Dimensionality and Model Fit Assessment of National Examination Council of Nigeria 2020 Biology in North Central Nigeria

Hamman, Mathew Joseph[1](PhD) josephhamman@yahoo.com +2348036798962
Nigeria Army College of Education
Omale Onuh[2] (PhD) omaleonuh@gmail.com +2348103587655
J.S.Tarka University Makurdi Benue State
and
Gloria Adaku Dike[3] +234837125165 gloria.dike@unn.edu.ng **University of Nigeria Nsukka**

Abstract
The primary objective of test development is to create a high-quality test by carefully selecting appropriate items through item analysis and ensuring that the test adheres to the assumptions of Item Response Theory (IRT). In this study, two research questions guided the research process, and an Expost-facto research design was employed. The research was conducted in North-Central Nigeria. The study population consisted of 18,252 Senior Secondary School Student III (SSS3) who registered for and took the NECO Biology Examinations in the year 2020. For this study, a sample of 1,825 students was selected from the total population of 18,252, representing 10% of the population. Data collected for analysis were processed using NORHAM for research question one, and model fit statistics was utilized for research question two. The research findings indicated that the NECO 2020 Biology examination did not meet the assumption of unidimensionality. However, the data could be effectively fitted into a 2-parameter logistic model. In summary, the study's main objective was to create a high-quality test by choosing suitable items through item analysis and ensuring that the test conformed to IRT assumptions. The research revealed that the NECO 2020 Biology examination did not satisfy the assumption of unidimensionality but could be adequately fitted into a 2-parameter logistic model.

Keywords: Assessment, Dimensionality, Model Fit, NECO, Biology

Introduction

Science and technology have always been recognized as critical factors in the process of development. Through its application, the resources of nations have been transformed into goods and services all over the world. Abdulkadir (2011) remarks that the current development in science and technology has greatly affected the lives of every human being, such that, to be ignorant of the basic knowledge of this development is to live an empty, meaningless, and probably unrealistic life. For any nation to attain the status of self-reliance, science must be an important component of that nation, irrespective of race, creed, or sex. Science can be defined as the systematic body of knowledge obtained by methods or techniques based on observation, and experimentation as its authority. It seeks to explain the natural phenomenon using inquiry processes or activities. The branches of science are: Biology, Physics, and Chemistry, among others, and can be broadly classified into natural and applied sciences.

Biology is an important aspect of the study of science as it deals with the study of life and living organisms with their environments (Satyaprakashai & Yaspal, 2014). It entails the study of the structure, function, growth, origin, evolution, distribution, and taxonomy of living things, as well as provides basic knowledge to an individual about the human body and that of other living organisms in the environment (FME,2012). In most Nigerian secondary schools, Biology is one of the major subjects students take at internal and external examinations. Biology, as an integral part of natural science is needed for our nation's technological breakthrough. According to Nsofor (2010), Biology covers a wide scope, serves as a springboard for many future careers in science and technology, and has applications, nearly in every field of life. For a nation to be considered a developed society in the technological age, the study of Biology is the brain behind its growth. It is important, since it forms a veritable armour against misconceptions and superstitions which muddle technological advancement anywhere. It is a natural science that deals with the living world: how the world is structured, how it functions and what these functions are. How it develops, how living things came into existence, and how they react to one another and their environments (Umar. 2011). Biology is a pre-requisite subject for many fields of study that contribute immensely to the technological growth of nations. This includes Medicine, Pharmacy, Nursing, Agriculture, Forestry, Biotechnology, Nanotechnology. and many other areas (Ahmed & Abimbola. 2011). As a scientific discipline, it is a fundamental subject in the development of any nation, including Nigeria. Its development potential is evident in Biomedical sciences, Agricultural sciences, Textile, Paper and Dyeing Technology, Biotechnology, Environmental sciences, and socio-cultural issues (TutorVista. 2010; Mork. 2011).

The quality and quantity of science education received by secondary school students are geared toward developing future scientists, technologists, engineers, and related professionals. Despite the importance of Biology for career development of Nigerian students, the achievement of students in Biology at senior secondary school level has been consistent on average over the years as reported by the NECO chief examiner, from 2018-2022. Achievement is a measure of success exhibited by senior secondary school students, in terms of the scores obtained on the test given to them to ascertain how much of the knowledge is learned at the end of an instruction. Scores on any test are unequally precise measures for examinees of different abilities; thus making the assumption of equal errors of measurement for all examinees implausible. Since individual ability is not measured, rather it is performing the test that has been measured by the examinee, which gives rise to questions of measurement theory used by the examination body in developing the test.

