

Item Analysis of the Science and Technology Components of the 2019 Basic Education Certificate Examination Conducted by National Examinations Council

Maxwell E. Uduafemhe^{1,*}, Danladi Uwelo², Sunday O. John³, Roseline Y. Karfe¹

¹Department of Industrial and Technology Education, Federal University of Technology Minna, Niger State, Nigeria

²Department of Educational Foundations, Faculty of Education, Nasarawa State University, Keffi, Nigeria

³National Open University of Nigeria, Minna Study Centre, Minna, Niger State, Nigeria

Received December 28, 2020; Revised February 22, 2021; Accepted March 12, 2021

Cite This Paper in the following Citation Styles

(a): [1] Maxwell E. Uduafemhe, Danladi Uwelo, Sunday O. John, Roseline Y. Karfe, "Item Analysis of the Science and Technology Components of the 2019 Basic Education Certificate Examination Conducted by National Examinations Council," *Universal Journal of Educational Research*, Vol. 9, No. 4, pp. 862 - 869, 2021. DOI: 10.13189/ujer.2021.090420.

(b): Maxwell E. Uduafemhe, Danladi Uwelo, Sunday O. John, Roseline Y. Karfe (2021). *Item Analysis of the Science and Technology Components of the 2019 Basic Education Certificate Examination Conducted by National Examinations Council*. *Universal Journal of Educational Research*, 9(4), 862 - 869. DOI: 10.13189/ujer.2021.090420.

Copyright©2021 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract One of the activities that take place behind the scenes as part of the steps taken by examination bodies to standardized tests is the determination of the psychometric properties of the items in a process referred to as item analysis. This study analysed the adequacy of the basic science and technology test items used for the 2019 Basic Education Certificate Examination in terms of difficulty levels, discriminating powers, guessing, and distracter indices. The study adopted evaluation and descriptive survey research designs. A sample of 976 students was selected from 59 schools across Niger State for the study using a multistage random sampling procedure. A 60-item Basic Science and Technology test administered by the National Examinations Council for the 2019 BECE was used to collect data for the study. The item difficulty levels, discriminating powers, guessing, and distracter indices of the items were generated based on a four-parameter Item Response Theory model using Xcalibre software package. The items adequately satisfied unidimensionality and local independence, when statistically examined using confirmatory factor analysis and tetra-chronic correlation respectively. The findings revealed that all the items satisfied the difficulty, discrimination, distracter and

guessing indices criteria to be used for the 2019 BECE. It was recommended among others that school proprietors and the government should endeavour to sponsor teachers to attend item generation workshops to be up-to-date with how to generate test items that meet the criteria for final examinations.

Keywords Item Analysis, Basic Education Certificate Examination, Difficulty Levels, Discriminating Powers, Guessing, Distracter Indices

1. Introduction

Basic Education Certificate Examination (BECE) occupies a prominent position in the educational life of the Nigerian child. This is not farfetched as it is the first public certification examination taken by students in Nigeria [1]. This examination, in some circles referred to as the Junior School Certificate Examination (JSCE) [1, 2], is taken by students who have completed their first three years of secondary education. According to [3], junior

secondary education is the education that a child receives immediately after primary education. The objectives of this level of education as spelt out in the National Policy on Education (NPE), are to: develop patriotic young people equipped to contribute to social development and the performance of their civic responsibilities; provide children with diverse basic knowledge and skills for entrepreneurship and educational advancement; inspire national consciousness and harmonious co-existence irrespective of differences in endowment, religion, colour, ethnic, and socio-economic background; inculcate values and raise morally upright individuals capable of independent thinking, and who appreciate the dignity of labour [3].

