Differential Item Functioning Technique for Detection of Item Bias in Economics among Secondary School Students in Enugu State.

[1]Boi Moses Afam, [2]Ezeamaenyi M.V.O.N., [3]Ebuoh C. N. & [4]Benjamin C. Okechineke

[1,2&3]Department of Mathematics and Computer Education, Enugu State University of Science and Technology (ESUT), Enugu State.
[4]Department of Science Education, Alex Ekwueme Federal University Ndufu Alike, AE-FUNAI, Ebonyi State.
[1]Email: afamm4god@yahoo.com

ABSTRACT

This study focused on differential item functioning (DIF) technique for detection of item bias in economics among secondary school students in Enugu State. Three research questions guided the study with a hypothesis tested at significance level of .05. Proportionate sampling technique was employed in obtaining the sample of 1008 students out of the 20,158 population for the study. The instrument for data collection was NECO (SSCE) 2018 economics objective of 60 multiple choice standardized items. Logistic Regression Analysis was used to answer for research questions 1 &2 while mean (x̄) and standard deviation (SD) was used for research question 3. Two-way ANOVA was used to test the hypothesis. The findings identified significant DIF in gender of students and location of school. The findings also revealed that there is no significant difference between the mean achievement scores of male and female students in the both urban and rural locations. Based on the findings, it was recommended amongst others that seminars, workshops & conferences be organized for researchers, teachers, lecturers, and other stake holders on the general principle of testing and measurement and regular evaluation practices be carried out to minimize DIF. Attention should also be paid to gender and school location, when constructing test items.

Keywords: Differential Item Functioning (DIF), Test, Item Bias, Economics, Detection

INTRODUCTION

Education is a purposeful activity directed at achieving certain aims, such as transmitting knowledge or fostering skills and character traits. According to Obanya (2019), education is a process of systematic all-around development of individuals as they fit into society, aided by genetic and environmental enabling factors. Meanwhile, evaluation entails judging the effectiveness, efficiency or otherwise of a process or outcome of a programme. One of the roles of evaluation as outlined by Anigbo (2014), is that it helps in diagnoses of learning difficulties.

A test, in a narrow sense, is a set of questions to which an individual is expected to respond to in order to obtain a measure that is a numerical value or attribute of such a person. A more encompassing definition of test sees test as a measuring instrument for knowledge, skills, feelings, intelligence or aptitude of any individual (Iketaku, 2013). A test can be a multiple choice, matching, true/false, completion/short answer and essay. Testing is one of the most important parameters by which a society judges the product of her educational system.

It is important to always find out the extent to which the learners have acquired the theoretical and practical skills needed for

both personal and national development. The process through which this is ascertained is known as examination or testing. Examination is the process which comes after a period of learning, and it is an organized assessment of an individual's performance, on the basis of his or her institutional procedural exposure. (Adeyemi and Akindele 2012),

If a test is seen as an instrument designed to measure human behaviour according to some rules, what then is a fair test? A fair test is a standard one on how testing should be administered and outcomes of such test should have nothing to do with race, religion, gender, location, ethnic or so socioeconomic background (SEB) of the examinees or testees (Roever, 2005). Fairness, as defined by the Standards, is "the principle that every test taker should be assessed in an equitable way" (American Educational Research Association, American Psychological Association and National Council on Measurement in Evaluation, 2014). It is a construct that applies not just to test items, but the development of test materials, administration of test items, and dissemination of test information.

The most serious concerns voiced so far against testing pivots around the social issues that test may show culture or class bias (Anastasia & Urbina, 2013).

The issues of bias in testing is currently appearing in public forums including courts of law, and decisions are being made that have an impact on critical issues such as who shall be educated and who shall be employed (Berk, 2015). To examine such an issue requires at least a statistical approach to test analysis which is able to find initially whether the test items are functioning differentially among test taking groups and finally detect the sources of this variance (Geranpayeh & Kunnan, 2006). One of the

approaches suggested for such purposes is Differential Item Functioning (DIF). The issue of test bias has been the subject of a great deal of investigations in recent years, and a technique called Differential Item Functioning (D1F) analysis has become the new standard in psychometric bias analysis. Test bias is often examined at the item level, with differential item functioning (DIF) analyses being part of the framework for probing item bias. (Geranpayeh & Kunnan, 2016).

