DIFFERENTIAL ITEM FUNCTIONING OF NATIONAL EXAMINATIONS COUNCIL (NECO) AGRICULTURAL SCIENCE QUESTIONS FOR 2015

Kelechi Joshua Princewill Ihechu & Adolphus Onuabuchi Madu

College of Education
Michael Okpara University of Agriculture
Umudike, Abia State, Nigeria
Email: donsignor4ray@yahoo.com & Adolphus4god@yahoo.com

Abstract

Differential item functioning is meant to find out items that are biased. This study investigated items that are functioning differently in relation to school type (private and public schools), school location (urban and rural schools) using National Examinations Council (NECO) Agricultural Science questions for 2015. The research design employed in this study was a comparative research type of design. The study sample comprised students in Imo State, Nigeria. Four hundred and forty seven (447) students were used. And the test contains 60 items which was administered to the students. Logistic regression was used to analyse the data. The research findings showed that out of sixty items in NECO Agricultural Science questions 11 items were biased in relation to school type and 9 items in relation to school location. The implication of these findings is that NECO Agricultural Science examinations questions have items that are biased along school type and location dimensions. From the result of the findings, it was then recommended that test experts and developers should explore the use of logistic regression in detecting items that are biased before administering them.

Keywords: Differential Item Functioning, Logistic Regression, Item Biased, Latent Trait

Introduction

As we have seen, psychological tests can be well-conceived and well-constructed, but none is perfect. The reliability of test scores can be compromised by random measurement error (unsystematic error), and the validity of test score interpretations can be compromised by response biases that systematically obscure the psychological differences among respondents.

Psychological tests are often used to make important decisions that affect the lives of people, for examples, which colleges (if any) will decide to accept you, in which class will your child be enrolled, and will an employer decide to hire you? To the extent that such decisions are based on tests that are biased in favor of or against specific groups of people, such biases have extremely important personal and societal implications.

The issue of educational measurement in research pointing towards enhancing the fairness of test or examination across sub groups of examinees is very essential because important decisions are made based on scores of the examinees. Test consists of a set of uniform questions or task to which a student or testees is to respond independently and the result of which can be treated in such a way as to provide a quantitative comparison of the performance in different students (Nworgu, 2011). The term testees or examinees can be used interchangeably. It implies an individual or group of individuals who are examined by a standardized or teacher made examination. Ogbebor, (2012) opined that, tesstees or test takers of the same latent trait should respond to test items correctly irrespective of their gender, school location and school type. Test fairness can be viewed as any test given to a set of testees with an equal chance to demonstrate what they know. Various aspects of fairness in testing have been highlighted in literature, including fairness in regards to standardization, test consequences/score use, and item bias (Kunnan, 2000; Shohamy, 2000).

A fair test is one that affords all examinees an equal opportunity to demonstrate the skills and knowledge which they have acquired and which are relevant to the test's purpose (Roever, 2005). The existence of bias is an issue to be addressed because tests are used as gatekeepers for educational opportunities and it is a very important issue that test items are fair for every examinee. Bias is the existence of some irrelevant elements present in items that cause differential performance for individuals of the same ability but from different ethnic, sex, type of school attended, location of schools and cultural or religious groups. An examination item is said to be biased if it functions differently for a specified subgroup of test takers. Ogbebor, (2012) states that biased test measure characteristics that are not necessary or items that are irrelevant to the test. Frequently, examination items are considered biased because they contain sources of difficulty that are not relevant to the construct being measured and these extraneous sources affect test-takers' performance (Zumbo, 2009).

Item bias or differential item functioning (DIF) has critical political, social and ethical implications for test developers, policy makers and examines. The study of item bias and DIF is critical as such, this research would help to provide an empirical foundation for the identification and subsequent elimination of examination items that appear to be relatively more difficult for one group of test-takers than another. Further research on these issues will allow us to comprehend more fully the possible substantive interpretation that can be made by focusing on test items considered to be biased.