In pursuit of these objectives, the Junior Secondary Education curriculum was designed to provide children with training through the following subjects: English Studies, one Nigerian Language (one of Hausa, Igbo, Yoruba, Edo, Ibibio, and Efik), Mathematics, Religion and National Values, Basic Science and Technology, Pre-Vocational Studies, French, Cultural and Creative Arts, Business Studies, and Arabic. Among these 10 subjects, Basic Science and Technology (BST) is of interest to this study. This is because BST, as hinted by [4], is the only subject at BECE level that avails students the capacity to: Explain events in nature; develop in an all-round manner covering cognitive, affective and psychomotor skills; possess the ability of reasoning and thinking logically; evolve global skills in which scientific know-how can utilise information gathering from nature to solve the problems of mankind, and stir instinctive inquisitiveness during experiments. This is in addition to the fact that science and technology are tools that springboard nations to economic development, self-sufficiency, and sustainability. Hence, the nation being mindful of the importance of science and technology and their contribution towards technological advancement, intentionally included BST as one of the major and compulsory subjects in the junior school system. In testing students during the BECE examination, which is a condition needed for a student to be certified to have completed junior secondary education, the National Examinations Council administers 120 multiple-choice questions (MCQs) covering four main components of BST, namely; basic science, physical and health education, basic technology and computer studies each comprising of 30 items respectively.

The National Examinations Council (NECO) was established in April 1999 by the then Head of State, General Abdulsalami Abubakar. Its core mandate includes: the general control of the conduct of the internal and external Senior Secondary School Certificate Examinations in Nigeria without prejudice to the existing powers and functions of the West African Examination Council; conducting a Standard National Assessment of Educational Performance at junior and senior secondary

school levels; conducting research leading to national improvement of testing and examination procedures at junior and senior secondary school levels [5]. In effect, NECO is the only national agency saddled with the responsibility of conducting a national examination used to certify students after their junior secondary education, that is accepted worldwide. For the tests conducted by NECO to meet international standards, a lot of activities take place behind the scene. One of such activities is the determination of the psychometric properties of the tests used in a process referred to as the item analysis.

Item analysis is the process in which the validity, reliability, discrimination, difficulty, and distracter indices of individual items of a test are examined as they affect the whole test. Item analysis was defined by [6] as a procedure which provides a test developer with the opportunity to examine both test item and the answers supplied by the testees to find out whether they are of adequate quality and if the content coverage is sufficient for the test they are developed for. Item analysis is done on multiple-choice questions of a test after it has been administered [7]. The upside to the practice of item analysis is that it allows for the observation of the item characteristics and the improvement of the quality of the test. Reference [8] revealed that item analysis helps the test developer to examine each specific item to appraise their quality concerning difficulty, discrimination, and distracter indices.

Item difficulty may be viewed as a measure of the fraction of test-takers that correctly answered an item. Similarly, item discrimination can be seen as the ability of an item to distinguish between the performance of brighter testees and those in the group that may be considered as not so bright. In the same vein, [8] defined item discrimination as a measure that differentiates between the performance of students in the high score group and those in the low score group. Reference [9] explained that item difficulty index (DIF I) also commonly referred to as p-value, is a measure of the fraction of the whole test-takers that correctly responded to an item. Also, [10] stated that the distracter index is a measure of how well the incorrect options in an item prevent the 'not-so-bright' test-takers from picking the right option (also called the key). Good distracters appeal to a higher proportion of low-achieving test-takers than they do to the bright ones, in that way they result in negative statistics [11].

It has been noted that on the part of teachers and test developers, carrying out item analysis has the capacity to conserve valuable time and energy in the sense that it functions as a means of identifying items that are too easy or too difficult, distinguishing between testees that were exposed to the content and test-taker who were not exposed to the content or distracters that are not plausible [12]. Sources such as [13], [14] and [6], revealed that there are two widely used approaches to item analysis. They are Classical Test Theory (CTT) and Item Response

Theory (IRT). Classical Test Theory uses two main statistics: the item difficulty index (the proportion of testees that responded correctly to the item) and the discrimination index (the point-performance on individual items and total test scores). IRT however, refers to two key things, both item statistics and testee's capability with the postulation about the existence of a relationship between his score on a single item and eventual overall performance in the test [6].

