International Journal of Education, Learning and Development

Vol.4, No.10, pp.25-35, December 2016

Published by European Centre for Research Training and Development UK (www.eajournals.org)

DETERMINATION OF DIFFERENTIAL ITEM FUNCTIONING BY GENDER IN THE NATIONAL BUSINESS AND TECHNICAL EXAMINATIONS BOARD (NABTEB) 2015 MATHEMATICS MULTIPLE CHOICE EXAMINATION

Kingsley Osamede, OMOROGIUWA (Ph. D)
Department of Educational Evaluation and Counseling Psychology,
Faculty of Education, University of Benin, Nigeria.

Edeki Philomena, IRO-AGHEDO
Psychometric Unit
National Business and Technical Examinations Board (NABTEB)
National Headquarters, Ikpoba Hill, Benin City, Nigeria.

ABSTRACT: *The purpose of the study was to investigate the differential item functioning (DIF) by gender in National Business and Technical Examinations Board (NABTEB) 2015 Mathematics Multiple Choice Test Items (Dichotomous) Examination in Nigeria. This was conducted by determining the items that functioned differentially by male and female examinees. A survey research design was employed. A sample of 17,815 examinee responses was selected from two states each from two geo-political zones out of 63,584 examinee responses from the six geo-political zones in Nigeria. This comprised 11,873 male and 5,942 female examinees. A 50-item multiple choice Mathematics test item was used to gather data. To detect the items that functioned differentially by male and female examinees, Area Index (Raju) method which is one of the item response theory methods of DIF detection was applied. The results of the analysis revealed that male and female examinees functioned differentially in seventeen items (34%) and no difference in 33 items (66%). Out of the seventeen items, six items were in favour of male students while 11 items were in favour of the female students. Based on the result of the findings, it was then recommended among others that for bias-free items to be produced, examination bodies, test experts and developers should make certain that activities and connotations reflected in the test are relevant to the construct being measured and explore the use of Area Index method of DIF to detect the items that function differentially by gender.*

KEYWORDS: Differential Item Functioning, Area Index (Raju) method, Multiple Choice Item, Bias Item, NABTEB

INTRODUCTION

In education, test is crucial in determining students' academic achievement. The test could be used for promotion, certification, recruitments, and placement and so on depending on the purpose using a valid and reliable item as measuring instrument. The qualities of measuring instruments lie mainly on the quality of items used in the instruments. There is need to ensure that items are not only valid and reliable but also fair to all across the subgroup of examinees (male and female). Indeed, the Federal Republic of Nigeria in the National Policy on Education (2014) stated that national examinations tests should be valid and fair to all students to which the test is set to measure the attributes that is needed from them. Test fairness is a crucial issue

ISSN 2054 - 6297(Print), ISSN 2054 - 6300(Online)

Vol.4, No.10, pp.25-35, December 2016

Published by European Centre for Research Training and Development UK (www.eajournals.org)

in testing and it reflects the same constructs for all examinees and scores have the same meaning for all individuals in the intended population.

An important step in the construction of assessment instruments is to ensure that no individual or group responding to the instrument is disadvantaged in any way (Kanjee, 2007). For instance, in an achievement test, students of equal ability are usually drawn from the same population but belonging to different subgroups such as male or female, should have the same probability of getting an item correct. This can only be hindered when the item is biased. Bias test items are those that differentially inhibit individuals from showing their true abilities and thereby measuring irrelevant construct. Such items are said to be displaying differential item functioning (DIF) which according to Reynolds (2006) systematically underestimates or overestimates the value of the variable the items are designed to measure. DIF exists in a test item when, despite controls for overall test performance, examinees from different groups have a different probability of getting an item correct or when students from two sub-populations with the same ability level have different expected scores on the same item (Penfield & Camilli, 2007).

