, J. I., Ndlovu, M., & Ayanwale, M. A. (2022). Computer Adaptive-Based Learning and Assessment for Enhancing STEM Education in Africa: A Fourth Industrial Revolution Possibility. In B. Chirinda, K. Luneta, & A. Uworwabayeho (Eds.), Mathematics Education in Africa: The Fourth Industrial Revolution (pp. 131-144). Springer International Publishing. https://doi.org/10.1007/978-3-031-13927-7_8
34. Onyibe, C., Nwachi-Ikpor, J., & Abdulhakim, A. (2015). Computer based testing technique in Nigeria: Prospects and challenges. Journal of Information Engineering and Applications, 5(10), 17-21.
35. Opesemowo, O. A., Obanisola, A., & Oluwatimilehin, T. R. (2022). From brick-and-mortar to online teaching during the COVID-19 pandemic lockdown in Osun state, Nigeria. Journal of Education in Black Sea Region, 8(1), 134-142. https://doi.org/10.31578/jebs.v8i1.286
36. Opesemowo, O. A., Ayanwale, M. A., Opesemowo, T. R., & Afolabi, E. R. I. (2023). Differential Bundle Functioning of National Examinations Council Mathematics Test Items: An Exploratory Structural Equation Modelling Approach. Journal of Measurement and Evaluation in Education and Psychology, 14(1), 1-18. https://doi.org/10.21031/epod.1142713
37. Reckase, M. D. (2010). Designing item pools to optimize the functioning of a computerized adaptive test. Psychological Test and Assessment Modeling, 52(2), 127-141.
38. Redecker, C., & Johannessen, Ø. (2013). Changing Assessment — Towards a New Assessment Paradigm Using ICT. European Journal of Education, 48(1), 79-96. https://doi.org/https://doi.org/10.1111/ejed.12018
39. Russell, M., Bebell, D., O'Dwyer, L., & O'Connor, K. (2003). Examining Teacher Technology Use:Implications for Preservice and Inservice Teacher Preparation. Journal of Teacher Education, 54(4), 297-310. https://doi.org/10.1177/0022487103255985
40. Stone, E., & Davey, T. (2011). Computer-adaptive testing for students with disabilities: A review of the literature. ETS Research Report Series, 2011(2), 1-24. https://doi.org/10.1002/j.2333-8504.2011.tb02268.x
41. Thompson, G. (2017). Computer adaptive testing, big data and algorithmic approaches to education. British Journal of Sociology of Education, 38(6), 827-840. https://doi.org/10.1080/01425692.2016.1158640
42. Thompson, N. A., & Weiss, D. A. (2011). A framework for the development of computerized adaptive tests. Practical Assessment, Research, and Evaluation, 16(1), 1 – 9. https://doi.org/10.7275/wqzt-9427
43. Veldkamp, B.P., & van der Linden, W.J. (2010). Designing item pools for adaptive testing. In van der Linden, W.J. & Glas, C.A.W. (Eds.) (2010). Elements of adaptive testing, (Statistics for Social and Behavioral Sciences Series). New York: Springer.
44. Wong, E. M. L., & Li, S. C. (2008). Framing ICT implementation in a context of educational change: a multilevel analysis. School Effectiveness and School Improvement, 19(1), 99-120. https://doi.org/10.1080/09243450801896809
45. Xinhu. (2020). Covid-19 UN Chief calls for unity of Security Council. The East African June 16, 2020. [https://www.theeastafrican.co.ke/scienceandhealth/Africa-virus-cases-pass-240000/3073694-5577250-tqrar/index.html].
46. Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A.-R. (2015). Design and Implementation Content Validity Study: Development of an instrument for measuring Patient-Centered Communication. J Caring Sci, 4(2), 165-178. https://doi.org/10.15171/jcs.2015.017