In NECO, IRT is used to analyse the items of the various tests she conducts. To do this, a pilot study of the items that are intended for use in future examinations is done through a process known as trial testing. This enables the examination body to take a critical look at individual items of the test with the view of making sure that only those of them that meet expected standards are used. However, there is a gap created by the lack of information from an independent observation standpoint, concerning the quality of the items used for testing candidates in Nigeria. The quality of the items used by NECO over the years has only been attested to by the reports of NECO herself and she is not under any legal obligation to make an outright declaration about the quality of the items she uses in conducting her examinations. It is therefore hardly surprising that researchers who hold the opinion that the quality of items is the cause of the unstable nature of the performance of students in key subjects like BST in BECE, assert dominance in the narrative of students' failure to meet expected success levels in public examinations in the country. This is not farfetched as researchers like [15] have asserted that there is the probability of the existence of items that could be said to possess elements of technical inadequacies in the tests developed and administered by examination bodies like NECO. Similarly, [10] are of the view that sometimes, students' failure in key subjects at BECE is because of the faults inherent in the psychometric properties of the test, not just due to their inabilities. Against this backdrop, it is therefore not out of place for an independent item analysis to be done with a view of ascertaining whether or not the items used in public examinations by NECO are of the expected quality. Hence the need for this study.

To this end, this study was undertaken to do item analysis of the BST test administered by NECO in the 2019 BECE. Specifically, the study explored; the adequacy of the difficulty and discrimination indices, level of the guessing factors, and the effectiveness of the distracters of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination. The findings of the study provided information, from an independent viewpoint, on the quality of items used by NECO in conducting her examinations.

2. Theoretical Framework

The theoretical framework of the study is hinged on the 4 parameters logistic model of item response theory (IRT), as developed by Barton and Lord in 1981 [16]. According to [17], the 4 parameters logistic model (4PLM) makes for a more detailed technique for testees' ability estimation than 1, 2, and 3 parameter logistic models. 4PLM estimates four parameters; discriminating power (a-parameter), difficulty level (b-parameter), guessing factor (c-parameter), and distraction factor (d-parameter). The model generally runs under four basic assumptions of monotonicity, unidimensionality, local independence, and invariance [18], but unidimensionality and local independence are the most important. Unidimensionality indicates that items in a test are developed to measure only one area of knowledge. Local independence on the other hand means that the performance of a testee will not affect his score for better or for worse in any other item.

In the 4-parameter logistic model, the probability of a correct response to a dichotomous item is given by:

$$P_{4PLM}(\theta) = c + (d - c) \frac{1}{1 + e^{-Da(\theta - b)}}$$

Where:

θ : is the test taker's ability.

c: is the guessing index.

d: distracter index.

e: is the base of the natural logarithm and is approximately equal to 2.714.

D: is the arbitrary constant (normally $D = 1.7$).

a: is the item discrimination index.

b: is the difficulty parameter [17].

IRT predicts test-takers' chances of providing the right response concerning their abilities and characteristics of the items. Reference [15] revealed that ability level is estimated on a scale that can easily be transformed. The scale is such that it ranges from negative infinity through a mid-point of zero to positive infinity. Nevertheless, in real life situations -3 to +3 is the range of values that ability is often limited to. It follows that items with difficulty levels of +3 and -3 are labelled as "very difficult" and "Very easy" respectively. To carry out item analysis using IRT, the items are dichotomously scored. That is, '1' is allocated for correct responses and '0' for wrong responses. One of the merits IRT has over CTT is that it provides a framework for evaluating how well individual items in a test or examination function. Hence, as noted by [19], some of the relevance and advantages of using IRT in item analysis is that it enables test designers to generate test items, have a well-maintained item bank, and hold equality in terms of item difficulty for alternate forms of tests. This can allow for results obtained overtime to undergo a comparative analysis.

3. Objectives of the Study

The study was guided by the following objectives: To determine the

1. Adequacy of the difficulty indices of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination.
2. Adequacy of the discrimination indices of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination.
3. Adequacy of the guessing factors of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination.
4. Effectiveness of the distracters of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination.

4. Methodology

The study adopted evaluation and descriptive survey research designs. According to [10], evaluation research design has to do with an organised process of passing value judgments based on the set goals, to ascertain whether the objectives should be maintained, modified, or improved on. Descriptive research design is the type of research in which data are gathered from a population to find out how and why they are affected by certain conditions, situations, or phenomena. The population of the study comprised all junior secondary school students in Niger State who were preparing to sit for BECE either that of Niger State Ministry of Education or that of NECO.