Indeed, DIF occurs when examinees from different groups have different likelihoods of success on an item, after they have been matched on the ability of interest (Clauser & Mazor, 2008). The presence of DIF is as a result of some characteristics in an item that result in differential performance for individuals of equal ability but from different group. Items may be judged relatively more or less difficult for a particular group by comparison with the performance of another group drawn from the same population. Differential item functioning of an item can therefore be understood as a lack of conditional independence between an item response and group membership (often gender, location or ethnicity) given the same latent ability or trait (Taiwo & Eyitayo, 2014)

It is crucial to match groups, since the comparison should establish a distinction between differences in item responses from divergences between two groups. For example, in a Mathematics test which needs calculation ability and English-reading comprehension, consider examinees with the same calculation ability. However, one group is more competent in English-reading comprehension than the other group. If the two groups show differences in the probability of answering some of the items of the test correctly, due solely to differences in English proficiency, the items can be said to possess DIF. It is essential for test developers or test users to investigate whether items influence examinees' performance in systematically biased ways for some particular subgroups due to any extraneous sources of variance. Thus, if there are DIF items, it means that irrelevant factors which probably have effects on the responses, but are not interested in, are driving the responses beyond the latent variable that is purportedly measured (Ackerman, 2002). If some items function unfavorably over specific groups, the explanations made from the test cannot be thought of as valid and fair.

One of the ways to investigate bias items at the item level is through DIF analysis. A DIF analysis is a means of statistically identifying unexpected differences in performance across matched groups of examinees. It compares the performance of matched majority (or reference) and minority (or focal) group examinees. There are several methods of detecting DIF. Some of these methods are based on Classical Test Theory (CTT) which include Mantel-Haenszel (M-H), Logistic Regression (LR) and Simultaneous Item Bias (SIBTEST). Other methods such as

ISSN 2054 - 6297(Print), ISSN 2054 - 6300(Online)

Lord's chi square test, Raju's area measures and IRT-Likelihood Ratio (IRT-LR) are examples of DIF detection methods based on Item Response Theory (IRT). Most of these methods provide similar but not identical information about DIF. This study focuses on the Raju's Area method. The Raju's area measure is based on quantifying the gap between item characteristics curves functions. According to Oshima and Morris (2008), the approach provides an intuitive and flexible methodology for assessing differential item functioning.

Statement of the Problem

Recurrent poor academic performances and achievement difference in Mathematics across Nigerian secondary, technical and business schools seem to be unabated. Some researchers attribute the poor performance of students in Mathematics to factors such as teachers, environment, parents, and facilities. However, there seems to be a persistent better performance of male groups over female groups in Mathematics test items. Ogbebor and Onuka (2013) found that male students perform better in Mathematics than their female counterparts. Could these differences in performance be as a result of the nature of test items used that make one gender to perform better than the other? One would wonder if test items are fair to all test takers or if some items are in favour of one group over the other. Thus, the researcher deems it necessary to investigate if the multiple choice Mathematics test items administered by NABTEB function differentially by gender Raju method of Item Response Theory based approach. NABTEB is one of the certification bodies in Nigeria responsible for conduct of examination leading to the award of the National Technical Certificate (NTC), Advanced National Technical Certificate (ANTC), National Business Certificate (NBC), Advanced National Business Certificate (ANBC).

Concept of Area Index for Two-Parameter Logistic Model (Raju's Area Method): Area index for two-parameter logistic model is used to measure the area between the two item characteristic curves (ICCs) of the reference and the focal groups as an index of the difference between the performances of the two groups matched on ability. The larger the area, the larger the difference between the two curves (Abedlazeez, 2010). An item is said to possess differential item functioning when the area index is greater than a critical value of 0.22, while an item does not possess differential item functioning when the area index is zero or close to zero (De Beer, 2004). Also, according to Ling and Lau (2003), when the b parameter (item difficulty) for one group (for example, Male) is greater than the other group (for example, Female), this shows that the item is more difficult for the male group and the item is said to favour the other group (that is, female), and vice versa.