Differential Item Functioning (DIF) is a statistical technique used to examine the existence of item bias and occurs when an item on a test functions differently for different groups, given the ability level. (Geranpayeh & Kunnan, 2016). Differential item functioning (DIF) which is also referred to as measurement bias, occurs when people from different groups (commonly gender or ethnicity) with the same *latent trait* (ability/skill) have a different probability of giving a certain response on a questionnaire or test. Zumbo (2013) explained that differential item functioning (DIF) occurs when examinees from different groups show differing probabilities of success on the item after matching on the underlying ability that the item is intended to measure. DIF analysis provides an indication of unexpected behaviour of items on a test. An item does not display DIF if people from different groups have a different probability to give a certain response; it displays DIF if and only if people from different groups with the same underlying true ability have a different probability of giving a certain response.

Differential item functioning (DIF) is a statistical characteristic of an item that shows the extent to which the item might be measuring different abilities for members of separate subgroups. Average item scores for subgroups having the same overall score on the test are compared to determine whether

the item is measuring in essentially, the same way for all subgroups. The presence of DIF requires review and judgment, and it does not necessarily indicate the presence of bias.

In the course of this study, amongst other variables, gender and location influence on achievement on Economics were examined. The issue of gender and students' academic achievement have been inconclusive. Some researchers are of the view that male students perform better than females, others disagree with this view, arguing that achievement is a factor dependent on several factors such as socio – economic background, and teaching method among others.

Therefore, one sees that the issue of gender and location of study have not yet been resolved particularly, in relation to students' achievement, hence, the need for further study in that regard, especially when trying out new approaches to solving issues of achievement.

The varying levels of students' academic achievements in schools across various subjects have been a subject of investigation all over the world. In Nigeria today, the achievement of candidates in senior school certificate examinations conducted by both West African Examinations Council (WAEC) and the National Examinations Council (NECO) has not been encouraging. Looking at the perception of people critically on such national examinations in Nigeria, there are likely to be serious item bias in what we called standardized examination.

Purpose of the Study
The main purpose of the study was to detect the item bias in NECO Economics question paper in senior secondary school certificate examinations (SSCE) using Differential Item Functioning Techniques (DIF).

Specifically, the study determined:
1. Test items on standardized NECO 2018 economics multiple choice question paper that exhibit differential item functioning (DIF) with regards to gender.
2. Test items on standardized NECO 2018 economics multiple choice question paper that shows DIF with respect to school location.
3. Mean achievement of students in standardized NECO 2018 economics question paper due to gender and school location.

Research Question
The following research questions guided the study.
1. To what extent do test items on standardized NECO 2018 economics multiple choice question paper exhibit differential item functioning (DIF) with regards to gender of students?
2. To what extent do test items on standardized NECO 2018 economics multiple choice question paper show DIF with regards to school location?
3. What are the mean achievements of students on standardized NECO 2018 economics multiple choice question paper with regards to gender and school location?

Hypothesis
There is no significant difference between the mean achievement score of male and female students in urban and rural senior secondary schools in NECO 2018 economics multiple choice question paper.

Methodology
The research design adopted was ex-post facto research design. This type of research though establishes relationship between variables and assigns cause and effect status to variables but has no control over variables. The area of study was Enugu State, with six education zones namely:

Agbani, Awgu, Enugu, Nsukka, Obollo-Afor and Udi education zones. The sample of the study was made up of 1008 SS3 economics students who registered for NECO examination 2018 in Enugu State. The instrument for data collection was the NECO (SSCE) 2018 economics objective question paper which has undergone standardization process before being administered to the examinees. Logistic regression analysis was used to answer research questions 1 & 2 by analyzing the responses of the students on the achievement test using the SPSS Mean and standard deviation were used to answer

research question 3. Two-way analysis of variance (ANOVA) was used to test the hypothesis.

Results
The results are presented according to research questions and hypotheses as follows:

Research Question One
To what extent do test items on standardized NECO 2018 economics multiple choice question paper exhibit differential items functioning (DIF) with regards to gender?

Table 1: Results of DIF analyses for on 60 NECO economics items with regards to gender.