A sample of 976 students was selected from 59 schools across the state for the study using a multistage random sampling procedure. The instrument used for the study was a 60-item multiple choice test consisting of the science and technology components of the Basic Science and Technology test administered by NECO in the 2019 BECE. The breakdown of the 60 items of the instrument showing the nine specific subject areas covered is presented in table 1. Each item of the instrument had five options lettered A – E. The testees were expected to answer each question within one minute, hence the time allotted was one hour. The instrument was not subjected to moderation and validation. This was because NECO being an internationally recognized and organized examination agency develops, moderates, validates, and standardizes her items before using them for the examinations she conducts. Hence, the items of the instrument were assumed to be of adequate quality. Copies of the instrument were administered to the sampled students under similar conditions as given by the NECO. The data collected were analysed using Xcalibre statistical analysis software package. The item difficulty levels, discriminating powers, guessing, and distracter indices of the items were generated based on a

four-parameter IRT model from the software package. Two important conditions; unidimensionality and local independence were statistically examined. Confirmatory factor analysis was used to test for unidimensionality by examining all the items used in the study to determine whether or not there is a dormant factor among them. Similarly, tetra–chronic correlation was used to check the local independence between items with the same ability. All the items were found to have met the conditions for unidimensionality and local independence.

Table 1. An overview of specific subject composition of the instrument

Component	Subject Area	No. of Items	Total
Basic Science	Chemistry	5	30
	Biology	20	
	Physics	5	
Basic Technology	Automechanics	5	30
	Building Construction	5	
	Electrical and Electronics	5	
	Metalwork	5	
	Technical Drawing	5	
	Woodwork	5	
	Total		

5. Results

After the data were collected and analysed, the results obtained from the four research questions were summarised and presented in tables 2 – 5 respectively.

Research Question One

How adequate are the difficulty indices of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination?

Table 2. Summary of the Difficulty Indices of the Basic Science and Technology Test Items used for the 2019 Basic Education Certificate Examination

Range of Difficulty Indices	Number of Items	Percentage (%) of Items	Remark
0.30 to 0.39	1	1.67	Easy
0.40 to 0.49	29	48.33	Adequate
0.50 to 0.59	30	50.00	Difficult

Table 2 shows a summary of the difficulty indices of the basic science and technology test items used for the 2019 Basic Education Certificate Examination. The table revealed that of the 60 basic science and technology test items, only one item (1.67%) had a difficulty index below 0.40 meaning that it was easy. Similarly, 29 (48.33%) of the items had difficulty indices that were between 0.40

and 0.49 and so fell into the adequate threshold. Also, 30 items of the test were difficult because they had difficulty indices that fell between 0.50 and 0.59. In a nutshell, all 60 items of the test had adequate difficulty indices.

Research Question Two

How adequate are the discrimination indices of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination?

Table 3. Summary of the Discrimination Indices of the Basic Science and Technology Test Items used for the 2019 Basic Education Certificate Examination

Range of Discrimination Indices	Number of Items	Percentage (%) of Items	Remark
0.30 to 0.39	4	6.67	Slightly low discrimination
0.40 to 0.49	20	33.33	Adequate discrimination
0.50 to 0.59	28	46.66	High discrimination
0.60 to 0.69	8	13.33	Very high discrimination

The summary of the discrimination indices of the basic science and technology test items used for the 2019 Basic Education Certificate Examination are shown in table 3. The results show that four items representing 6.67% had slightly low discrimination with their discrimination indices falling below 0.4. In the same vein, 33.33% (20 items) of the items discriminated adequately with their discrimination indices falling between 0.40 and 0.49. Also, 28 (46.66%) items had high discrimination powers based on having indices that ranged between 0.50 and 0.59. Similarly, 13.33% (8) of the items had discrimination indices that were between 0.60 and 0.69. This means that their discrimination was very high. Essentially, all the 60 items effectively discriminated between bright candidates and the not so bright ones.

