DIFFERENTIAL ITEM FUNCTIONING OF GENERAL STUDIES MATHEMATICS EXAMINATION IN GOMBE STATE POLYTECHNIC BAJOGA, NIGERIA

By

UMAR YAHAYA UMAR
Department of Educational Foundations, Faculty Technology Education
Abubakar Tafawa Balewa University, Bauchi, Bauchi State Nigeria
Email: uyahaya77@gmail.com,

KAKU ZAKKA
Department of Education
University of Maiduguri, Borno State Nigeria
Email:kakuzakka@gmail.com

and

ABUBAKAR MOHAMMED
Department of Educational Foundations, Faculty Education
Federal University Kashere, Gombe State
Email: abusaabiiq@gmail.com

Abstract
The study investigated the incidence of Differential Item Functioning (DIF) of General Studies Mathematics Examinations of 2017/2018, 2018/2019 and 2019/2020 by gender in Gombe State Polytechnic Bajoga. The study was guided by two (3) research questions and one (1) hypothesis. An ex-post facto research design was employed for the study. The population of the study consisted of one thousand two hundred and twenty four (1,224) students. All the populations were used as a sample. The research questions were answered using Binary Logistic Regression analysis. The hypothesis formulated was tested using Analysis of Variance (ANOVA). The findings of the research questions one, two and three revealed that 34 items functioned differently by gender on General Studies Mathematics Examinations of 2017/2018, 2018/2019 and 2019/2020 respectively. Furthermore, the finding of the research hypothesis revealed that there is no significant difference among General Studies MathematicsExaminations in 2017/2018, 2018/2019 and 2019/2020 with regards to gender.In view of the findings, the study recommended that, the institution should either employ test experts or pilot test the items to eliminate bias and unreliable items before administration to the examinees.

Keywords: Differential Item Functioning, General Studies, Mathematics and Gender

Introduction

Generally, it is understood that in the field of education only educational achievements of the students are measured, and are expressed in terms of marks. However, in present times various variables such as intelligence, interest, aptitude, attitude, personality of the students, usefulness of the curriculum, teaching methods, educational policy, educational activities of the person concerned such as administrators, teachers, Guardians are also measured.

Every aspect of our life is touched by measurement in it numerous factors. For instance, the date and time of birth of a baby, his weight is measured by a balance, he grows in height and weight from year to year, his intelligence, achievement and many other traits are measured in his school, and he assesses the prospects of different occupation before making his own choice. Measurement is the act or process of ascertaining the quality or extent of something. It is an amount of something which has been counted, measured or assessed. The set of numbers used depends upon the nature of the characteristics being measured and the type of measuring instruments used. The set of objects or individual is determined by the purposes for which the measurements are to be made (Sidhu, 2012).

Differential Item Functioning (DIF) is the simple and idealized situation in which the response (Y) to a particular item on the test is determined solely by the target proficiency intended to be measured by the test, denoted here by θ, and random error. We can express the probability distribution of Y conditional on θ by $f\ (Y \mid \theta)$. Differential Item Functioning (DIF) as a statistical procedure that checks whether examinees with comparable total test scores belonging to different groups answer similarly the individual items of the test. Differential Item Functioning (DIF) analyses it is usual that there are, at least, two groups of interest: the focal group and the reference group. The former generally refers to a minority or traditionally considered disadvantaged group, while the latter is the majority or privileged group (Wu &Ercikan, 2006).

Differential Item Functioning (DIF) refers to the differences in item functioning aftergroups have been matched with respect to their ability. It occurs when test items function differently for students from two different

comparison groups that are matched by the construct being measured. Differential item functioning refers to the violation of the invariance assumption in Item Response Theory Models, and happens when the probability of endorsing an item for test takers of equal ability level varies in different groups (Battuaz, 2017).