Differential item functioning (DIF) is an approach used to assess the existence of item bias, it is a systematic error in the predictive or construct validity of an item that may be attributable to factors irrelevant to the test. Camilli, (2003), states that DIF specify whether individuals of equal ability have the same probability of getting a given item correct. The modern approach for detecting item bias is by providing evidence of DIF. According to Roever, (2005), locating items on which a group of examinees perform significantly better than another group is logically the first step in detecting item bias. If an item on which a particular subgroup performs significantly better than another subgroups, it is said to have functioned differentially with respect to the two groups. Ogbebor (2012), noted that DIF occurs when a test item measures an ability which is unfamiliar to the subject matter, such that students' scores on the item is now sustained by abilities which are unfamiliar to the subject matter.

Perhaps the best way to evaluate construct bias is a procedure called differential item functioning analysis. Differential item functioning analysis is a feature of a psychometric approach called Item Response Theory (IRT). An important aspect of IRT is the assumption that it is possible to estimate respondents' trait levels directly from empirical sources of data. The trait levels are, in essence, estimates of participants' true scores for the psychological attribute that is being measured. If we assume that we know the trait levels for all the people in two groups and we have their responses to a test item, then we can see if the trait levels and the item responses match-up in the same way for both groups. If they do not, then it is possible that the item is biased.

IRT is based on the idea that there is a function relating a participant's trait level to the probability that he or she will answer a question on a test correctly. For example, we might find that an individual with a trait level that is one standard deviation above the mean has a .80 probability of answering a particular item correctly, but that an individual with a trait level that is one standard deviation below the mean has only a .20 probability of answering the item correctly. If you have a group of people take a test and you know their respective trait levels, then you can use specialized statistical software to draw an item characteristic curve (ICC) to illustrate this function for each item. Furthermore, if you have two groups of people, then you can draw ICCs separately for each group. To evaluate the presence of construct bias, you would compare the ICCs of the two groups. If the item is not biased, then the two groups' ICCs should be very similar. That is, the probability that two people will answer an item correctly should be the same if the two people have the same trait level. However, if the item is biased, then the two groups ICCs will be dissimilar. That is, the probability that two people (e.g., a male and a female) will answer an item correctly might be different even if the two people have the same trait level. Such a situation would clearly reflect the presence of construct bias.

Studies have shown that differences in test of student achievement and low test scores in some subject areas such as Agricultural Science and Economics could be attributed to social and cultural influences that create sex role stereotypes that reduce female interest and achievement in traditionally male-dominated subjects (Williams et al., 2002; Hirschfeild et al.,2005). Studies have also shown that there are significant differences in the academic performance of students from rural and urban areas. Obe, (2004), observed that there is a significant difference in the performance of students from rural and urban schools in their academic performance; he therefore concluded that children from urban schools were superior to their rural counterparts. Owoyeye, (2002) also found out that there was a significant difference between academic performance of students in rural and urban area in public examinations. However, Ajayi and Ogunyemi, (2000) and Gana, (2007) in their different studies on the relationship between academic performance and school location revealed that there was no significant difference of students in urban and rural schools. While Ajayi, (2009) also found out that there was no significant difference between students academic achievement of rural and urban secondary school students.

A lot of research works have been conducted in this area of item bias. Pedrajita, (2009), in a study "using Logistic regression to detect test items in Chemistry Achievement", the result from the study revealed that there are gender bias and class bias in Chemistry Achievement test. Nworgu, (2011), revealed that current research evidence has implicated test used in national and regional examination as functioning differently with respect to different subgroups. This means that students' scores in such examinations are determined largely by the group to which an examinee belongs and not by ability. Gierl's, (2009), a study on DIF in Alberta examined 30 education Social Studies Diploma students, the study evaluated the effects of DIF between male and female, the results indicated that the majority of multiple choice items did not display DIF using the three-tiered ratings. Thus, 65 of 70 items displayed negligible effects, 5 items with moderate DIF, three favoured male and two favoured female, this indicate that the test contained items that functioned differently for male and female.