Test is one of the most important parameters with which society adjudges the product of her education system. Test has always been an important part of the school system that even the habitual absentees normally turn up in school and present themselves to be tested on examination days. The essence of the test is to reveal the latent ability of the examinee and to make grounds for assessment across the country to be as uniform as possible, which may be lacking in many measuring instruments (Olabode and Adeleke 2015). One of the primary purposes of tests in our educational system is to provide a means of measuring or evaluating a group of examinees' abilities and skills that are as fair and objective as possible. The test has been fully accepted in

most modern societies as the most objective method of decision-making in schools, industries and government establishments.

Though test result is accepted to be used in most societies as one of the most objective methods of decision-making, nonetheless, the use of test has sparked some concerns among the members of the public in recent years. These concerns have tended to erode people's faith in the power and efficacy of tests. Most of the serious allegations levied against tests pivot around social issues, that tests may show culture, location, gender, school type or class bias (Anastasia and Urbina, 2006). A test must not be biased against a segment of the population taking the test items. In many cases, test items are biased because they contain sources of difficulty that are irrelevant to the construct being measured and these irrelevant factors affect the performance of the examinees. The question is 'what could have caused this problem?' This gives rise to the present study.

An aspect which appeared not to have been investigated with detailed attention by researchers in Biology Education is the effect of the scoring procedure being adopted by public examination bodies in Nigeria on students' performance. According to Oguoma, Metibemu and Okoye (2016), Public examination bodies in Nigeria use Classical Test Theory for the scoring of students' performance.

Classical Test Theory assumes that each person has a true score (T) that would be obtained when there were no errors in measurement. A person's true score is defined as the expected number of correct scores over an infinite number of independent administrations of the test. Unfortunately, test users never observe a person's true score, but only an observed score (x). It is assumed that the observed score = true score plus or minus some error. with the advent of Item Response Theory (IRT), students' performance in external examinations could be assessed using the IRT Model.

The IRT model is based on the assumption that the items are measuring a single trait, ranging from negative to positive infinity. IRT assumes that all items in a test must be developed to measure one and only one trait. Secondly, the IRT model assumes local independence of item responses. Thus, the performance across items in the same instrument should not be related, except as a result of the influence of the trait level that they all are designed to measure (Nenty, 2004). The assumptions of unidimensionality and item local Independence are technically related, in that items' local dependence implies a separate dimension in factor analysis. These assumptions are very useful for the determination of test dimensionality. It is therefore important that the dimensionality and item local independence of Chemistry be assessed; and using a unidimensional model in the scoring of the test, which is inherently a multidimensional test on examinees' performance. De Ayala (2009) states that test dimensionality refers to the number of traits or constructs assessed by the items in the test. A test assessing only one trait is termed unidimensional, while a test assessing more than one trait of examinees is referred to as a multidimensional test (DeMars, 2010). An item is considered unidimensional if the systematic differences within the items' variance are only due to one variance source, that is, one latent variable. Unidimensionality assumes that, only one dominant latent trait is being measured and that this trait is the driving force for the responses observed for each item in the measure. (Aduloju,2019). It has been reported by Liu (2015) that some large-scale tests, such as NECO are

nearly unidimensional for the constructs that are measured; in either case, theoretical models and estimation programs are available for calibrating items. The crucial benefits of IRT models are realized to the degree that the data fit the different models, 1-, 2-, and 3 parameters. Model-data fit is a major concern when applying Item Response Theory (IRT) models to real test data. Though there is an argument that the evaluation of fit in IRT modelling has been challenging, the use of IRT model checking and item fit statistics serve as crucial factors to effective IRT use in psychometrics for information on items and model selections. Obtaining evidence of model-data-fit when an IRT model is used to make inferences from a data set is recommended as the standard for educational and psychological testing by the American Association of Educational Research, American Psychological Association, and National Council on Measurement in Education (2014). Failure to meet this requirement invalidates the application of IRT in real data set evaluation. It is on this basis that the researcher seeks to assess the dimensionality and model data fit in NECO 2020 Biology paper.

