

---

**ANALYSIS OF GENDER-RELATED DIFFERENTIAL ITEM FUNCTIONING OF 2021 NECO  
BECE MATHEMATICS MULTIPLE-CHOICE ITEMS IN NIGERIA**

**OLUTOLA Adekunle Thomas PhD, ALI Binta Ado PhD, APENANIBO Terkaa Jude**  
aolutola@fudutsinma.edu.ng; baali@fudutsinma.edu.ng; apenat.jude@fudutsinma.edu.ng  
Department of Educational Psychology and Counselling  
Faculty of Education, Federal University Dutsin-Ma, Katsina State

**Abstract**

*This study examined the Gender-related Differential Item Functioning of 2021 NECO BECE Mathematics Multiple-Choice Items in Nigeria. The survey research design was adopted. The target population comprised five thousand one hundred and seventy-eight (5,178) junior secondary school three (JSS 3) students in Federal Unity Colleges in North Central, Nigeria. A sample of one thousand two hundred and seventy-five (1,275) was drawn through proportionate stratified random sampling. The 2021 NECO (BECE) Mathematics multiple-choice test served as the instrument for data collection. The instrument was not validated based on it's a standardized test developed by National Examinations Council (NECO). The instrument was administered to students in a school which is not part of the sample. The 2021 NECO (BECE) Mathematics multiple-choice test was the primary instrument. Content validity was affirmed by mathematics education experts, and reliability for the study sample was established with a Cronbach's alpha of 0.78. Data obtained for the study were analyzed using Binary Logistic Regression, chosen for its capacity to detect both uniform and non-uniform DIF and provide interpretable odds ratios. Findings revealed that three (3) items functioned differentially in favour of male candidates and one (1) item favour female candidates, which show that, there is significant difference in the items which functioned differentially in the 2021 Mathematics NECO (BECE) multiple choice items in North-central, Nigeria based on gender. Based on the finding of the study conclusions were made and appropriate recommendations were made.*

**Keywords:** Differential Item Functioning, Gender, Mathematics Achievement, NECO, BECE, Multiple-Choice Test

---

**Citation:** Olutola, A. T., Ali, B. A., Apenanibo, T. J. (2025). Analysis of Gender-Related Differential Item Functioning of 2021 NECO BECE Mathematics Multiple-Choice Items in Nigeria. *Confluence Journal of Education and Knowledge Management*, 2992-2518 (Print) 2992-2526 (Online), Volume 2 (2), pages: 131 - 140. DOI:<https://doi.org/10.5281/zenodo.19098496>

---

**Introduction**

Testing remains one of the most important processes in educational assessment, providing the basis for decisions about students' placement, progression, and certification. Among the different forms of assessment is multiple-choice tests item that are particularly valued for their objectivity, efficiency in covering wide curriculum content, and ease of scoring (Aminu & Okonkwo, 2021). In Nigeria, the Basic Education Certificate Examination (BECE), conducted by the National Examinations Council (NECO), is a key assessment at the end of basic schooling, determining students' educational achievement and readiness for senior secondary education.

However, for such high-stakes examinations to be considered fair and valid, the test items must be free from biases and measure the intended construct across all groups, irrespective of irrelevant personal attributes such as gender, school location, or school type (Obioma & Salau, 2020; Queen soap & Orluwene, 2019). When items behave differentially for subgroups of examinees who possess the same underlying ability is referred to as Differential Item Functioning (DIF) (Zhang & Stout, 2016; Boone, Staver & Yale 2020). Olutola and Nuraddeen (2025) opines that, DIF occurs when examinees from

different groups of the same ability show different probabilities of success on the item after matching on the underlying ability that the item is intended to measure.

Gender is an important variable in the school system. Gender is one of the most extensively studied variables in DIF research (Olutola, Ihechu & Nuraddeen 2022). Numerous studies have shown mixed findings regarding gender differences in mathematics performance. Some researchers have reported that males outperformed females in certain mathematics domains (Okeke & Nwankwo, 2018; Akanwa, Ihechu & Nkwocha, 2022), while others found no significant difference or even female advantage in specific cognitive tasks (Ezenwa, Nwankwo & Eze, 2021; Adewale & Yusuf, 2019). These inconsistencies underline the importance of evaluating test fairness at the item level rather than relying solely on overall scores.