Item bias is of a particular concern on test of Agricultural Science subject in students' academic achievement, here differences in performance between, private and public urban and rural is commonly found. Therefore, this study is focused on differential item functioning based on school type and location in Agricultural Science National Examinations questions.

Research Questions

The following research questions are raised in order to achieve the objective of the study.

1. What are the presence of differential item functioning in terms of school type (private and public)
2. What are the presence of differential item functioning in terms of school location (rural and urban)

Methodology

A comparative research design was adopted for the study. This design is considered appropriate because it attempts to establish cause effect relationship among the variables in the study. The target population comprises all students in SSS 3 in Imo State, Nigeria, who enrolled for the Senior School Certificate Examination in 2016. The accessible population of the study was made up of 1096 male and 896 female Agricultural Science students of the senior secondary schools who enrolled for the Senior School Certificate Examination in 2016.

Based on Robert and Morgan's (1970) formular for determining sample size from a known/finite population, a sample of 272 Agricultural Science students in SSS 3 who enrolled for the Senior School Certificate Examination in 2016. Multi-stage sampling techniques were used for the study. This sample was drawn from one zone out of three educational zones in Imo state using simple random sampling technique. The sample of the schools, teachers and students also involved the use of simple random sampling and purposive sampling leading to different stages of sampling. In the first stage, a list of the Local Government Areas under the selected zone was made before a simple random sampling of the Local Government was conducted. Secondly Okigwe Educational Zone was cluster into urban and rural areas. Purposive sampling was employed to select three (4) private schools and three (4) public schools from the urban area, and four (4) private school and three (4) public schools from the rural areas. The total schools used for this study was 16 secondary schools in Okigwe Educational Zone in Imo State Nigeria. An intact class was used in each of the school sampled.

The response to each item of the NECO Agricultural Science examinations for all the students in the schools selected was used. The NECO Examinations is a standardized examination taken nationwide in Nigeria.

Logistic regression was used to analyze the data. It involved the following steps:

i. Identify Reference and Focal groups of interest usually two at a time.
ii. Design the DIF study to have samples which are large as possible
iii. Choose DIF statistics which are appropriate for the data
iv. Carry out the statistical analyses
v. Interpret DIF statistics/results and delete items or make item changes as necessary.

Results

Research Question 1

What are the presences of differential item functioning in terms of school type (private and public)

Table 1: Logistic Regression to Detect School Type Bias

Item	B	S.E	Sig	Exp (B)	Lower	Upper
1	.157	.225	.483	1.170	.754	1.817
2	.243	.238	.308	1.275	.799	2.035
3	.095	.190	.616	1.100	.758	1.597
4	-.076	.190	.691	.927	.639	1.346
5	-.235	.231	.309	.791	.503	1.243
6	.311	.211	.142	1.364	.902	2.065
7	-.177	.190	.454	. .837	.577	1.216
8	-.339	.191	.004*	..712	.490	1.035
9	.417	.195	.343	1.517	1.035	. 2.224
10	.92	.197	.639	1.097	.746	1.613
11	.242	.218	.268	1.273	.831	1.952