Traditional Differential Item Functioning analyses typically contrast the performance of two groups that differ in race, ethnicity, gender, family income, first language spoken at home, or other similar demographic variables (Lan Yu, 2006).Differential Item Functioning occurs when a test or a survey item functions differently for a reference group (e.g whites) of examinees or respondents compared to a focal group (Blacks) of examinees or respondents, after controlling for the level of the attribute being measured (Dodeen& Johnson, 2003; Kamata& Vaughn, 2004). For example, an item exhibit Differential Item Functioning if the probability of white responding to a specific category differs from Blacks when they both are operating at the same overall level on the construct.

Differential item functioning is the tendency of the testee standing in the latent trait to perform differently in a test item. According to Dogan (2005), Differential item functioning occurs when test-takers form different subgroups with identical overall test score or ability level, differ systematically with regard to the probability of solving particular test items whose sources of extraneous error are not checked will likely function differently for subgroups with identical overall test scores.

Uniform DIF is a form of DIF which occur when a sub-group of examinees with ability levels, uniformly answer a particular item or subset of items than the other group, that particular sub-group is said to be advantaged over the other group and can be considered as having a superior ability over the less favoured group and is termed as "reference" group, while the less (inferior group) advantage group is the "focal" group in bias analysis comparatively (Walker, 2011). While **Non-uniform DIF** it occur when there is an interaction between the higher ability and lower ability group for an item seem to be difficult for those at higher level in one group, and more difficult for those at lower level in the other group.

General Studies is a broad curriculum meant to help aspirants sharpen their general knowledge, communication skills, and people skills. These courses and the exam are your gateways to many institutions of higher learning in most countries. Initially, General studies may seem

confusing, but you'll find they offer a lot of benefits. Many universities and polytechnics emphasize the overall improvement of students' abilities by making sure that they have a firm grasp of how to do well in various subject areas. The course is knowledge-orientated and aims to broaden thinking and increase curiosity. That makes it suitable for students who want to find out more about the world around them. The components of general studies include social sciences, law, finance, the humanities, Mathematics and other multidisciplinary subjects. It also combines methods and concepts from many different fields, like sociology, health sciences, geography, history, and so on. General studies teach students about the "big picture" of knowledge and how different disciplines are connected unlike traditional disciplines. General studies focus less on highly specific or technical knowledge and more on teaching students about the world as a whole.

Mathematics as the mother of science is the study of size, numbers and patterns. Mathematics is the most international of all subjects, and mathematical understanding influence decision making in all areas of life; being it private, social and civil (Anthony &Walshaw, 2009). It is the subject that enables scientist and technologist to develop relationships among biological, chemical, geographical and physical qualities; understanding and explain natural phenomena. The knowledge of mathematics is an essential tool in any society today. The contribution that mathematical knowledge and skills has made to economics, industrial and technological growths of modern world are quite obvious to almost everyone.

Mathematics has helped to transform mans` rural society to modern society. Mathematics is the foundation of science which is the bedrock of modern development. It is well known that the level of social and economic development of any country is intimately connected with the level of development of that country in science and technology. This by implication means that, the level of social and economic development is closely connected with the level of development in the mathematical science (Kuku, 2012).

The reason of the study arises as a result of massive poor performance of the students in mathematics across all the discipline in the institution. For instance, in Science Laboratory Technology (SLT), about 60.1% fail mathematics and in Computer Science Programme (CSP)

60.1% fail mathematics Public Administration (PA) 61.9% fail mathematics, Business Administration and Management (BAM) 62.2% fail mathematics, Office Technology and Management (OTP) 47.1% fail mathematics and Accounting option (ACC) 70.7% fail mathematics respectively (GSPB 2018).

The following questions and hypothesis guided the study

1. Which of the items in General Studies Mathematics examination functions differently by Gender in 2017/2018 academic session
2. Which of the items in General Studies Mathematics examination functions differently by Gender in 2018/2019 academic session
3. Which of the items in GNS Mathematics examination functions differently by Gender in 2019/2020 academic session?