Research Question Three

How high are the guessing factors of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination?

Table 4. Summary of the Guessing Factors of the Basic Science and Technology Test Items used for the 2019 Basic Education Certificate Examination

Range of Guessing Factors	Number of Items	Percentage (%) of Items	Remark
0.00 to 0.09	1	1.66	Extremely low guessing.
0.10 to 0.15	22	36.67	Very low guessing.
0.16 to 0.20	37	61.67	Very low guessing.

The summary of the guessing factors of the basic science and technology test items used for the 2019 Basic Education Certificate Examination is shown in table 4. The results show that 1.66% of the items had a guessing factor less than 0.10, making the probability of guessing the correct response to it extremely low. In the same vein, the results showed that the probability of guessing the correct responses to 22 (36.67%) of the items which had guessing factors ranging between 0.10 and 0.15 was very low. Similarly, the correct response to 37 (61.67%) of the items cannot easily be guessed because their guessing factors were between 0.16 and 0.20. In a nutshell, since none of the 60 items had a guessing factor greater than 0.26, they are all good items.

Research Question Four

How effective are the distracters of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination?

Table 5. Summary of the Distracter Indices of the Basic Science and Technology Test Items used for the 2019 Basic Education Certificate Examination

Range of Distracter Indices	Number of Items	Percentage of Items	Remark
0.05 and less	59	98.33	High distraction.
Above 0.05	1	1.67	Low distraction.

Table 5 shows the summary of the distracter indices of the basic science and technology test items used for the 2019 Basic Education Certificate Examination. The result revealed that only one item with a distracter index greater than 0.05 which means that the alternatives did not sufficiently appeal to low achievers among the testees. However, the remaining 59 (98.33%) items had distracter indices of 0.05 or less. This means that the alternatives or distracters of the 59 items appealed sufficiently to low achievers than they do to high achievers. In essence, the distracters of the items were effective.

6. Summary of Findings

The results of the analysis produced the following findings:

1. The difficulty indices of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination are adequate.
2. The discrimination indices of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination are adequate.
3. The guessing factors of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination are adequate.

4. The distracters of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination were effective.

7. Discussion of Findings

The information in table 2 provided the answer to research question one. Findings revealed that the difficulty indices of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination are adequate. This implies that although NECO does not disclose the psychometric properties of the items she uses for the tests she conducts, it does not in any way suggest that the quality of the items used may be of low quality. These findings are consistent with the findings earlier made by [15] who revealed that the items used by NECO are of adequate quality for use in the tests she conducts since their psychometric properties; difficulty, discrimination, distracter indices as well as guessing factors were adequate. Also lending a voice to the adequacy of the NECO's items, [20] and [21] reported that the difficulty parameter of test items constructed by public examining bodies in Nigeria, NECO inclusive, were adequate. This means that even though there is room for improvement, the items constructed by NECO are basically of the right standard for the candidates that sat for the examination.

On a similar note, the information in table 3 provided the answer to research question two. The findings revealed that the discrimination indices of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination are adequate. This implies that NECO as an examination body spares no resources in ensuring that standards are met and maintained in her test items. Similar findings were made by [13] and [21] who reported that the difficulty parameter of test items constructed by public examining bodies in Nigeria, NECO inclusive, were adequate. Hence, accepting the adequacy of the quality of the items, it can be concluded that the quality of the items NECO uses for BECE helps in ensuring that quality and credible grades are awarded to test takers.

The answer to research question three was provided by the information in table 4. The finding revealed that the guessing factors of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination are adequate. Once again this is indicative of the fact that the items used by NECO adequately screen higher achievers from low achievers. The findings of [22] in a study designed to assess the items prone to guessing in SSCE economics multiple-choice tests among students in Kwara State, Nigeria, lend a voice to this when it reported that the items used by NECO are of adequate quality for use in the tests she conducts. [15] also supports this finding when they reported that the items used by

NECO are of adequate quality for use in the tests she conducts since their psychometric properties; difficulty, discrimination, distracter indices as well as guessing factors were adequate. Guessing has a way of negatively impacting on the credibility of test scores. This is because low achieving test-takers could sometimes score a set of items correctly by merely guessing their responses [23]. A good test is normally designed to reduce guessing to the barest minimum just like NECO appeared to have done in the test used for this study.