12	-.227	.190	.033*	.797	.549	1.158
13	.663	.201	.531	1.941	1.310	2.876
14	1.039	.361	.004*	2.826	1.393	5.733
15	.249	.202	.219	1.283	.863	1.908
16	-.959	.266	.000* .	.383	.227	.646
17	-.023	.191	.905	.977	.672	1.422
18	-.319	.191	.094	.727	.500	1.056
19	.241	.199	.226	. 1.272	.861	.1.879
20	.317	.193	.101	1.373	.941	2.004
21	.163	.247	.509	1.177	.725	1.911
22	.164	.354	.001*	.897	.6785	1.976
23	-.543	.307	.077	.581	.318	1.061
24	.218	.261	.402	1.244	.747	2.073
25	-.494	.325	.129	.610	.323	1.155
26	-.131	.202	.507	.877	.590	1.304
27	.083	.196	.672	1.087	.740	1.596
28	-458	.266	.085	.632	.375	1.066
29	-.111	.271	.682	.895	.527	1.522
30	.046	.190	.808	1.047	.721	1.521
31	.299	.197	.129	1.349	.916	1.985
32	.122	.256	.635	1.129	.683	1.866
33	.166	.191	.386	1.181	.811	1.718
34	-.141	.216	.513	.868	.568	1.326
35	.204	.198	.290	1.233	.836	1.817
36	.242	.223	.278	1.273	.823	1.971
37	-.140	.201	.486	.869	.587	1.289
38	.374	.287	.192	1.454	.829	2.550
39	.257	.201	.202	1.293	.871	1.918
40	-.326	.198	.100	.722	.489	1.064
41	.086	.191	.653	1.89	.750	1.583
42	.136	.278	.626	1.145	.664	1.974
43	-1.488	.459	.001*	.226	.092	.555
44	.460	.218	.034*	1.585	1.035	2.427
45	.065	.215	.761	1.068	.700	1.628
46	.461	.201	.021*	1.586	1.070	2.350
47	-.209	.280	.455	.811	.469	1.404
48	.263	.207	.203	1.301	.867	1.953
49	.414	.191	.031*	1.513	1.039	2.202
50	-.506	.228	.027*	.603	.386	.943

51	.103	.272	.705	1.109	.650	1.891
52	.106	.245	.666	1.112	.688	1.796
53	.134	.216	.536	1.143	.749	1.744
54	-.071	.193	.711	.931	.638	1.358
55	-.161	.207	.437	.851	.567	1.278
56	.255	.248	.305	1.290	.793	2.099
57	.168	.211	.425	1.183	.783	1.788
58	-.014	.246	.007*	.986	.609	1.595
59	.564	.207	.677	1.758	1.171	2.639
60	-.060	.195	.760	.942	.643	1.381

From Table 1 shows the items in relation to school type (private and public), identified by logistic regression method using SPSS version 21. Out of sixty items in NECO Agricultural Science questions DIF was present in eleven items. These items are item 8, 12, 14, 16, 22, 43, 44, 46, 49, 50, and item 58.

Research Question 2

What are the presences of differential item functioning in terms of school location (rural and urban)

Table 2: Logistic Regression of Sixty NECO Item for School Location

Item	**B**	**S.E**	**Sig**	**Exp (B)**	**Lower**	**Upper**
1	.017	.243	.965	1.017	.550	1.760
2	-.236	.246	.000*	.788	.512	1.28
3	1.220	.200	.340	3.388	2.290	5.012
4	-.403	.191	0.35	.669	.460	.973
5	.194	.233	.406	1.214	.769	1.917
6	-.217	.209	.002*	.805	.535	1.212
7	-.840	.194	.254	.432	.295	.632
8	-.339	.191	.075	.712	.490	1.035
9	-.618	.199	.002*	.539	.365	.796
10	-.370	.197	.059	.690	.469	1.015
11	-.506	.217	.019*	.603	.394	.921
12	.098	.190	.604	1.103	.760	1.602
13	.107	.199	.591	1.113	.754	1.643
14	-.254	.315	.419	.776	.419	1.437