   H_{01}: There is no significant mean difference in DIF General Studies Mathematics in 2017/2018, 2018/2019 and 2019/2020 academic session across Gender

Methods

The study employed an Ex-Post Facto research design. An ex-post-facto design sought to find out the factors that are associated with certain occurrences of already existing condition or state of affairs and searching back in time for plausible causal factors retrospective (Cohen, Lawrence & Morrison, 2007). The population of the study comprises all Pre-National Diploma students of Gombe State Polytechnic Bajoga. The total population of the examinees for both the pure science and social science are 1,224 students 2017/2018, 2018/2019 and 2019/2020 session (Gombe State Polytechnic Bajoga 2018). The entire population was used as a sample size. The researcher used the whole total population of the students to serve as a sampling technique (TPS). This is because the size of the population can be manageable by the researcher. The constructed question paper of Gombe State Polytechnic Bajogawas used as an instrument for data collection. Logistic Regression (LR) and ANOVA method was employed using SPSS for data analysis. Logistic Regression has been chosen due to the fact that, it is used to analyze dichotomous data such as 1 for right and 0 for wrong response and chi-square test to evaluate the level of significance of items exhibiting DIF, frequency cross tabulation was also used in identifying the percentage of correct response for each

item that functioned differently by gender and ANOVA to determine the differences between the years of examination.

Result

Research Question One: Which of the items in General Studies Mathematics examination functions differently by Gender in 2017/2018 academic session?

Table 1

Summary of Binary Logistic Regression Analysis of 2017/2018, GNS Mathematics Examination by Gender

Item No.	ΔR^2	Sig.	Remarks	Item No.	ΔR^2	Sig.	Remarks
1	0.001	0.683	No DIF	11	0.813**	0.000	DIF
2	0.007	0.125	No DIF	12	0.003	0.289	No DIF
3	0.061*	0.000	DIF	13	0.001	0.622	No DIF
4	0.000	0.686	No DIF	14	0.000	0.907	No DIF
5	0.006	0.157	No DIF	15	0.082**	0.000	DIF
6	0.000	0.788	No DIF	16	0.097**	0.000	DIF
7	0.003	0.304	No DIF	17	0.074**	0.000	DIF
8	0.008	0.129	No DIF	18	0.006	0.183	No DIF
9	0.094**	0.000	DIF	19	0.001	0.571	No DIF
10	0.002	0.421	No DIF	20	0.084**	0.000	DIF

Negligible DIF: $\Delta R^2 < 0.035$; *Moderate DIF: $0.035 \leq \Delta R^2 < 0.07$,**Large DIF: $\Delta R^2 \geq 0.07$

From the above table, the items that functioned differently with regards to gender in GNS Mathematics examination of 2017/2018 academic session are items 3, 9, 11,15, 16, 17 and 20 with significant values of 0.000, 0.000, 0.000, 0.000, 0.000, 0.000 and 0.000 respectively.

Research Question Two: Which of the items in General Studies Mathematics examination functions differently by Gender in 2018/2019 academic session?

Table 2: *Summary of Binary Logistic Regression Analysis of 2018/2019, GNS Mathematics by Gender*

