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

EXAMINING THE ITEM BIAS OF MATHEMATICS EXAMINATIONS CONSTRUCTED BY WAEC AND NECO IN NIGERIA

Prof. Bandele Samuel Oye¹ and Dr. Aborisade Olatunbosun James²

¹Ekiti State University, Ado-Ekiti, Ekiti State, Nigeria ²Faculty of Education, Ekiti State University, Ado-Ekiti, Ekiti State, Nigeria

ABSTRACT: This study examined the item bias of mathematics examination items constructed by West African Examination Council (WAEC) and National Examination Council (NECO) in Nigeria on the basis of probability of success of examinees of comparable ability. The study employed descriptive research design of the survey type. A sample of 600 Senior Secondary School three (SSS III) students was randomly selected from 12 Senior Secondary Schools in Ekiti State, Nigeria using multistage sampling technique. Two Instruments which included 50 mathematics test items randomly selected from past WAEC questions and 50 mathematics test items randomly selected and tested at 0.05 probability level. Two-parameter Logistic Bilog MG software statistical analysis was used to generate Item Response to examine whether the items exhibit item bias or not . The results show that 44 (88%) of NECO items exhibit item bias while 35 (70%) of the WAEC items exhibit item bias. Hence, it was recommended that examination bodies should commit itself to eliminating or reducing biased items in National Examinations.

KEYWORDS: Item Bias, WAEC, NECO, IRT and ICC Curves

INTRODUCTION

In Nigeria, examination and award of ordinary level certificate are majorly done by the West African Examination Council (WAEC) and the National Examination Council (NECO). These examination bodies construct test items on various subjects offered at the secondary school level (ordinary level) which they administer on students for certification. Students that take this examination are supposed to perform without bias to sex, discipline and age and so on. However, candidates who participate in the examinations conducted by these examination bodies are in different settings and therefore differently toned for personal and environmental reasons. As a result of this, the problem of test item bias cannot be ruled out in these examinations. It has been claimed that some of the national examinations unfairly favour examinees of some particular group than the other (Emaikwu, 2012). The National Policy on Education stated that the National Examination tests should be as valid as possible and as fair as possible to all students (FRN, 2014).

A critical look at the perception of people on such national examination in Nigeria indicates the serious nature of item bias. A test item that is not uni-dimensional is of course not free from bias. If the test makes the members of one group look worse than their attainment would actually be on the job or in the classroom, the test is said to be biased against the group. Bias is the presence of some characteristics of an item that results in differential performance for individual comparable ability but from different sex, location, discipline, age religion and so on. Therefore, when important decisions are made based on test scores, it is critical to avoid Published by European Centre for Research Training and Development UK (www.eajournals.org)

bias, which may unfairly influence examinees scores. According to Scheuneman and Bleistein (1994) in Aborisade (2016), it is necessary to apply item bias detection procedure to examination items or tests as part of the evaluation of the overall instrument using IRT approach.

Statement of the Problem

Item bias affects the vital psychometric properties of measurement results in terms of validity and reliability, examination bodies are expected to construct test items in such a manner that items are free from writing errors such as wordiness, irrelevancy, offensiveness, and excessive stimulations, so that when an inadequacy exists between groups examination scores, the discrepancy will be what the test purports to measure in the examinees.

Students that take these examinations are supposed to be of comparable abilities location, age, gender and so on notwithstanding. Bias in testing is appearing in public forum including court of law (Berk, 2007). Emaikwu (2012) reported that the Nigeria Senate in year 2010 summoned the then minister of education to the senate chamber to explain why the massive failure occurred in that year's national examinations; the issue of test item bias and test-wiseness featured prominently among other reasons given for massive failure in some sections of the country.

Examinees that are of comparable ability levels such as a set of students that have exposed to the same course content should have 'equal probability of success' irrespective of the subgroup of the population to which they belong. If the examination items contain sources of difficulty that are not relevant to the construct being measured, these extraneous sources affects examinees performance (Aborisade, 2016).However, some examination bodies do not include item bias detection in their item analysis. Could this be case with the test items constructed by WAEC and NECO? Therefore, analysis of item bias of the items constructed by WAEC and NECO should be done to ascertain the level of validity of the examination items.