Item	B	S. E	Wald	Sig	Exp (B)	95% C.I for Exp (B) Lower	Upper
1.	.082	.156	.532	.661	.916	.567	1.262
2.	_.048	.152	.052	.709	1.034	.786	1.861
3.	.072	.181	.329	.525	1.176	.706	1.458
4.	_.016	.155	.016	.921	1.114	.766	1.304
5.	.031	.153	.038	.916	1.003	.761	1.410
6.	_.066	.444	.089	.777	.962	.719	1.379
7.	.211	.145	.486	.462	.798	.669	1.282
8.	_196	.147	1.428	.312	1.198	.988	1.713
9.	.241	.153	.836	.444	1.135	.864	1.451
10.	.408	.162	7.129	.006*	1.503*	1.112	2.016
11.	.039	.174	.153	.616	1.139	.681	1.964
12.	.186	.142	1.494	.236	1.190	.899	1.589
13.	.071	.726	.048	.821	.941	.457	.1.616
14.	_162	.149	1.981	.271	1.173	.888	1.567
15.	.107	.172	.336	.551	.898	.636	1.276
16.	.079	.191	.266	.691	.825	.631	1.346

17.	‒081	.152	.371	.544	.819	.616	1.333
18.	‒201	.270	1.271	.267	.722	.577	1.200
19.	‒066	.148	.526	.441	1.113	.784	1.382
20.	‒016	.180	.009	.951	.989	.711	1.390
21.	‒182	.141	.266	.703	.921	.663	1.269
22.	‒414	.180	3.416	.078	1.294	.972	1.909
23.	.069	.162	.108	.847	.911	.677	1.316
24.	‒028	.124	.019	.981	1.100	.753	1.349
25.	‒018	.164	.018	.906	1.031	.654	1.378
26.	.098	.129	.476	.640	.810	.574	1.230
27.	‒472	.111	2.818	.096	.691	.499	1.072
28.	.235	.170	2.222	.109	.716	.755	1.082
29.	.306	.189	2.269	.141	1.355	.916	1.899
30.	.122	.145	.764	.124	1.53	.389	1.563
31.	.151	.162	.556	.654	1.140	.800	1.617
32.	‒123	.175	.556	.654	1.391	.908	1.710
33.	.319	.161	4.242	.006*	1.374	1.018	1.762
34.	‒124	.132	.679	.448	1.128	.836	1.501
35.	.114	.184	.002	.976	.991	.745	1.333
36.	.041	.189	.091	.812	.714	.709	1.311
37.	‒112	.163	.540	.461	1.811	.850	1.516
38.	.100	.662	.102	.454	.309	.850	1.804
39.	‒038	0.541	.081	.268	1.134	.756	1.401
40.	.247	.249	.809	.143	1.268	.778	2.067
41.	‒256	.148	1.490	.461	.776	.538	1.101
42.	.081	.157	.322	.360	.911	.651	1.925
43.	‒026	.641	.022	.816	1.023	.768	1.363
44.	.070	.150	.119	.226	1.073	.728	1.741
45.	.213	.551	2.002	.361	.974	.591	1.075

46.	.342	.162	4.743	.039	1.411	1.039	1.916
47.	_013	.125	4.342	.307	1.371	1.001	1.859
48.	.014	.161	1.341	.214	.232	.610	1.133
49.	.178	.156	1.564	.412	.380	.816	1.111
50.	_033	.149	.026	.786	1.624	.764	1.371
51.	.189	.156	1.103	.524	1.218	.813	1.671
52.	.021	.189	.619	.526	1.018	.691	1.502
53.	.312	.148	2.004	.135	.808	.306	1.081
54.	_061	.344	.648	.828	1.051	.536	1.701
55.	.236	.144	2.934	.221	1.622	.939	1.706
56.	_182	.581	1.056	.308	1.276	.963	1.600
57.	.042	.166	3.969	.029*	1.406	1.406	1.042
58.	.031	.207	.022	.880	1.040	.696	1.446
59.	_067	.581	.174	.678	1.066	.738	1.455
60.	.204	.149	1.906	.681	1.227	.981	1.737

(a) Variables entered on step 1: Gender DIF exist.

Table 1 shows four items that identified significant DIF in gender of students using logistic regression method of analysis of SPSS version 20. Items 10, 33, 46 and 57 reveal significant differences between male and female students with significant level less than 0.05.

For item 10, the logistic regression model or the equation shows significant difference at $(X^2= N =1,008, P < 0.05)$.

For item 33, the logistic regression model or the equation shows significant difference at $(X^2= 5.461, N =1,008, P < 0.05)$.

For item 46, the logistic regression model or the equation showed significant difference at $(X^2= 50.029, N =1,008, P < 0.05)$.

For item 57, the logistic regression model or the equation showed significant difference at $(X^2=72.304, N =1,008, P < 0.05)$.

Research Question Two

To what extent do test items on standardized NECO 2018 economics multiple choice question paper show DIF with regards to location?