Item No.	?R^2	Sig.	Re-marks	Item No.	?R^2	Sig.	Re-marks	Item No.	?R^2	Sig.	Re-marks
1	0.008	0.341	No DIF	26	0.035*	0.027	DIF	51	0.067*	0.000	DIF
2	0.044*	0.020	DIF	27	0.028	0.099	No DIF	52	0.007	0.397	No DIF
3	0.018	0.161	No DIF	28	0.015	0.218	No DIF	53	0.048*	0.003	DIF
4	0.084	0.000	No DIF	29	0.097**	0.000	DIF	54	0.001	0.792	No DIF
5	0.035*	0.037	DIF	30	0.003	0.600	No DIF	55	0.021	0.129	No DIF
6	0.001	0.688	No DIF	31	0.000	0.908	No DIF	56	0.002	0.604	No DIF
7	0.029	0.069	No DIF	32	0.009	0.336	No DIF	57	0.005	0.454	No DIF
8	0.000	0.910	No DIF	33	0.002	0.600	No DIF	58	0.024	0.101	No DIF
9	0.003	0.541	No DIF	34	0.015	0.208	No DIF	59	0.033	0.060	No DIF
10	0.008	0.358	No DIF	35	0.001	0.808	No DIF	60	0.040*	0.009	DIF
11	0.000	0.929	No DIF	36	0.001	0.760	No DIF	61	0.003	0.575	No DIF
12	0.000	0.850	No DIF	37	0.145**	0.000	No DIF	62	0.027	0.141	No DIF
13	0.004	0.482	No DIF	38	0.000	0.894	No DIF	63	0.076**	0.000	DIF
14	0.004	0.543	No DIF	39	0.052*	0.006	No DIF	64	0.099**	0.000	DIF
15	0.000	0.964	No DIF	40	0.032	0.059	No DIF	65	0.004	0.528	No DIF
16	0.002	0.664	No DIF	41	0.124**	0.000	No DIF	66	0.001	0.753	No DIF
17	0.022	0.115	No DIF	42	0.001	0.760	No DIF	67	0.001	0.797	No DIF
18	0.057*	0.002	DIF	43	0.026	0.094	No DIF	68	0.022	0.129	No DIF
19	0.011	0.283	No DIF	44	0.001	0.770	No DIF	69	0.000	0.922	No DIF
20	0.012	0.256	No DIF	45	0.052*	0.000	No DIF	70	0.011	0.262	No DIF
21	0.012	0.257	No DIF	46	0.086**	0.000	No DIF				
22	0.076**	0.000	DIF	47	0.004	0.528	No DIF				
23	0.025	0.069	No DIF	48	0.032	0.062	No DIF				
24	0.001	0.802	No DIF	49	0.101**	0.000	No DIF				
25	0.005	0.447	No DIF	50	0.011	0.304	No DIF				

Negligible DIF: $\Delta R^2 < 0.035$; *Moderate DIF: $0.035 \leq \Delta R^2 < 0.07$,**Large DIF: $\Delta R^2 \geq 0.07$

In the above table, the items that functioned differently with regards to gender in GNS Mathematics examination of 2018/2019 academic session are items 2, 5,18, 22, 26, 29, 37, 39, 41, 45, 46, 49, 51, 53, 60, 63 and 64 with significant values of 0.020, 0.037, 0.002, 0.000, 0.027, 0.000, 0.000, 0.006, 0.000, 0.000, 0.000, 0.000, 0.000, 0.000, 0.003, 0.009, 0.000, and 0.000 respectively.

Research Question three: Which of the items in GNS Mathematics examination functions differently by Gender in 2019/2020 academic session?

Table 3

Summary of Binary Logistic Regression Analysis of 2019/2020, GNS Mathematics Examination by Gender

Item No.	ΔR^2	Sig.	Remarks	Item No.	ΔR^2	Sig.	Re-marks
1	0.019	0.531	No DIF	16	0.001	0.918	No DIF
2	0.187**	0.000	DIF	17	0.075**	0.000	DIF
3	0.003	0.823	No DIF	18	0.019	0.531	No DIF
4	0.034	0.406	No DIF	19	0.034	0.406	No DIF
5	0.098**	0.000	DIF	20	0.102**	0.000	DIF
6	0.002	0.830	No DIF	21	0.001	0.918	No DIF
7	0.267**	0.000	DIF	22	0.001	0.918	No DIF
8	0.013	0.603	No DIF	23	0.000	1.000	No DIF
9	0.001	0.894	No DIF	24	0.001	0.918	No DIF
10	0.020	0.543	No DIF	25	0.034	0.406	No DIF
11	0.003	0.823	No DIF	26	0.001	0.918	No DIF
12	0.136**	0.000	DIF	27	0.060*	0.004	DIF
13	0.011	0.645	No DIF	28	0.005	0.757	No DIF
14	0.236**	0.000	DIF	29	0.074**	0.000	DIF
15	0.003	0.823	No DIF	30	0.027	0.461	No DIF