Raju formula for area index between two curves is as follows:

$$Area = 2\frac{(a_2 - a_1)}{Da_1 a_2} L_n\left[1 + e^{Da_1 a_2 \frac{b_2 - b_1}{a_2 - a_1}}\right] - (b_2 - b_1)$$

Where: a_1: discrimination parameter for males (reference group), a_2: discrimination parameter for females (focal group), b_1: difficulty parameter for males (reference group), b_2: difficulty parameter for females (focal group), D = 1.7 (constant: scaling factor)

ISSN 2054 - 6297(Print), ISSN 2054 - 6300(Online)

Vol.4, No.10, pp.25-35, December 2016

Published by European Centre for Research Training and Development UK (www.eajournals.org)

Objectives of the Study

The objective of this study is to determine the differential item functioning of the 2015 NABTEB Mathematics test items in Nigeria. Specifically, this study is aimed to find out if there is a difference in the number of items functioning differentially by gender in the 2015 NABTEB multiple choice Mathematics test. To carry out this study, the following research questions were posed:

1. What percentage of items in the 2015 NABTEB multiple choice Mathematics examination function differentially by gender?
2. Is there any difference in the number of items functioning differentially by gender in favour of males and those in favour of females in the 2015 NABTEB multiple choice Mathematics examination?

Question two was hypothesized.

Research Design

The research design adopted for this study was the survey research design. This design is considered appropriate because only a part of the population was studied and findings from this were used to generalize for the entire population.

Population of the Study

The population of the study comprised 63,584 candidates that enrolled and sat for the National Business and Technical Examinations Board (NABTEB) 2015 May/June Mathematics multiple choice examination in the six geo-political zones in Nigeria.

Sample and Sampling Technique

The total number of sampled candidates used in the study was 17,815 students. This comprised 11,873 male and 5,942 female students. Multistage sampling technique was employed for effective selection of the sample in the study. At the first stage, Stratified sampling technique was used to stratify the states in Nigeria into six (6) geo-political zones which are North-West, North-Central, North-East, South-West and South-East South-South Zones. At the second stage, simple random sampling technique was used to select two (2) zones from the six geo-political. This was gotten from the 33% of the total number of the geo-political zones in Nigeria (that is 6). The two zones that were selected are North-Central and South-South zone. At the third stage, there are seven states that make- up the North-Central and six states that make-up the South-South zone. Simple random sampling technique was used to select two (2) states (Niger State and Kwara State) from North-Central zone and two (2) states (Edo State and Akwa Ibom from South-South which constitutes 33% of the total number of the states in the selected zones.

Research Instrument

The instrument that was used to gather data was a 50 -item multiple choice Mathematics test from National Business and Technical Examinations Board 2015 May/June examination.

Validity of the Instrument

The 50 multiple choice Mathematics test items was validated and standardized by the Examinations Development Department, NABTEB. Thus, the items are considered appropriate in terms of subject contents and instructional objectives.

ISSN 2054 - 6297(Print), ISSN 2054 - 6300(Online)

Reliability of the Instrument

Being an instrument of a standardized national examination which was conducted by NABTEB, the instrument was deemed reliable. Hence, the reliability of the instrument was not established by the researcher.

METHOD OF DATA COLLECTED

The researcher collected the candidates' responses and key on the Mathematics test items from NABTEB.

Method of Data Analysis

Principal Component Analysis was performed to test for uni-dimensionality of the multiple choice Mathematics test items. Item parameters were estimated using the computer program Xcalibre Version 4.2.0.1 (Assessment Systems Corporation, 2013). Raju Area Measure method was used to determine the presence of differential item functioning. An item is said to possess differential item functioning when the area index is greater than a critical value of 0.22, while an item does not possess differential item functioning when the area index is zero or close to zero (De Beer, 2004). Also, according to Ling and Lau (2003), when the b parameter (item difficulty) for one group (for example, Male) is greater than the other group (for example, Female), this shows that the item is more difficult for the male group and the item is said to favour the other group (that is, female), and vice versa. The hypothesis was tested using chi-square statistic at 0.05 alpha level of significant.