The data in table 5 provided the answer to research question four. The finding revealed that the distracters of the Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination were effective. This result is not farfetched as researchers like [15], [20] and [21] who previously studied the psychometric properties of tests constructed by the examination bodies in Nigeria, had reported that most of the items used by NECO for instance, had distracters that functioned effectively. This has some implications in education. Seeing that the educational system is strengthened by the quality of testing. Hence, much still needs to be done in the area of improving candidates' performance in key subjects like science and technology. Hence, the abysmal performance sometimes recorded among candidates in public examinations is not in the quality of tests as insinuated by [10]. Perhaps, other factors like students attitude, low morale on the part of teachers, teacher efficacy [24] and the poor state of school infrastructure may have to be looked into.

8. Conclusions

Based on the findings of the study, it was concluded that the basic science and technology items constructed by NECO for the Basic Education Certificate Examination have adequate difficulty, discrimination, distracter and guessing indices. Their standards are adequate for use in a national examination like the Basic Education Certificate Examination. The unstable nature of the performance of candidates in the examination could be due to some other factors such as students' attitude, low morale on the part of teachers and the poor state of school infrastructure.

Recommendations

The following recommendations emerged:

1. School teachers at all levels of the secondary school should imbibe the strategy of incorporating national examination bodies' like NECO's past questions into internal evaluation processes to familiarize their students with the standards they will be facing in their final examination.
2. School proprietors and the government should endeavour to expose teachers to marking

coordination exercises organized by examination bodies during marking exercises to enable them to be acquainted with what examiners expect from candidates in national examinations.

- School proprietors and the government should endeavour to sponsor teachers to attend item generation workshops to be up-to-date with how to generate test items that meet the criteria for final examinations.

Acknowledgements

The authors are grateful to the National Examinations Council (NECO) Nigeria, whose Basic Science and Technology test items used for the 2019 Basic Education Certificate Examination were adapted for this study.