Table 2: DIF analyses for School Location on 60 NECO economics items.

Item	B	S. E	Wald	Sig	Exp (B)	95% C.I for Exp (B)	
						Lower	Upper
1.	.119	.174	.568	.491	.890	.644	1.332
2.	. 485	.167	10.227	.000*	1.723	1.351	2.496
3.	.296	.172	3.557	.062	1.341	.992	1.911
4.	.555	.251	7.283	.310	.661	.544	.812
5.	.112	.029	.309	.581	1.091	.817	1.496
6.	_ .286	.148	5.863	.298	.682	.905	1.907
7.	_ .681	.451	1.640	.229	.883	.694	1.321
8.	_ .113	.163	.684	.000*	.894	.684	1.332
9.	_ .119	.174	.747	.026*	.896	.726	1.791
10.	.216	.289	.610	472	1.116	.808	2.096
11.	.226	.351	2.006	.149	.789	.529	1.078
12.	.448	.159	2.999	.489	.648	.577	.868
13.	.631	.728	.273	.628	1.451	.668	1.876
14.	_ .103	.192	1.110	.316	1.312	.843	1.677
15.	.716	.781	6.680	.169	.595	.734	1.779
16.	.192	.129	1.011	.316	1.213	.834	1.776
17.	_ .184	.148	.986	.231	1.169	.866	1.550
18.	.021	.129	.018	.901	1.022	.734	1.244
19.	.159	.174	1.661	.003*	1.172	.987	1.652
20.	.185	.174	1.661	.290	1.173	.897	1.566
21.	.239	.116	3.962	.079	1.341	.918	1.938
22.	`.082	.183	.402	.645	.926	,646	1.327
23.	.403	.172	4.172	.085	.783	.529	1.038
24.	_ 048	.156	.089	.787	1.046	.757	1.555
25.	.062	.451	.139	.719	1.066	.783	1.421
26.	.066	.154	.131	.018*	1.066	.781	1.824
27.	.061	.152	.131	718	1.058	.781	1.198
28.	_ .665	.173	3.263	.196	1.928	1.693	2.688
29.	.089	.300	.145	.706	1.082	.792	1.591
30.	.006	.166	.112	.589	.990	.637	1.356
31.	.122	.146	.749	.388	1.731	.856	1.622
32.	_ .201	.177	1.863	.240	1.226	.871	1.357
33.	_ .401	.150	4.241	.039*	.717	.650	.998
34.	.719	.143	6.419	.306	.658	.512	.966
35.	.172	.519	1.230	.251	.840	.632	1.229
36.	.046	.162	.065	.720	1.048	.767	1.314
37.	.024	.562	.028	.001*	1.031	.578	1.843
38.	.601	.322	5.457	.298	1.956	1.025	3.711
39.	.218	.251	2.626	.616	.563	.488	.889
40.	.401	.252	1.496	.222	1.362	.831	2.334
41.	_ .176	.138	.821	.461	.848	.590	1.412
42.	.116	.157	.452	.511	1.221	.798	1.596
43.	.06	.142	.154	.690	.942	.702	1.385
44.	.256	.361	4.352	.410	1.735	1.258	2.583

45.	.296	.156	2.169	.183	1.285	.931	1.711
46.	.169	.117	1.218	.296	1.413	.881	1.866
47.	.160	.163	.864	.426	.862	.883	1.361
48.	− .144	.517	.876	.002*	.873	.936	1.394
49.	.472	.166	10.155	.218	1.665	1.213	2.691
50.	.353	.148	6.974	.279	.694	.519	.936
51.	− .913	.166	6.781	.261	.391	.288	.549
52.	.151	.189	.666	.497	.896	.590	1.282
53.	.319	.144	2.112	.226	.728	.639	.966
54.	− .019	.422	.009	.041*	.892	.691	1.685
55.	.266	.181	4.128	.811	.536	.498	.823
56.	.112	.463	7.531	.200	.729	.561	.959
57.	− .610	.125	4.216	.091	.524	.600	.835
58.	.148	.711	.846	.734	.868	.677	1.399
59.	.172	.160	1.244	.243	1.209	.889	1.548
60.	− .402	.248	3.198	.000*	.639	.551	.998

(a) Variable(s) entered on step 1: School Location, DIF exist.

Table 2 shows ten items that identified significant DIF. In School Location of students, using logistic regression method of analysis of SPSS version 20, items 2, 8, 9, 19, 29, 33, 37, 48, 54 and 60 reveal significant difference between urban and rural students with significant level less than 0.05.