15	-.116	.201	.562	.860	.600	1.320
16	-.432	.249	.084	.650	.398	1.059
17	-.611	.193	.002*	.543	.372	.793
18	.370	.191	.053	1.447	.995	2.105
19	.122	.198	.538	1.130	.766	1.667
20	-.017	.193	.928	.983	.674	1.434
21	-.332	.252	.188	.717	.438	1.176
22	.424	.234	.070	1.528	.966	2.417
23	.080	.293	.785	1.083	.610	1.925
24	-.333	.266	.212	.717	.425	1.209
25	-.087	.314	.781	.916	.495	1.695
26	.276	.202	.171	1.318	.888	1.957
27	.199	.196	.311	1.220	.830	1.792
28	.020	.258	.938	1.020	.615	1.693
29	.324	.269	.228	1.382	.816	2.341
30	-.316	.191	.097	.729	.502	1.059
31	.029	.191	.883	1.029	.701	1.511
32	-.143	.259	.580	.867	.522	1.439
33	-.054	.191	.012*	.948	.651	1.379
34	-.329	.218	.131	.720	.470	1.103
35	.209	.198	.290	1.233	.836	1.817
36	.093	.223	.678	1.097	.706	1.698
37	.181	.200	.366	1.198	.810	1.773
38	.540	.289	.062	1.715	.975	3.020
39	-.068	.202	.737	.934	.629	1.388
40	.333	.197	.091	1.395	.948	2.050
41	-.133	.191	.487	.876	.602	1.273
42	.213	.278	.447	1.237	.718	2.132
43	-.644	.377	.088	.525	.251	1.100
44	-.153	.218	.483	.858	.559	1.316
45	-.405	.219	.065	.667	.434	1.025
46	-.100	.201	.619	.905	.611	1.341
47	-1.069	.314	.001*	.343	.185	.636
48	-.080	.208	.701	.923	.614	1.387
49	-.463	.192	.061	.629	.432	.918
50	-.251	.224	.263	.778	.502	1.207
51	-.195	.276	.476	.822	.479	1.411
52	-.575	.255	.024*	.563	.341	.929
53	-.053	.216	.002*	.948	.621	1.449

54	.598	.194	.365	1.819	1.244	2.661
55	.223	.206	.280	1.249	.834	1.871
56	-.054	.250	.827	.947	.581	1.544
57	.124	.211	.558	1.132	.749	1.710
58	-.258	.248	.299	.773	.475	1.257
59	.265	.206	.198	1.304	.870	1.953
60	-.251	.196	.201	.778	.530	1.143

From table 2 shows the items in relation to school type (private and public), identified by logistic regression method using SPSS version 21. Out of sixty items in NECO Agricultural Science questions DIF was present in nine items. These items are item 2, 6, 9, 11, 17, 33, 47, 52 and 53.

Discussion

Logistic regression statistics detected items that have DIF against subgroups such as public and private schools examinees, and it was revealed that out of the sixty items in NECO Agricultural Science examinations question paper, ten items showed DIF these items are item 8, 12, 14, 16, 22, 43, 44, 46, 49, 50, and 58. Seven item which are item 8, 12, 22, 44 , 46, 49,and 58, favoured private school students while the public school student were disadvantaged, while four items which are item 14,16,43, and 50 favoured public schools than the private schools. The private schools on these items were disadvantaged. This finding is in line with the finding of Ogbebor and Onuka (2013), who found out that there were presences of school type and school location bias in NECO economics questions. The findings of this study agrees with the work of Pedrajita, (2009) when he used Logistic regression to detect test items bias in Chemistry Achievement", the result from the study revealed that there is school type bias in the Chemistry Achievement test that was administered to the testees out of 22 items that were biased 11 items favoured public schools while eleven also favored private schools.

Logistic regression also detected items that have DIF against subgroup such as urban and rural school students, and it was revealed that out of the sixty items in NECO Agricultural Science examinations question paper, nine items showed DIF these items are 2, 6, 9, 11, 17, 33, 47, 51 and 53. From the findings, its observed that these items that showed DIF are due to the structure of the questions and stem, thus these could be the characteristics that affected the test takers response to getting the item correctly. Nworgu, (2011), revealed that current research evidence has implicated test used in national and regional examination as functioning differently with respect to different subgroups. This means that students' scores in such examinations are determined largely by the group to which an examinee belongs and not by ability. Adedoyin (2010) in his study on investigating gender biased items in public examinations; he found that out of 16 test items that fitted the 3PL item response theory statistical analysis, 5 items were gender biased.