REFERENCES

- Osuafor A. M., Okoli J. N. Challenges encountered by non-science teachers in teaching basic science and technology in the Nigerian universal basic education (UBE) curriculum, *African Journal of Teacher Education*, Vol. 3, No.3, pp. 1-8, 2013.
- Orluwene G. W., Igwe B. N. The influence of educational systems on the academic performance of JSCE students in Rivers State, *Journal of Education and Practice*, Vol. 6, No. 6, pp. 85-89, 2015.
- Federal Republic of Nigeria. National policy on education, NERDC Press, Lagos, 2013.
- Osokoya M. M. Teaching methodology in basic science and technology classes in South-West Nigeria, *Asian Journal of Education and e-Learning*, Vol. 1, No. 4, pp. 206-214, 2013.
- National Examinations Council (NECO) (Establishment) Act, 2002.
- Siri A., Freddano M. The use of item analysis for the improvement of objective examinations, *Procedia-Social and Behavioral Sciences*, Vol. 29, No. 1, pp. 188-197, 2011.
- Hassan S., Hod R. Use of item analysis to improve the quality of single best answer multiple choice question in summative assessment of undergraduate medical students in Malaysia, *Education in Medicine Journal*, Vol. 9, No. 3, pp. 33-43, 2017.
- Sabri S. Item analysis of student comprehensive test for research in teaching beginner string ensemble using model based teaching among music students in public universities, *International Journal of Education and Research*, Vol. 1, No. 12, pp. 1-14, 2013.
- Musa A., Shaheen S., Elmardi A., Ahmed A. Item difficulty & item discrimination as quality indicators of physiology MCQ examinations at the Faculty of Medicine Khartoum University, *Khartoum Medical Journal*, Vol. 11, No. 2, pp. 1477-1486, 2018.
- J. Joe-Kinane, B. Nkem, G. W. Orluwene. A comparative study on item characteristics of 2014 – 2016 mathematics objective tests in junior secondary school certificate examination questions in Rivers State, *International Journal of Mathematics Trends and Technology*, Vol. 52, No. 8, pp. 545-551, 2017.
- Chauhan P., Chauhan G. R., Chauhan B. R., Vaza J, V., Rathod S. P. Relationship between difficulty index and distracter effectiveness in single best-answer stem type multiple choice questions, *International Journal of Anatomy and Research*, Vol. 3, No. 4, pp. 1607-1610, 2015. DOI: <http://dx.doi.org/10.16965/ijar.2015.299>.
- Quaigrain K., Arhin A. K. Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation, *Cogent Education*, Vol. 4, No. 1, pp. 1-11, 2017. <http://dx.doi.org/10.1080/2331186X.2017.1301013>.
- Bichi A. A., Embong R., Talib R., Salleh S., Ibrahim A. Comparative analysis of classical test theory and item response theory using chemistry test data, *International Journal of Engineering and Advanced Technology*, Vol. 8, No. 5C, pp. 1260-1266, 2019. DOI: 10.35940/ijeat.E1179.0585C19.
- Janssen G., Meier V., Trace J. Classical test theory and item response theory: Two understandings of one high-stakes performance exam, *Colombian Applied Linguistics Journal*, Vol. 16, No. 2, pp. 167-184, 2014. DOI: <http://dx.doi.org/10.14483/udistrital.jour.calj.2014.2.a03>.
- Aborisade O. J., Fajobi O. O. Comparative analysis of psychometric properties of mathematics items constructed by WAEC and NECO in Nigeria using item response theory approach, *Educational Research and Reviews*, Vol. 15, Vol. 1, pp. 1-7, 2020. DOI: 10.5897/ERR2019.3850.
- Barton M. A., Lord F. M. An upper asymptote for the three-parameter logistic item-response model. *Research Bulletin 81-20*, Princeton Educational Testing Service, New Jersey, 1981.
- Liao W. W., Ho R. G., Yen Y. C., Cheng H. C. The four-parameter logistic item response theory model as a robust method of estimating ability despite aberrant responses, *Social Behavior and Personality: an international journal*, Vol. 40, No. 10, 1679-1694, 2012.
- Columbia University. Item response theory, 2019. Retrieved on 11th April 2020 from. <https://www.mailman.columbia.edu/research/population-health-methods/item-response-theory>.
- Bortolotti S. L. V., Tezza R., de Andrade D. F., Bornaia A. C., de Sousa Júnior A. F. Relevance and advantages of using the item response theory, *Quality & Quantity*, Vol. 47, No. 4, pp. 2341-2360, 2013. DOI 10.1007/s11135-012-9684-5.
- Metibemu M. A. Comparison of classical test theory and item response theory in the development and scoring of senior secondary school physics tests in Ondo State, Unpublished Ph.D thesis, Institute of Education, University of Ibadan, 2016.
- Ogbebor U. C. Construction of mock economics test for senior secondary school students in Delta State, Nigeria using classical test and item response theories. Unpublished

Ph.D Thesis, Institute of Education, University of Ibadan, 2017.

- [22] Jimoh M. I., Daramola D. S., Oladele J. I., Sheu A. L. Assessment of items prone to guessing in ssce economics multiple-choice tests among students in Kwara State, Nigeria, *Anatolian Journal of Education*, Vol. 5, No. 1, pp. 17-28, 2020. <https://doi.org/10.29333/aje.2020.512a>.
- [23] Cevik E. E., Kurnaz M. A. Analysis of the Responses of Science Teacher Candidates to Force Concept Inventory by Concentration Factor. *Universal Journal of Educational Research* vol 7, No 1, pp. 111-117, 2019. DOI: 10.13189/ujer.2019.070115.
- [24] Valente S., Veiga-Branco A., Rebelo H., Lourenço A. A., Cristóvão A. M. The relationship between emotional intelligence ability and teacher efficacy. *Universal Journal of Educational Research*, Vol. 3, No. 8, pp. 916-923. 2020. DOI: 10.13189/ujer.2020.080324.