Presentation of Results

Unidimensionality of the test items was considered before analyzing DIF. This was done because undimensonality is an assumption of item response theory (IRT). The method used in this study for assessing the undimensonality was principal component analysis which was done on the dichotomous items using a sample size of 17,815 students. The examinees' performance in the Mathematics examination was accounted for by a single latent trait/ ability due to the dominating factor. The scree plot showed the unidimensionality of the items.

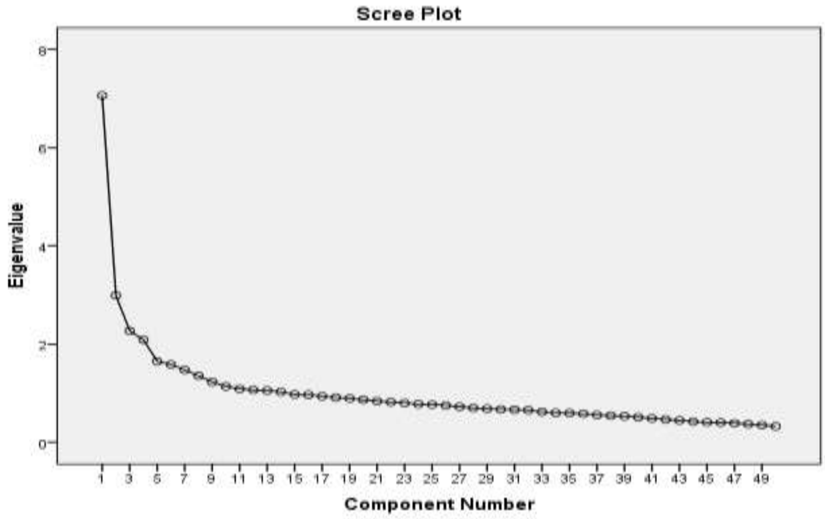

Figure1: Scree plot of eigen value (2015 NABTEB Mathematics)

ISSN 2054 - 6297(Print), ISSN 2054 - 6300(Online)

Vol.4, No.10, pp.25-35, December 2016

Published by European Centre for Research Training and Development UK (www.eajournals.org)

Table 1: Summary of Area Index of 2015 NABTEB multiple choice Mathematics

Item	a1(male)	b1(male)	a2(female)	b2(female)	Area Index	Decision	Favoured group
1	0.453	-3.308	0.365	-3.503	0.291	DIF	Female
2	1.859	-0.12	0.682	-0.828	1.060	DIF	Female
3	0.803	-2.227	0.811	-2.183	0.043	NON-DIF	
4	0.876	-2.855	0.828	-2.844	0.025	NON-DIF	
5	0.831	-2.72	0.813	-2.921	0.228	DIF	Female
6	0.55	-2.917	0.71	-2.482	0.229	DIF	Male
7	3.003	-0.382	0.818	-1.548	3.239	DIF	Female
8	0.74	-1.863	0.686	-2.031	0.207	NON-DIF	
9	0.602	-1.602	0.868	-1.217	0.233	DIF	Male
10	0.622	-2.083	0.709	-1.963	0.104	NON-DIF	
11	1.293	-1.008	1.252	-1.094	0.122	NON-DIF	
12	1.147	-0.901	1.063	-1.081	0.297	DIF	Female
13	0.707	-2.23	0.846	-2.03	0.157	NON-DIF	
14	0.831	-2.074	0.837	-2.024	0.049	NON-DIF	
15	3.166	-0.328	0.971	-1.184	3.783	DIF	Female
16	0.804	-1.833	0.793	-1.85	0.017	NON-DIF	
17	1.052	-1.483	1.146	-1.383	0.096	NON-DIF	
18	0.969	-1.423	0.874	-1.639	0.310	DIF	Female
19	1.077	-1.554	0.979	-1.767	0.341	DIF	Female
20	0.619	-2.317	0.657	-2.223	0.082	NON-DIF	
21	1.007	-1.871	1.152	-1.818	0.212	NON-DIF	
22	0.477	-1.46	0.871	-0.774	0.167	NON-DIF	
23	0.752	-1.935	0.839	-1.803	0.112	NON-DIF	
24	0.516	-2.292	0.598	-1.789	0.335	DIF	Male
25	1.167	-1.044	1.065	-1.187	0.183	NON-DIF	
26	0.807	-1.324	0.885	-1.141	0.144	NON-DIF	