Statement of Problem

The basic aim of test development is to construct a test of desired quality by choosing the appropriate items through item analysis and ensuring their reliability and validity. In developing quality test items to effectively measure students' achievement, it is pertinent that the best practices in test construction be employed by NECO. Researchers over the years have pointed out that, some best practices in an item and test analysis are too infrequently used by most examinations, NECO inclusive. It is expected that examinations, such as NECO should be valid, reliable, and all other psychometric properties that made up tests are ensured. Since it is presumed that the examination body uses the IRT model to calibrate and standardise their instrument, one would expect that the item is unidimensional and is fitted into the right IRT model that will reflect the true abilities of the test takers. This way, the researcher intends to determine the Assumption of Unidimensionality of the NECO Biology items for 2020 and determine the data model fit of NECO 2020 Biology test items

Research questions

1. Do items of the NECO Biology test of 2020 fulfill the assumption of unidimesionality?
2. Which of the IRT model data do NECO 2020 Biology test items fit?

Item Response Theory (IRT)

The item response theory (IRT) was pioneered by Fredrick, Rasch, and Larzarfeld in the 1950s and 1960s. Item Response Theory is a general statistical theory about examinees, item and test performance, and how performance relates to the abilities that are measured by the items in the test. It is a theory of testing based on the relationship between individual performances on a test item and the test takers' levels of performance on an overall measure of the ability that the item was designed to measure. IRT is a collection of mathematical models and statistical methods

that are used to analyze items, administer psychological measures, and measure individuals on psychological constructs.

Assumptions of IRT

The IRT model is based on the assumption that the items are measuring a single continuous latent variable θ ranging from negative to positive infinity. IRT assumes that all items in a test must be developed to measure one and only one trait. Secondly, the IRT model assumes local independence of item responses. Thus, the performance across items in the same instrument should not be related except as a result of the influence of the trait level that they all are designed to measure. The assumptions of unidimensionality and local independence are technically related, in that an item's local dependence implies a separate dimension in factor analysis. Three fundamentals of IRT are: Item Response Functions (IRFs), Information Functions (IF), and Item Invariance (II).

Essen, Idaka, and Metibemu (2017) investigate item-level diagnostic statistics and model-data fit with the one- and two-parameter models using IRTPROV3.0 and BILOG- MG V3.0. The findings revealed that only 1 item fitted 1- 1-parameter model in BILOG- MG V3.0 and IRTPRO V3.0. 26 items fitted 2-parameter models when using BILOG-MG V3.0. Five items fitted 2-parameter models in IRTPRO. While Umobong and Udeme (2017) examine the dimensionality of the NECO Biology examinations based on the Rasch Measurement Model. The findings of the study show that the majority of the items on the examinations had item fit statistics and factor structures that met the recommended values, and that the standard errors of measurement were in line with the recommendation of the Rasch Model. Evidence derived from the analysis of the dimensionality of Biology examinations showed that most of the items measured unidimensional constructs and; thus, were valid and reliable.

Chikezie (2017) assesses the unidimensionality of the West African Senior Certificate Examination (WASSCE) in Chemistry. The findings revealed that the test violated the assumption of unidimensionality and the chi-square goodness of fit test showed that 94% of the items were statistically significant, and did not fit the IRT 3- 3-parameter model. Also, Taiwo and Afolabi (2014) assess unidimensionality and occurrence of Differential Item Functioning (DIF) in Mathematics and English Language items of the Osun State Qualifying Examination (OSQ). The results show that OSQ Mathematics ($-0.094 \leq r \leq 0.236$) and English Language items ($-0.095 \leq r \leq 0.228$) were unidimensional.

Methodology

The study adopts an Expost-facto design. The study was carried out in the North Central of Nigeria. The population for the study comprises 18,252 Senior Secondary School Student III (SSS3) who registered and sat for the NECO Biology examinations in the years 2020, over two years. The sample for this study consists of 1,825 students out of the 18,252 who registered and sat ‘for the NECO 2020 Biology examination. The sample size was arrived at by taking 10% of the population. This is following Borg and Gall (cited in Emaikwu, 2015) who stated that, for a population that is

up to 5,000 and above, 10% of the population is large enough to be considered a representation of the population. Data collected were analyzed using NORHAM for research question one, model fit statistics built in IRT Pro for research question two

Results and Discussion

Research Question 1: do items of the NECO Biology test of 2020 fulfil the assumption of unidimensionality?