For item 2, the logistic regression model or equation showed significant at (X^2=25.536, N =763, P < 0.05).

For item 8, the logistic regression model or equation showed significant at (X^2=7.463, N =1008, P < 0.05).

For item 9, the logistic regression model or equation showed significant at (X^2=16.051, N =1008, P < 0.05).

For item 19, the logistic regression model or equation showed significant at (X^2=86.086, N =1,008, P < 0.05).

For item 29, the logistic regression model or equation showed significant at (X^2=134.159, N =1,008, P < 0.05).

For item 33, the logistic regression model or equation showed significant at (X^2=8.163, N =763, P < 0.05).

For item 37, the logistic regression model or equation showed significant at (X^2=12.269, N =763, P < 0.05).

For item 48, the logistic regression model or equation showed significant at ($X2$= 15.281, N =763, P < 0.05).

For item 54, the logistic regression model or equation showed significant at (X^2=4.616, N =763, P < 0.05).

For item 60, the logistic regression model or equation showed significant at (X^2=5.398, N =763, P < 0.05).

Research Questions Three.
What are the mean achievement scores of students on standardized NECO 2018 economics multiple choice question paper with regards to gender and location? '

Table 3: Mean scores and standard deviation of students in NECO 2018 economics question paper due to gender and location.

Gender	Location					
	Urban		Rural		Total	
	Mean	SD	Mean	SD	Mean	SD
Male	56.90	14.16	56.74	17.21	56.82	15.72
Female	55.11	16.93	55.31	15.69	55.21	16.27
Total	56.01	15.60	56.03	16.31	56.02	15.92

Table 3 shows that male and female students in urban and rural schools obtained nearly equal mean scores ranging from 55.31 to 56.90. The total mean scores range from 55.21 to 56.82.

Research Hypothesis

There is no significant difference between the mean achievements scores of male and female students in Urban and Rural senior secondary schools in NECO 2018 economics multiple choice question paper.

Table 4: Two-way ANOVA results due to Gender and Location.

Scores	Type III sum of squares	df	Mean square	F	Sig	Decision
Corrected model	0.685	3	.228	.011		
Intercept	1923.446	1	1923.446	95.333		
Gender	.347	1	.347	.017	.211	NS
Location	.167	1	.167	.008	.221	NS
Gender x location	.021	1	.021	.001	5.31	NS
Error	2025.671	1004	20.176			
Total	2752.235	1006				
Corrected total	16.629	1007				

Table 4 shows that for gender, the F computed value of 0.017 is significant at 0.211 which is higher than 0.05 set for this study. Hence, gender is not significant. For location the F computed value of 0.008 is significant at 0.221 level which is higher than 0.05 for this study. Hence, location is not significant. For interaction (gender x location), the F computed value of 0.008 is significant at 0.361 level which is higher than 0.05 set for this study. Hence, interaction effect is not significant.

Summary of the Findings of the Study and Conclusion

The following are the summaries of the findings of the study.

1. Four items identified significant DIF in gender of students. They are item 10, 33, 46 and 57. Item 10 and 46 favoured female students, while items 33 and 57, favoured male students.

2. Ten items identified significant DIF in school location of students. They are item 2, 8, 9, 19, 26, 33, 37, 48, 54 and

60. Rural students were advantaged in items 2, 8 and 48 than urban students. Urban students were advantaged in items 9, 19, 26, 33, 37, 54 and 60 more than their rural counterparts.
3. There is no significant difference between the mean achievements scores of male and female students in urban and rural senior secondary schools in NECO 2018 economics multiple choice question paper.
4. It is evident that students from urban schools outperformed their counterparts from rural schools.

In conclusion, the findings of the study show that secondary school students' academic achievement in Enugu State have some elements of DIF that distinguished between subgroups in NECO (SSCE) Economics objective test items. Hence, researchers in the education sector should place more emphasis on the best way of constructing valid and reliable test items which are fair for effective assessment of students' performance in achievement test.

Recommendations
Based on the findings of the study, the following recommendations were made.
1. Seminars, workshops and conferences should be organized for researchers, teachers, lecturers, psychometricians on the general principle of testing and measurement to avoid or minimized DIF. This is because DIF is relatively new in Nigeria.
2. Evaluators, educational practitioners, researchers and examining bodies involved in test development instrument should use logistic regression for data analysis to ensure consistency or reliability or validity of the test items.
3. Studies of this nature are recommended to be carried out to provide further empirical evidence on the fairness of test items.