Purpose of the Study

The purpose of the study was to find out whether items constructed by WAEC and NECO exhibit item bias. The study, also, identified items that are bias on the basis of the probability of success of examinees from comparable ability level using IRT approach.

Research Hypotheses

The following hypotheses were formulated for the study and tested at 0.05 probability level.

- Ho₁: Examination items constructed by WAEC will not exhibit item bias
- Ho₂: Examination items constructed by NECO will not exhibit item bias

METHODOLOGY

The study employed a descriptive research of the survey type. The population for the study consisted of all the final year students in the public senior secondary school in Ekiti State, Nigeria who had completed the learning process of the curriculums in the syllabi of the two examination bodies. A total of 600 students were selected as sample for the study using

Vol.6, No.2, pp.1-7, May 2018

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

multistage sampling technique. The instrument comprised 50 mathematics test items randomly selected from past WAEC questions and 50 mathematics test items randomly selected from past NECO questions. The Instruments were reviewed and vetted for face and content validity. Experts were asked to examine and check the adequacy of the distribution of the items selected as well as correct the item classification. The instruments were trial tested using 120 students in three secondary schools outsides the sampled schools in Ekiti State, Nigeria. The Kudar Richardson formula 20 (KR₂₀) was used to estimate a reliability coefficient of 0.75 for the WAEC objective test and coefficient of 0.72 for NECO objective test. The instruments were administered to the 600 sampled students under similar conditions as given by the examination bodies.

Data Analysis

Bilog MG Logistic software statistical analysis was used to generate Item Characteristic Curves (ICCs) for each NECO and WAEC examination items to test the hypotheses.

RESULTS

Hypothesis 1

Examination items constructed by WAEC will not exhibit item bias.

Fig I: showing the sample of two- parameter IRT Item Characteristics Curves (ICC) generated for each of the WAEC items as shown for the item one below:



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

Item	a-parameter	b-parameter	Df	Chi-square	Prob.
1	0.436	4.165	5.0	9.93	0.0773
2	0.985	0.948	6.0	25.56	0.0002*
3	1.121	1.516	7.0	131.44	0.0000*
4	0.478	0.567	7.0	42.51	0.0000*
5	0.784	0.824	6.0	8.47	0.2058
6	0.118	1.827	7.0	119.63	0.0000*
7	0.643	0.870	7.0	13.61	0.0587
8	0.401	1.367	8.0	77.54	0.0000*
9	0.533	0.592	7.0	6.58	0.4737
10	0.746	0.972	5.0	1.94	0.8574
11	0.829	1.047	6.0	6.69	0.3508
12	0.248	1.225	7.0	68.31	0.0000*
13	0.447	0.923	7.0	54.10	0.0000*
14	0.726	0.617	7.0	13.74	0.0560
15	1.203	0.837	4.0	26.95	0.0000*
16	0.112	0.211	8.0	141.78	0.0000*
17	0.611	1.905	7.0	15.06	0.0352*
18	0.516	0.952	7.0	31.89	0.0000*
19	0.745	1.067	6.0	14.95	0.0206*
20	0.463	1.860	6.0	44.79	0.0000*
21	0.841	1.029	6.0	16.44	0.0116*
22	0.460	1.987	6.0	18.77	0.0046*
23	0.00	0.000	0.0	0.00	0.0000*
24	0.658	0.797	7.0	13.25	0.0662
25	0.709	1.075	6.0	23.39	0.0007*
26	0.736	1.604	6.0	5.49	0.4828
27	0.517	1.630	7.0	6.69	0.4623
28	0.339	1.348	7.0	46.14	0.0000*
29	0.559	1.796	6.0	5.37	0.4968
30	0.419	0.481	7.0	22.91	0.0018*
31	0.399	1.122	8.0	47.97	0.0000*
32	0.820	0.794	7.0	5.50	0.5989
33	0.466	0.769	6.0	60.28	0.0000*
34	0.877	1.493	5.0	5.40	0.3691
35	0.494	2.801	6.0	21.75	0.0013*
36	1.263	1.248	4.0	22.55	0.0002*
37	0.828	0.895	5.0	17.31	0.0039*
38	0.086	-3.910	7.0	107.48	0.0000*
39	0.505	1.816	7.0	13.07	0.0704
40	0.101	1.620	7.0	265.55	0.0000*
41	0.254	0.582	8.0	44.13	0.0000*
42	0.642	1.141	6.0	36.91	0.0000*
43	0.433	2.487	7.0	28.45	0.0002*
44	0.371	1.994	8.0	73.46	0.0000*
45	0.701	1.496	6.0	2.57	0.8600
46	0.002	644.185	0.0	0.00	0.0000*
47	0.702	1.164	6.0	17.22	0.0085*
48	0.494	1.903	6.0	32.76	0.0000*
49	0.124	4.051	8.0	110.19	0.0000*
1 50	0.708	1.0.855	1.6.0	1 18.69	10.0047*