Table 1**: Dimensionality assessment of NECO Biology 2020**

Dimension	GFI	RMSR Criterion	RMSR	Diff. in RMSR	Reduction in RMSR	% Reduction
1	0.94091	0.10454	0.00954			
2	0.95499	0.10454	0.00833	0.00121	0.126834	12.68
3	0.96419	0.10454	0.00743	0.00090	0.108043	10.80
4	0.96922	0.10454	0.00689	0.00054	0.072678	7.27

The Table depicts the number of dimensions underlying Multiple-choice Biology achievement Test Table shows that 1-dimension, 2-dimension, 3-dimension and 4-dimension models fitted the data (RMSR for 1-dimension, 0.00954 was less than the criterion 0.10454, a similar trend was observed for 2-dimension, 3-dimension and 4-dimension respectively and GFI for the four dimensions hypothesized to underlie the test were greater than 0.90). To ascertain the optimal dimensions embedded in the test data, the fitness of the data to the four hypothesized dimensions models are compared. The table depicts that from one dimension to the two-dimensions, the RMSR value decreased by 12.68%. According to Tate criteria, this is a significant amount of reduction, showing that two -dimensions significantly fitted the data better than one dimension. Moreover, when the 3-dimension was hypothesized to underlie the data set, the percentage of reduction in RMSR between the two-dimension and three-dimension model was approximately 10.80%, showing that three- dimensions significantly fitted the data better than the two- dimensions.

More importantly, when the fitness of four dimensions was compared to that of three- dimensions, the reduction in RMSR value was less than the criterion of 10%. The highest dimensional model that still produced an approximately 10% or greater percentage reduction in the RMSR over the preceding model was the three-dimensional model. Therefore, the 3-dimension was the most parsimonious model that fits the test well. This implies that the performance of examinees on the Biology achievement test substantially depended on three dimensions of abilities or traits. The implication is that multidimensional is evident in the test data.

Research Question 2: Which of the IRT model data do NECO 2020 Biology test items fit?

Table 2: Calibration Analysis of Biology

Item	S-X^2	Df	P–Value	Remarks	Item	S-X^2	df	P–Value	Remarks
1	63.370	33	0.001	**Misfit**	31	49.216	32	0.027	**Misfit**
2	30.347	32	0.550	Fit	32	25.994	33	0.802	Fit
3	135.755	32	0.000	**Misfit**	33	57.893	33	0.005	**Misfit**
4	44.651	33	0.060	Fit	34	26.246	33	0.607	Fit

5	41.535	32	0.121	Fit	35	84.868	33	0.000	**Misfit**
6	25.744	33	0.812	Fit	36	26.017	33	0.617	Fit
7	40.350	33	0.051	Fit	37	68.430	33	0.000	**Misfit**
8	39.514	33	0.202	Fit	38	97.070	33	0.000	**Misfit**
9	42.843	33	0.117	Fit	39	27.852	32	0.743	Fit
10	85.654	32	0.000	**Misfit**	40	44.807	32	0.066	Fit
11	108.945	33	0.000	**Misfit**	41	44.515	33	0.064	Fit
12	36.833	32	0.301	Fit	42	27.507	33	0.737	Fit
13	38.886	33	0.222	Fit	43	31.550	33	0.539	Fit
14	44.478	32	0.070	Fit	44	42.056	33	0.134	Fit
15	43.171	32	0.142	Fit	45	24.919	33	0.843	Fit
16	42.951	33	0.115	Fit	46	40.036	33	0.178	Fit
17	27.222	33	0.750	Fit	47	44.540	33	0.087	Fit
18	40.013	33	0.187	Fit	48	65.770	32	0.000	**Misfit**
19	35.256	33	0.362	Fit	49	33.718	32	0.384	Fit
20	41.328	32	0.125	Fit	50	54.221	33	0.011	**Misfit**
21	23.035	33	0.902	Fit	51	44.165	33	0.071	Fit
22	34.311	32	0.381	Fit	52	44.939	33	0.080	Fit
23	36.255	33	0.319	Fit	53	33.421	33	0.447	Fit
24	290.954	33	0.000	**Misfit**	54	32.363	33	0.449	Fit
25	37.873	33	0.236	Fit	55	43.552	33	0.104	Fit
26	41.326	32	0.125	Fit	56	37.879	32	0.219	Fit
27	42.643	31	0.080	Fit	57	31.605	32	0.486	Fit
28	30.769	32	0.529	Fit	58	34.443	32	0.352	Fit
29	30.904	31	0.471	Fit	59	29.426	31	0.546	Fit
30	37.629	32	0.227	Fit	60	27.360	31	0.654	Fit

Key: $s\text{-}x^2$ = Chi-square statistic
df = degree of freedom

Table 2 shows that out of the 60 items, 48 items representing 76% fitted the 3-PLM. The table also revealed that the remaining 12 items representing 24% were statistically significant and did not fit the 3-PLM.