Negligible DIF: $\Delta R^2 < 0.035$; *Moderate DIF: $0.035 \leq \Delta R^2 < 0.07$,**Large DIF: $\Delta R^2 \geq 0.07$

The outputs of the analysis in the table present the Nagelkerke R Square (ΔR^2) and the significant values respectively. The decision on how the DIF is classified was based on the categorization by Jodoin and Gierl (2001) of effect levels determined with logistic regression in the following ways: R < .035, a negligible level of DIF is present; 036 < R < .070, a medium level of DIF is present; R > .071, a magnitude level of DIF is present. It can be observed that, the items that functioned differently with regards to gender in GNS Mathematics examination of 2019/2020 academic session are items 2, 5, 7, 12, 14, 17, 20, 27 and 29 with significant values of 0.000, 0.000, 0.000, 0.000, 0.000, 0.000, 0.000, 0.004 and 0.000 respectively.

H_{O1}: There is no significant mean difference in DIF GNS Mathematics in 2017/2018, 2018/2019 and 2019/2020 academic session across Gender

Table 4

Result of ANOVA on Mean DIF Indices of GNS Mathematics Examinations in 2017/2018, 2018/2019 and 2019/2020 Academic Session by Gender

Source	SS	df	MS	F	Sig.
Between Groups	.030	2	.015	2.128	.124
Within Groups	.837	117	.007		
Total	.868	119			

The outcome from the analysis shows that the mean DIF indices of GNS Mathematics examinations in 2017/2018, 2018/2019 and 2019/2020 academic session by Gender do not differ significantly. The F value was found to be 2.128 with significant value (p-value) of .124 respectively. The null hypothesis stated above is thereby retained, meaning there is no significant difference among GNS MathematicsExaminations in 2017/2018, 2018/2019 and 2019/2020 with regards to gender in their respective mean DIF indices.

Discussion

The findings of this study agreed with the findings of previous studies in terms of DIF significant and insignificant in Gender. For instance.Yousif (2011) where they found notable gender differences for item associated with specific passages, reporting that females scored higher than males with humanities-oriented reading passages, but lower than male in science-oriented passages. Abduaziz (2010) where he indicated that females showed a statistical significant and consistent advantage over males on numerical ability, whereas males showed a consistent advantage over females on spatial and deductive abilities. Similarly, the finding is also in accord with the work of Madu and Bassey (2010), analyzed mathematics multiple choice test items set and administered by National Examination Council (NECO) for differential item functioning by gender whereby the analysis indicates that there are some differences; item in Algebra, Geometry and Trigonometry number, Numeracy and Problem-solving seems to favour males while items in statistics and probability seems to favour females. Mado (2011), where the

DIF analysis flagged some items as having gender bias in multiple-choice items in mathematics of some public examinations.Similarly, pae (2011) revealed that items 20 and 68 seemed to be more difficult for females, while items 39, 78, and 86 were more difficult for the males. Although there were two outliers (items 38 and 78), it is hard to assume a test bias because the source of the outlying position is unknown. Furthermore, Abba (2014) found that items in English language multiple choice were flagged as having gender, location and school type bias which significantly and consistently favoured male and female students, Urban student than Rural, day than boarding students. Shanmugan (2018) studied on determining gender differential item functioning for mathematics in coeducational school culture and found two moderate DIF items that assess subtraction favoured girls.