ISSN 2054 - 6297(Print), ISSN 2054 - 6300(Online)

27	0.903	-1.48	1.097	-1.407	0.316	DIF	Male
28	1.017	-1.248	1.224	-1.535	0	NON-DIF	
29	1.251	-0.405	0.745	-0.851	0.282	DIF	Female
30	1.468	-0.579	1.13	-0.843	0.071	NON-DIF	
31	0.695	-1.6	0.85	-1.197	0.243	DIF	Male
32	0.926	-1.519	0.976	-1.305	0.177	NON-DIF	
33	1.161	-1.481	1.237	-1.446	0.075	NON-DIF	
34	0.79	-1.397	0.897	-1.325	0.094	NON-DIF	
35	0.93	-1.674	1.33	-1.446	0	NON-DIF	
36	2.394	-0.236	1.451	-0.458	1.561	DIF	Female
37	2.353	-0.694	1.252	-1.21	1.455	DIF	Female
38	1.26	-1.187	1.349	-1.118	0.096	NON-DIF	
39	0.654	-1.447	0.775	-1.158	0.195	NON-DIF	
40	1.088	-1.472	1.031	-1.541	0.074	NON-DIF	
41	1.2	-0.886	1.462	-0.892	0	NON-DIF	
42	0.081	2.667	0.106	2.694	0.020	NON-DIF	
43	1.192	-1.286	1.152	-1.399	0.161	NON-DIF	
44	1.474	-1.15	1.62	-1.176	0	NON-DIF	
45	0.999	-1.295	0.944	-1.353	0.059	NON-DIF	
46	0.527	-0.75	0.618	-0.161	0.384	DIF	Male
47	0.779	-1.568	0.883	-1.34	0.168	NON-DIF	
48	0.925	-1.339	1.013	-1.325	0.064	NON-DIF	
49	1.026	-1.415	1.057	-1.402	0.017	NON-DIF	
50	0.243	2.667	0.298	2.694	0.056	NON-DIF	

Table 1 shows the Summary of Area Index of 2015 NABTEB multiple choice Mathematics. It displays the items that exhibit DIF and the group it favoured. The finding shows that seventeen

ISSN 2054 - 6297(Print), ISSN 2054 - 6300(Online)

Vol.4, No.10, pp.25-35, December 2016

Published by European Centre for Research Training and Development UK (www.eajournals.org)

items out of fifty multiple choice Mathematics items functioned differently by gender. Research question one revealed that Seventeen items representing 34% functioned differently by gender with area indices greater than the critical value of 0.22 and thirty-three items representing 66% do not function differentially with area indices less than 0.22. The findings showed that out of the seventeen items that functioned differentially by gender, six items representing 35.3% were in favour of male students (that is, 11 items representing 64.7% functioned against the male) while 11 items representing 64.7% were in favour of the female students (that is, 6 items representing35.3% functioned against the female students). There is no significant difference in the number of items functioning differentially by gender in favour of males and those in favour of females in the 2015 NABTEB multiple choice Mathematics examination.

Table 2: Chi-square summary of Differential Item Functioning in favour of male and female students

Gender	Item favoured due to DIF	Df	Chi-square	Sig.(2-tailed)
Male	6 (8.5)	1	1.47	0.225
Female	11 (8.5)			
Total	17			

$\alpha = 0.05$

Table 2 shows a chi-square value of 1.47 and a p-value of 0.225, testing the hypothesis at an alpha level of 0.05, the p-value is greater than the alpha level, so the null hypothesis which states that 'there is no significant difference in the number of items functioning differentially by gender in favour of males and those in favour of females in the 2015 NABTEB multiple choice Mathematics examination' is retained.