Discussion of Findings

The study's findings indicate that the Biology Achievement Tests consist of multidimensional items, suggesting that these items measure multiple dominant traits in test-takers. This result contradicts the findings of Umobong and Udeme (2017), who concluded that Biology examination items were unidimensional constructs and thus valid and reliable. However, it aligns with the research conducted by Chikezie (2017), which also found that the test violated the assumption of unidimensionality. Chikezie's study further revealed that 94% of the items did not fit the IRT 3-parameter model according to the chi-square goodness-of-fit test.

In contrast, the results for OSQ Mathematics and English Language items showed unidimensionality (i.e., they measured a single trait). In the analysis of research question 2, it was found that 76% of the test items were statistically insignificant and therefore fitted the 3-PLM. This finding is consistent with the findings of Agah (2015) and Ayanwale et al. (2018), who reported similar percentages of items fitting the 3-PLM in their respective studies. It also aligns with the results of Eleje and Esomonu (2018) and Atsua et al. (2018), who used 2loglikelihood values to establish model fit and found the 3-PLM to be the most appropriate model for their test items. Furthermore, the finding agrees with Osarumwase's (2019) study, which reveals that the NABTEB May/June 2017 Biology test items fit the 3-PLM.

However, the findings of this study contradict those of Okwu and Iweka (2018), Bichi et al. (2016), and Chikezie (2017), whose researches indicate that the items in their test instruments fit the 1- and 2-PLM, respectively. The present study's results may be attributed to the fact that most items in the two tests are unidimensional and thus fit the 3-PLM. Additionally, the use of large sample sizes for item parameter estimation may have contributed to significant item fit statistics in the present study.

Conclusion/Recommendation

Based on the results of this study, it was concluded that the test items are multi-dimensional, fitted to 3-PLM and can be utilized in comparing students' latent abilities for sound educational decisions in our schools. It was recommended that examination bodies, researchers that wish to use IRT in solving measurement problems, especially those involving tests and scales should make efforts to conform to the IRT assumption.

Reference

Agah, J. (2015). Relative Efficiency of Test Scores Equating Methods in Comparison of Students' Continuous Assessment Measures. A Ph.D. Thesis. Department of Science Education, University of Nigeria, Nsukkka, Nigeria.

Abdulkadir, A. A. (2011). Youths' Interest in Business and Technical Skill Acquisition. It's Implication for National Development. *Journal of Science Technology and Mathematics Education, 7(3), 238-244.*

Aduloju, M.O. (2020), Assessment of Unidimensionality and Local Independence of PHY111 Test of University of Agricultural Makurdi using IRT Framework. *Nigerian Journal of Educational Research and Evaluation Vol 19, 124-137.*

Ahmed, M. A., & Abimbola, I. O. (2011). Influence of Teaching Experience and School Location on Biology Teachers' Ratings of the Difficulty Levels of Nutrition Concepts in Ilorin, Nigeria. *Journal of Science, Technology, Mathematics and Education (JOSTMED), 7(2), 52-61.*

Anastasia, A, & Urbina, S. (2006). Psychological testing. New Delhi: Prentice Hall

Atsua, T.G, Uzoeshi, V.I & Oludi, P. (2018). Equating 2015 and 2016 Basic Education Certificate Examination on Civic Education using Classical Test Theory and Item Response Theory in Oyo state, Nigeria. *Journal of Pristine, 14*(1), 2250-9593.

Ayanwale, M.A. Adeleke, J.O. & Mamadelo, T.I. (2018). An Assessment of Item Statistics Estimates of Basic Education Certificate Examination through Classical Test Theory and

Item Response Theory Approach. *International Journal of Educational Research Review, 3(4),* 55 – 67.

Bichi, A.A., Hatiz, H. & Bello, S.A. (2016). Evaluation of North-West University, Kano Post-UTME Test Items Using Item Response Theory. *International Journal of Evaluation and Research in Education (IJERE), 5(4),* 261 – 270.