Conclusion

It could be recall that this research work aimed at finding out whether General Studies mathematics 2018, 2019 and 2020 multiple choice items function differently by gender in Gombe state polytechnic Bajoga. Therefore, the findings of this study indicated uniform and non-uniform DIF. For instance, when a sub-group of examinees with ability levels, uniformly answer a particular item or subset of items than the other group, that particular sub-group is said to be advantaged over the other group and can be considered as having a superior ability over the less favoured group and is termed as "reference" group, while the less (inferior group) advantage group is the "focal" group in bias analysis comparatively. While Non-uniform DIF it occur when there is an interaction between the higher ability and lower ability group for an item seem to be difficult for those at higher level in one group, and more difficult for those at lower level in the other group.

Recommendation from the Study

1. The study recommended that the institution should either employ test experts or pilot test the items so as to eliminate bias and unreliable items before administration to the examinees.
2. Mathematics text books should be in abundance in the school library
3. The study also recommended that educational administrators, academic planners, and admission officers to carefully admit students

according to their programme of choice so as to create interest in the students.

Reference

Abba, B. (2014). Analysis of differential item functioning of senior secondary certificate (NECO) English examination in Dawakin Kudu educational zone Kano State.Unpublished M.Ed thesis.Bayero University Kano.

Abdulaziz, M. A. (2010). Prevalence of refractive errors among pre-school children at King Abdulaziz Medical City, Riyadh, Saudi Arabia, *Saudi Journal of Ophthalmology 24(2)*, 45-8: DOI: 10.1016/j.sjopt.2010.01.001

Anthony, G. S., Walshow, M. (2009).Effective pedagogy in mathematics.BIE education series 9. International bureau of education

Battuaz, M. (2017). On World's Test on Differential Item Functioning Detection Methods. Retrieved from http://onworld's test on DIF

Cohen, L., Lawrence, M., & Morrison, K. (2007).Research Method in Education.*British Journal of Educational Studies 55* (4)

Dogan, N. (2005). Comparing differential item functioning based on manifest groups and latent classes. Educational science: theory and practical. Retrieved from www.estp.com.tr

Dodeen, H. & Johnson, G. A. (2003).An analysis of sex-related differential item functioning in attitude assessment.*Assessment and evaluation in higher education,* 28(2) 129-134

Gombe State Polytechnic Bajoga (2019). News Bulletin: 2 (1)

Kuku, A. O. (2012). Mathematics as a time-tested resource for scientific, technological, socio-economic and intellectual development. Lecture notes delivered at the University of Ibadan, Ibadan University Press.

Lan, Y., Chen, S. Y., & Lei, P. W. (2006). Comparing methods of assessing differential item functioning in a computerized adaptive testing environment.*Journal of Educational Measurement*, *43*(3), 245-264. https://doi.org/10.1111/j.1745-3984.2006.00015.

Madu, B.C. &Bassey, V. I. (2010). Gender differential item functioning in mathematics multiple choice test items set and administered by National Examination Council (NECO). *Journal of mathematical sciences and education* 1(1) 101-114

Pae, H. (2011). Differential Item Functioning and Unidimensionality in the Pearson Test of English. USA: University of Cincinnati

Shanmugan, K. S. (2018). Determining gender differential item functioning for mathematics in coeducational school culture.*Malaysian journal of learning and instruction,* 15(2) 83-109

Sidhu, S. S. (2012). New approaches to measurement and evaluation. New Delhi: Sterling Publishers Pvt Ltd.

Walker, (2011).The impact of management on administrative and survey measures of organizational performance.Retrieved from DOI 10.1080/14719037.2010.532968

Wu A.D, &Ercikan, k. (2006). "using multiple durable matching to identify cultural sources of DIF". *International journal of testing*, 6(3) 287-300.

Yousif, A.N. (2011). Comprehension Performance of Gender-Neutral Texts. Retrieved from *faculty.ksu.edu.sa/yousif/publication/microsoftword-Gender Difference in Reading Comprehension.pdf* on 28-06-2012

ASSEREN Journal of Educational Research and Development (AJERD)
AJERD Assessment Form