DISCUSSION OF FINDINGS

The findings of research question one indicated that 17 out of the 50 multiple choice Mathematics items functioned differentially for male and female students. The finding of this study agrees with the findings of Adedoyin (2010), who in his study investigated gender biased items in public examinations, and found that out of 16 items that fitted the 3PL item response theory statistical analysis, 5 items were gender biased. The finding also agreed with that of Adebule (2013) that out of the 40 items examined for the first factor program structure in computer science, only seven items representing 17.5% displayed DIF, comparing male and female examinees. The finding is also in agreement with the report of Birjandi and Mohadeseh (2007) that in the general reading comprehension, 7 out of the 13 DIF flagged items favoured females and 6 proved much easier for males.

The finding from hypothesis revealed that there is no significant difference in the number of items functioning differentially by gender in favour of males and those in favour of females in the 2015 NABTEB multiple choice Mathematics examination. The result of the analysis also showed that the gender DIF items are due to the fact that they contain sources of difficulty that are irrelevant or extraneous to the construct being measured, and these irrelevant factors affect performance. The researcher found that the reason of differential performance is due to the characteristics of the Mathematics test items such as the content of the item and cognitive complexity. Cognitive skills assessed by items seem the most effective factor that produced

ISSN 2054 - 6297(Print), ISSN 2054 - 6300(Online)

gender DIF. The finding is in agreement with the finding of Ling and Lau (2004) who investigated the gender DIF in multiple choice and open- response science item types for elementary, middle and high school levels and found out that the indicative of possible sources of DIF is due to the differences in content category, visual-spatial component and item type dimensions.

The findings did not correlate with the findings of Adebule (2013) who investigated DIF in a 3-20 item multiple choice Mathematics test items selected from Ekiti State Unified Mathematics examination for 2008/2009 and 2009/2010 academic sessions. The study concluded that the items of ESUME did not function differentially among the testees on the basis of gender, age, parental qualification and location. The trend of this study also did not correlate with the findings of Madu (2012) who investigated differential item functioning (DIF) by gender in Mathematics examination conducted by West African Examinations Council (WAEC) in 2011 in Nigeria. Using a sample of 1,671 students and Scheuneuman Modified Chi-square Statistics (SSχ2), the results of the analysis indicated that items significantly function differentially by gender for male and female examinees in 39 items and 11 items did not exhibit DIF.

This study determined the differential item functioning in Mathematics multiple choice test items in terms of gender. Two research questions were raised to guide the study while one hypothesis was tested. The data for the study were collected from the students' responses in the 2015 NABTEB Mathematics examination. Therefore, the data were analyzed using Xcalibre 4.2.0.1 IRT item parameter estimation software, Microsoft Excel, SPSS Version 20, Chi-square statistics and Area index statistic. The result of the analysis revealed that:

- The multiple choice Mathematics test items functioned differentially by gender.
- Out of fifty multiple choice items Mathematics items, seventeen items (34%) functioned differently by gender with area indices greater than the critical value of 0.22 and thirty-three items (66%) do not function differentially with area indices less than 0.22.
- Out of the seventeen items that functioned differentially by gender, six items representing 35.3% were in favour of male students while 11 items (64.7%) were in favour of the female students.

Tthere is no significant difference in the number of items functioning differentially by gender in favour of males and those in favour of females in the multiple choice Mathematics examination

CONCLUSION

DIF is an issue that must be properly addressed in examinations and tests designed for heterogeneous groups. Through the application of IRT methodology (Area Index Measure), it was clear that there were presence of DIF in the 2015 NABTEB Mathematics test items. It is obvious that threat in the validity of test items has been created. Such threats could influence or introduce traits irrelevant to the construct of interest. This could jeopardize classification of subgroup of candidates test scores negatively. It was also concluded that Multiple Choice Mathematics test items administered by NABTEB in 2015 do not show a significant difference in the number of items functioning differentially by gender in favour of males and those in favour of females. Therefore, test developers, ministry of education and examination bodies should ensure that items are free from differential item functioning (DIF).