Chikezie, I. J.(2017). Assessment of Unidimensionality of West African Senior School Certificate Examination in Chemistry with Principal Component Analysis and Item Response Theory Model. *African Journal of Theory and Practice of Educational Assessment.* vol.5, November 2017, 47-57

Cyrinus B. E., Idaka E. I., and Michael A. M.(2017) Item level diagnostics and model-data fit in item response theory (IRT) using BILOG - MG v3.0 and IRTPRO v3.0 programmes GLOBAL JOURNAL OF EDUCATIONAL RESEARCH VOL 16, 87-94 DOI*: http://dx.doi.org/10.4314/gjedr.v16i2.2*

De Ayala R.J. (2009) The Theory and Practice of Item Response Theory. Psychometrika. 75. 778-779. 10.1007/s11336-010-9179-z.

Demars,C.(2010). Item Response Theory. Oxford University Press,2010. Scientific Research An Academic Publisher *.https'//doi.org/10.1093/prof:oso/9780195377033.001.0001*

Essen C. B., Idaka E. I & Metibemu M. A. (2017) Item level diagnostics and model - data fit in item response theory (IRT) using BILOG - MG v3.0 and IRTPRO v3.0 programmes https://www.ajol.info/index.php/gjedr/article/view/162434

Eleje, L.I. & Esomonu, N.P. (2018). Test of Achievement in Quantitative Economics for Secondary Schools: Construction and Validation Using Item Response Theory, *Asian Journal of Education and Training 4(1),* 18 – 28.

Emaikwu, S.O. (2011). Fundamentals of Test, Measurement and Evaluation with Psychometric Theories. 2nd Edition, Makurdi. SAP Ltd (selfers Academic Press). 78 – 85.

Federal Republic of Nigeria (2012). National Policy on Education, Abuja: NERDC Press.

Liu, Y. (2015). Interval estimation of latent variable scores in item response theory. Journal of Educational and Behavioral Statistics. https://doi.org/10.3102/1076998617732764.

Lord, F. M., & Novick, M. R. (1968).Statistical theories of mental test scores (with contributions by A. Birnbaum). Reading, MA Addison-Wesley.

Lord, F.M. (1952). A theory of test scores. Psychometric Monograph, 7. Lord. F. M. (1971).A theoretical study of 151-166. two-stage testing. Psychometrika,36, 227-242.

Lord, F.N., Novick, M.R., Statistical theories of mental test scores, Reading: MA, Addison-Wesley, 1968.

Nenty, H.J.(2004). From classical test theory(CTT) to Item response theory (IRT): An introduction to a desirable transition. In A.O. Afemikhe& J.G Adewale, Issues in Educational Measurement and Evaluation in Nigeria. Educational Research and Study Group, Institute of Education, University of Ibadan, Ibadan Nigeria.

Nsofor, C .C. (2010). Effects of Improvised Instructional Media on Secondary School Students' Achievement in Biology Concepts in NigerState. *An Unpublished Ph.D Thesis, Department of Science Education, Federal University of Technology, Minna.*

Osarumwense, H.J. & Duru, C.P. (2019). Assessment of Model-fit for 2016 and 2017 Biology Multiple Choice Test Items of the National Business and Technical Examination Board. *International Journal for Innovation Education and Research, 7(4),* 12 – 22.

Oku, K., & Iweka, F. (2018). Development, Standardization and Application of Chemistry Achievement Test using One-Parameter Logistic Model (1-PLM) of Item Response Theory (IRT). *American Journal of Educational Research, 6*(3), 1-58. Retrieved from. http://pubs.sciepublicon education/6/3/11/index.html#cor on 05/09/18.

Oguoma, C.,Metibemu, M., & Okoye, R.(2016). An assessment of the dimensionality of the 2014 West Africa Secondary School Examination mathematics objective test in Imo State, Nigeria. African Journal of Theory and Practice of Educational Assessment.

Taiwo,O.A. &Afolabi, A.E.(2014). Assessing Unidimensionality and Differential Item Functioning in Qualifying Examination for Senior Secondary Schools students, Osun State, Nigeria. World Journal of Education. ISSN (Print) 1925-0746(online) Sciedu press. *Doi.https//doi.org/10.5430/wje.v4n4p30*

Umobong, E. & Udeme, E. (2017). The dimensionality of National Examinations Council's Biology Examinations: Assessing Test Quality in Modern Trend Approach. African Journal of Theory and Practice of Educational Assessment.

Umar, A. A. (2011). Effects of Biology Practical Activities on Students' Process Skill Acquisition in Minna, Niger State, Nigeria. *JOSTMED,* 7(2), 118-126.