The finding of this study agrees with Felder, Mohr, Dietz and Ward (2004) who find out that urban student enjoy greater success than rural student, a result also supported by Tremblay, Ross and Berthelot, (2001), Kolcic, (2006) and Considine and Zappala, (2002). On the other hand the findings of this study disagree with Lee and McIntire, (2001) whose findings revealed that there is no significant difference between performance of rural students and urban students. This implies that items used in assessing student ability has element of biasness that disadvantaged the rural school examinees and favors the urban schools examinees.

Conclusion

Based on the forgoing findings the following conclusions were made. There were presences of school type and school location bias in NECO Agricultural Science questions.

Recommendations

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

i. Test experts and developer should explore the use of differential item functioning method, particularly the use of logistic regression to detect both uniform and no uniform biased items.

ii. A study of this should be conducted to provide further empirical evidence on the validity of the method in detecting biased test items. Evaluators and educational practitioners who are engaged in the development of assessment tools should use logistic regression for bias correction
iii. Measurement practitioners should make use of logistic regression for developing a valid , reliable gender fair test school type fair test with biased items revised or replaced
iv. The subject curriculum should be made clear for teachers to be able to teach the concept effectively
v. Teachers should exposure learners to more than one textbook.

References

Adedoyin O. O. (2010). Using IRT Approach to Detect Gender Biased Items in Public Examinations: *Educational Research and Reviews Academic Journals*. 5(7): 385-399

Ajayi K. & Ogunyemi B. (2000). The Relationship between Instructural Resources and Soci-Economics Status in Selected Population of High School. Dissertation Abstract International, 25(2): 22-23.

Ajayi I. A., (2009). Unit Cost of Secondary Education and Student Academic Performance in Ondo State, Nigeria. Ph.D Dissertation. University of Ibadan, Ibadan

Camilli G. & Shepard L. A. (2004). Methods for Identifying Biased Test Items. Newbury Park, CA: Sage.

Considine G. & Zappala G. (2002). Influence of Social Land Economics Disadvantaged in the Academic Performance of School Students in Australia. *Journal of sociology,* 38, 129 – 148. Retrieved on August 16, 2007 from http:// jos. Sage pub.com

Embretson S. & Reise S. P. (2000). Item Response Theory for Psychologists. Mahwah New Jersey: Lawrence Erlbaum Associates Publishers.

Felder R. M. Mohr P. H. Dietz E. J. & Ward L. B. (2004). A Longitudinal Study of Engineering Student Performance and Retention. II Differences between Students from Rural and Urban Backgrounds. *Journal of Engineering Education,* 83(3):15–21,

Gana E. S. (2007). The Effect of using Visual Designed Training Model on the Learning of Mathematics at J.S.S. Ph.D. Thesis, Unpublished. University of Ibadan,

Gierl M. J. (2009). Test Bias, Sage Publications, Inc., California.

Gierl M. J. Bisanz B. G. Boughton K. A. & Khaliq S. N. (2001). Using Differential Bundle Functioning to Identify and Interpret Gender Difference on Science Achievement Tests. Paper Presented at the Annual Meeting of the American Education Research Association.

Hambleton R. & Rodgers J. (2005). Item Bias Review. Practical Assessment, Research, and Evaluation, 4(6). Retrieved March 18, 2009, from http://PAREonline.net/getvn.asp?v=4andn=6.

Hambleton R. & Rodgers J. (2005). Item Bias Review. Practical Assessment, Research Evaluation, 4(6). Retrieved November 18, 2006, from http://PAREonline.net/getvn.asp?v=4&n=6

Hambleton R. K. & Swaminathan H. (2005). Item Response Theory: Principle and Application. Boston: Kluwer.

Hambleton R. K., Swaminathan H. & Rogers H. J. (2001). Fundamentals of Item Response Theory. Newbury Park, California: Sage Publications.