ISSN 2054 - 6297(Print), ISSN 2054 - 6300(Online)

RECOMMENDATIONS

On the basis of the findings and conclusion, the following recommendations are made:

1. Test experts and developers should consider the use of Area index measure in determining differential item functioning. This approach provides an intuitive and flexible methodology for detecting DIF.
2. Educational measurement experts in Nigeria should rise to the challenges placed by the measurement community and be fully aware of the usefulness of IRT in constructing and scoring of tests or examinations.
3. For bias-free items to be produced, the NABTEB examination developers should make certain that activities and connotations reflected in the test are relevant to the life experiences of examinees responding to the items. Test items should be written in a straight forward, uncomplicated, easily read manner. Excessive wordiness can obviously prevent the examinees from responding appropriately to test items and therefore create bias in the examination.
4. Examination bodies should organize training for item developers on the construction of valid, reliable and fair test especially in the area of DIF. In addition, items flagging DIF should be revised, modified or eliminated from the test.

REFERENCES

Abedlazeez, N. (2010). *Exploring DIF: Comparison of CTT and IRT methods.* International Journal of Sustainable Development, 1(7), 11-46.

Adebule, S.O. (2013). *A study of differential item functioning in Ekiti State unified mathematics examination for senior secondary schools.* Journal of Education and Practice, 4(17), 43-46.

Adedoyin, O. O. (2010). *Investigating the invariance of person parameter estimates based on classical test and item response theories.* International Journal of Education Science, 2(2), 107-113.

Birjandi, P. and Mohadeseh, A. (2007). *Differential item functioning (test bias) analysis paradigm across manifest and latent examinee groups (on the construct validity of IELTS.* Human Sciences, 55, 153-172.

Clauser, B. E. and Mazor, K. M. (2008). *Using statistical procedures to identify differentially functioning test items.* Educational Measurement Issues and Practice, 17, 31-44.

De Beer, M. (2004). *Use of differential item functioning (DIF) analysis for bias analysis in test construction.* South African Journal of Industrial Psychology, 30 (4), 52-58.

Kanjee, A. (2007). *Using logistic regression to detect bias when multiple groups are tested.* South African Journal of Psychology, 37, 47-61.

Ling, S. E and Lau, S. H. (2004). *Detecting differential item functioning (DIF) in standardized multiple-choice test: An application of item response theory (IRT) using three parameter logistic model.* Journal of Applied Psychology, 94 (7), 452-459.

Madu, B. C (2012). *Analysis of gender-related differential item functioning in mathematics multiple choice items administered by West African Examination Council (WAEC).* Journal of Education and Practice, 3(8), 71-79.

Ogbebor, U. and Onuka, A. (2013). *Differential item functioning method as an item bias indicator.* Educational Research, 4 (4), 367-373.

Oshima, T. C. and Morris, S. B. (2008). *Raju's differential functioning of items and tests (DFIT).* The National Council on Measurement and Evaluation Instructional Module, pp. 43-50.

Penfield, R. D. and Camilli, G. (2007). Differential item functioning and item bias. In C. R. Rao and S. Sinharay (Eds.*), Handbook of statistics*: Vol. 26. Psychometrics (pp. 125-167). Amsterdam: Elsevier.

Reynolds, R. C. (2006). *Measurement and Assessment in Education*. Boston: Pearson.

Ajaigbe T. O. and Afolabi, R. I. (2014). *Assessing unidimensionality and differential item functioning in qualifying examination for senior secondary school students, Osun State, Nigeria.* World Journal of Education, 4(4), 30-37.

ISSN 2054 - 6297(Print), ISSN 2054 - 6300(Online)