4. Examinations should be properly validated by relevant test experts to ensure consistency or reliability or validity of the test items.
5. NECO should make sure that items to be used to examine their students are free from any kind of DIF.

Discussion of findings
Gender bias occurs in educational tests when a student is offended or unfairly penalized because the student is a male or female. A logistic regression analysis was carried out on some students to predict the group of gender performance using SPSS Version 20. The result from table 1 reveals that four items that has DIF against subgroups that consist of male and female examinees. Items 10, 33, 46 and 57 differentiated significantly between male and female students in which items 33 and 57 favoured the male students and placed the female examinees at the disadvantaged group. Items 10 and 46 favoured the female students and placed their male counterparts at disadvantaged group. This is because there was confusion in choosing between the correct answers and the distracters which gave rise to DIF and also lack of comprehension of the different concepts of the topic as well as the theme that are not familiar to the subgroups of examinees, teacher incompetence in the teaching economics subject content for master. These could cause DIF between the subgroups of examinees.

For research question 2, the result from table 2 reveals 10 items that has DIF against subgroups that consist of urban and rural examinees. Items 9, 29,33,37,54 and 60 favour the urban students and placed the rural examinee at the disadvantaged group. Item 2, 8, 19 and 48 favour the rural students and placed the urban students at the disadvantaged group. The significant differences in school location could be as a result of the structure of the constructed

questions which may be unfamiliar to the affected group of examinees. It could also be due to lack of exposure to the topic content and concept that may have reflection on the items. A qualified personal, moving to a particular location could also cause DIF when the affected groups is being taught by unqualified teachers or the affected group may have received limited lesson on the subject being taught. This finding implies that students that attend urban secondary school do perform better than their counterparts in the rural secondary schools. This further indicates that NECO (SSCE) 2108 economics multiple choice objective test items in assessing student ability has elements of biases that places urban students' examinees at advantaged group and places rural examinees at disadvantaged group.

Table 3 shows that there is no significant difference between the mean achievement scores of male and female students in urban and rural senior secondary school. In NECO 2018 economics multiple choice question paper. This means that male and female students in either urban or rural schools achieved alike in their mean scores. This finding agrees with the findings of Ikelaku (2019) and Obiake (2019) that both male and female achieved equally in maths. This finding disagrees with that of Shakness and DeAngelo (2011) that there exists gender difference in student's performances in economics.

Conclusion
The findings of the study shows that secondary school students' academic performance in Enugu State have some elements of DIF that distinguished between subgroups in NECO (SSCE) objective test items. Hence, researchers in the education sector should place more emphasis on the best way of constructing valid and reliable test items which are fair of effective assessment of students' performance in achievement test.

Reference
Adeyami, J. and Akindele, I. (2012). Examination Malpractices in Nigeria of Educational System; Causes, Effects and the Way Out. *Nigerian Journal of Clinical and Counselling Psychology*, 14(1)

Anastasia, A. and Urbina, S (2013), *Psychological Testing*. New Delhi: Prentice Hall.

Anigbo, L.C. (2014). *Teachers' handbook on measurement and evaluation*. Enugu. Executive Press Resources limited.

Berk (2015). Item Bias Detection Methods for Small Samples. *Dissertation abstract International.*

Boi, M.A. (2021). Differential Item Functioning Technique For Detection Of Item Bias In Economics Among Secondary School Students In Enugu State. *Unpublished Masters thesis*, Department of Mathematics and Computer Education ESUT, Enugu.

Iketaku, R.I. (2019). Standardization of a Mathematics Achievement Test for Senior Secondary Schools. *Unpublished Ph.D. Dissertation*, Department of Science Education ESUT, Enugu.

Nworgu, B.G. (2013). Educational Measurement and Evaluation: Theory and Practice. Nsukka: University Trust Publishers. 266 pp

Obanya, P. (2019). *Education, education, education: The answer to all societal problems*. Mosuro Publishers.

Roever, C. (2005). "That is not fair" Fairness Bias and Differential Item Functioning in Language Testing. Retired Nov. 18, 2006 from http://www2.howaii.edu/2roever/brewnbag.pdf. Joint committee on Testing Practices (2004) code of fair testing practice in education Washington, D.C; Author

Zumbo, B.E. (2013). *A handbook on the theory and methods of differential item functioning (DIF): scores* Ottawa, ON: Director of Human Resources and Evaluation, Department of National Defence.