Table 1: Summary of the ICC Parameter Estimates of WAEC Items

*The items with p ≤ 0.05 exhibit item bias

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

Items with $p \le 0.05$ exhibit item bias while item with p > 0.05 did not exhibit item bias. A cursory look at table 1 above showed that 35 items (70%) of the 50 WAEC examination items exhibit item bias while 15 items (30%) did not exhibit item bias.

Hypothesis 2

Examination items constructed by NECO will not exhibit item bias.

Fig 2: showing the sample of two-parameter IRT Item Characteristics Curves (ICC) generated for each of the NECO items as shown for the item one below:



Table 2: Summary of the ICC Parameter Estimates of NECO Items

Item	a-parameter	b-parameter	Df	Chi-Square	Prob.
1	2.162	-0.159	3.0	26.75	0.0000*
2	0.994	1.291	6.0	7.01	0.3201
3	0.427	1.853	7.0	54.66	0.0000*
4	1.452	0.187	4.0	18.61	0.0009*
5	1.528	0.404	4.0	45.35	0.0000*
6	1.023	0.378	5.0	58.57	0.0000*
7	0.126	2.382	8.0	95.44	0.0000*
8	0.767	1.584	6.0	28.98	0.0001*
9	0.414	0.486	7.0	136.04	0.0000*
10	0.128	5.912	7.0	69.70	0.0000*
11	0.180	4.295	7.0	35.41	0.0000*
12	1.2222	0.283	4.0	42.94	0.0000*
13	0.187	-0.460	8.0	39.49	0.0000*
14	0.219	6.010	6.0	11.88	0.0646
15	0.606	0.111	7.0	48.94	0.0000*
16	0.664	1.852	7.0	19.04	0.0081*
17	0.139	3.129	7.0	90.48	0.0000*
18	0.572	1.886	7.0	49.54	0.0000*
19	0.152	6.917	7.0	53.65	0.0000*
20	0.472	1.812	7.0	22.36	0.0022*
21	0.281	1.799	7.0	66.91	0.0000*