Hirschfeld M., Moore R. L. & Brown E. (2005). Exploring the Gender Gap on the GRE Subject Test in Economics. *Journal of Economic Education, 26, Pp. 316*

Kimball M. M. (2009). A New Perspective on Women's Math Achievement. Psychol. Bull. 105: 198-214.

Kolcic I. O. P. (2006). Academic Performance and Scientific Involvement of Final Year Medical Students coming from Urban and Rural

Background of Andjija Stampar School of Public Health, Medical School, university of Zagreb, Croatia htt://rrh. Deakin.edu.au/

Kunnan A. J. (2000). Fairness and Justice for All. In A. J. Kunnan (Ed.), Fairness and Validation in Language Assessment: Selected Papers from the 19th Language Testing

Lee J. & Mclritire W. (2001). Interstate Variation in the Mathematics Achievement of Rural and Non-Rural Students. *Journal of Research in Rural Education,* 16 (3)

Nworgu B. G. (2011). Differential Item Functioning: A Critical Issue in Regional Quality Assurance. Paper Presented in NAERA Conference.

Nworgu B. G. & Odili J. N. (2005). Analysis of DIF in WAEC SSCE Biology Multiple Choice Test. Review of Education 16, (2) 140-152.

Ogbebor U. C. & Onuka, E. O. (2012). Differential Item Functioning Economics Question Paper of National Examinations Council in Delta State Nigeria. Unpublished M.Ed Thesis. University of Ibadan.

Ogbebor U. & Onuka A. (2013) Differential Item Functioning Method as an Item Bias Indicator. Educational Research. 4(4), 367-373,

Owoeye J. S. (2000). The Effect of Integration of Location, Facilities and Class Size on Academic Achievement of Secondary School Students in Ekiti State, Nigeria. Ph.D Thesis, Unpublished. University of Ibadan

Pedrajita J. Q. (2009). Using Logistic Regression to Detect Biased Test Items. *The International of Educational and Psychological Assessment Journal*

Perrone M. (2006). Differential Item Functioning and Item Bias: Critical Consideration in Test Fairness. Applied Linguistics, 6(2): 1-3.

Robert, V. K. & Morgan, D. W. (1971), Determining Sample Size for Research Activities. *Journal of Educational and Psychological Measurement, 30, 607-610*

Roever C. (2005). "That's Not Fair!" Fairness, Bias, and Differential Item Functioning in Language Testing. Retrieved June 18, 2016, from the University of Hawai'i System Web site: http://www2.hawaii.edu/~roever/brownbag.pdf

Schumacker R. (2005). Test Bias and Differential Item Functioning. Retrieved November 18, 2006, from http://www.appliedmeasurementassociates.com/White%20Papers/TE ST%20BIAS%20AND%20DIFFERENTIAL%20ITEM%20FUNCTIONI

NG.pdf

Shohamy E. (2000). Fairness in Language Testing. In A. J. Kunnan (Ed.), Fairness and Validation in Language Assessment: Selected Papers from the 19th Language Testing

Tremblay S., Ross N. A. & Berthelot J. M. (2001). "Individual and Community Factors Affecting Grade Three Academic Performance: A Multi-Level Analysis." *Education Quarterly Review*, 7 (4), 25-36.

Williams M. L., Walaver C. & Duggal V. (2002). Gender Differences In Economic. An Extension of the Analysis. Journal of Economic Education. 23(3)

Williams V. (2007). The "Unbiased" Anchor: Bridging the Gap between DIF and Item Bias. Applied Measurement in Education, 10, 253-267.

Zumbo B. D. (2009). A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-Like (ordinal) Item Scores. Ottawa, Canada: Directorate of Human Resources Research and Evaluation.

Reference to this paper should be made as follows: Kelechi Joshua Princewill Ihechu & Adolphus Onuabuchi Madu (2016), Differential Item Functioning of National Examinations Council (Neco) Agricultural Science Questions For 2015. *J. of Education and Leadership Development Vol. 8, No. 2, Pp. 1 – 17.*