International Journal of Quantitative and Qualitative Research Methods

Vol.6, No.2, pp.1-7, May 2018

			-	-	
22	0.510	1.064	6.0	51.92	0.0000*
23	0.011	48.837	0.0	0.00	0.0000*
24	1.070	1.737	6.0	5.93	0.4311
25	0.962	0.241	5.0	26.48	0.0001*
26	1.091	-0.081	5.0	31.04	0.0000*
27	0.773	1.223	6.0	26.00	0.0002*
28	1.287	1.529	5.0	19.40	0.0016*
29	0.390	1.270	6.0	62.85	0.0000*
30	0.641	1.419	7.0	29.46	0.0000*
31	0.184	0.913	6.0	122.02	0.0000*
32	0.414	1.431	6.0	100.05	0.0000*
33	0.843	1.931	6.0	10.33	0.1114
34	0.113	2.096	7.0	94.17	0.0000*
35	0.450	1.379	7.0	36.56	0.0000*
36	0.005	114.653	0.0	0.00	0.0000*
37	0.464	1.378	6.0	49.07	0.0000*
38	0.564	1.818	6.0	79.36	0.0000*
39	0.099	4.068	7.0	67.91	0.0000*
40	0.799	0.246	6.0	44.48	0.0000*
41	0.625	1.470	7.0	31.72	0.0000*
42	0.148	7.078	7.0	41.88	0.0000*
43	0.323	1.757	7.0	59.24	0.0000*
44	1.159	0.307	4.0	28.76	0.0000*
45	0.185	7.069	6.0	9.13	0.1662
46	0.612	0.240	7.0	49.24	0.0000*
47	0.549	-0.128	7.0	40.38	0.0000*
48	0.267	6.970	4.0	5.45	0.2441
49	0.508	2.395	6.0	25.28	0.0003*
50	0.354	2.703	7.0	24.57	0.0009*

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

*The items with $p \le 0.05$ exhibit item bias.

Items with $p \le 0.05$ exhibit item bias while item with p > 0.05 did not exhibit item bias. A cursory look at table 2 above showed that 44 items (88%) of the 50 NECO examination items exhibit item bias while 6 items (12%) did not exhibit item bias.

DISCUSSION

The result of hypothesis 1 reveals that mathematics examination items constructed by WAEC exhibit item bias. The result of the study is in conformity with the finding of Aborisade (2016) who found out that examination items constructed by WAEC exhibit items bias.

The results of hypotheses 2 showed that majority of the Mathematics examination items constructed by NECO exhibit item bias. The result of the study is in conformity with the finding of Ogbebor and Onuka (2013) in their study that investigated items that are bias using National Examination Council (NECO) Economics questions for 2010 reported that using Logistic Regression statistic detected items that are biased against sub-groups of students. Also, Obinne, Nworgu and Umobong (2013) revealed that research evidence has implicated test used in

Vol.6, No.2, pp.1-7, May 2018

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

National and regional examination as functioning differently with respect to different subgroup of the examinees.

CONCLUSION AND RECOMMENDATION

The findings of the study showed that majority of the mathematics items constructed by WAEC or NECO exhibit item bias. Based on the findings of this study, the following recommendations are made:

- (1) WAEC and NECO should ensure that item bias analysis is done for all the examination items constructed by them in order to produce bias-free items
- (2) Teachers and Examiners who are involved in the construction of national examination items should be trained on item writing. This would enable them construct bias- free tools which will improve the quality of students' assessments
- (3) IRT framework should be incorporated into educational assessment in Nigeria especially in National Examinations. This would help provide for an objective assessment.

REFERENCES

- Aborisade O.J. (2016). Comparative Analysis of item Bias of the mathematics examination constructed by WAEC and NECO in Nigeria. An unpublished PhD Thesis of the Ekiti State University, Ado- Ekiti, Nigeria.
- Berk, R.A (2007). Item Bias Detection Method for Small Samples. Dissertation International.
- Emaikwu, S. O. (2012). Issues in Test Item Bias in Public Examinations in Nigeria and Implications for Testing. *International Journal of Academic Research in Progressive Education and Development* 1 (1) 175-187.
- Federal Republic of Nigeria (2014): National Policy on Education, 2014 edition., Lagos, NERDC.
- Obinne, A.D.E., Nworgu, B. G., &Umobong, M. E. (2013). An Investigation into Differential Item Functioning of Tests Conducted by the two major Examination Bodies in Nigeria. Advances in Educational Research 2 (1), 001-008.
- Ogbebor, U & Onuka, A. (2013). Differential Item Functioning method as an item bias indicator. Educational Research, 4(4); 367-373.