In Nigeria and other developing countries, both genders can influence students' attitudes towards mathematics, potentially leading to test bias. While NECO strives to ensure fairness, empirical evidence is required to verify that its items function equally well for both male and female candidates. High-stakes examinations such as the National Examinations Council BECE are expected to provide an unbiased measure of students' academic ability. However, several studies in Nigeria and elsewhere have raised concerns about potential item bias in standardized mathematics tests (Ihechu & Madu, 2016). When DIF exists in an examination, candidates of equal ability from different gender groups may have unequal chances of answering items correctly, thereby questioning the fairness and validity of the test.

Gender-related DIF is particularly important in mathematics because of the persistent debate over gender gaps in STEM fields (UNESCO, 2022). Although previous studies have examined DIF in BECE and other national examinations, most have focused on earlier years or combined multiple variables such as school location and type, leaving limited recent evidence on gender DIF in NECO (BECE) Mathematics.

High-stakes standardized tests like the NECO (BECE) are pivotal in determining the educational and professional trajectories of Nigerian students. For these examinations to be valid and equitable, they must provide a fair and unbiased measure of academic ability for all candidates, irrespective of gender. Despite the National Examinations Council's (NECO) commitment to fairness, the potential for test bias, particularly Differential Item Functioning (DIF) related to gender, remains a significant concern.

The existence of DIF in mathematics tests is especially critical due to mathematics' role as a gateway subject for Science, Technology, Engineering, and Mathematics (STEM) careers. Persistent socio-cultural beliefs and gender stereotypes in Nigeria can influence students' engagement and performance in mathematics, potentially leading to biased test items that systematically favour one gender over another (Bello & Olatunji, 2023). When items with DIF are present, they compromise the validity of the examination, as students of equal ability do not have an equal probability of success, thereby undermining the fairness of the entire assessment system. While previous studies have examined DIF in Nigerian examinations, there is a scarcity of recent, focused evidence on gender-based DIF in the NECO (BECE) Mathematics test. For instance, studies like Madu (2012) and Abedalaziz (2010) found evidence of DIF favouring males, suggesting a persistent issue. However, other researchers, such as Yusuf and Dauda (2022) and Lawal and Salami (2023), reported no significant gender DIF, creating an inconsistent and inconclusive body of literature. This inconsistency underscores the need for continuous, test-specific DIF analysis.

The core problem this study addresses is the lack of contemporary, empirical evidence on the fairness of the 2021 NECO (BECE) Mathematics test specifically concerning gender. This gap is critical given the high-stakes nature of the examination and the ongoing concern about gender equity in STEM education in Nigeria. Crucially, a recent and comprehensive study by Olutola (2021) on the 2019 NECO Mathematics objectives test found that 12 out of 60 items exhibited significant DIF, with a majority favouring male students. This finding directly highlights that gender bias is not a hypothetical concern but a documented

issue in recent NECO mathematics papers. It is therefore imperative to investigate whether this pattern of bias persisted in the subsequent 2021 examination.

Therefore, this study is necessitated by the critical need to provide empirical, up-to-date evidence on the fairness of the 2021 NECO (BECE) Mathematics items. The problem this study investigates is whether items in the 2021 examination function differentially for male and female students of equal mathematical ability, thereby identifying potential bias and contributing to the improvement of test quality in Nigeria.

### **Purpose of the Study**

The main purpose of this study was to examine analysis of gender-related differential item functioning of 2021 NECO (BECE) Mathematics multiple-choice items in Nigeria.

Specifically, the study seeks to:

- i. determine the items in the 2021 National Examinations Council Basic Education Certificate Examination Multiple-Choice Mathematics items that functioned differentially by gender.

### **Research Hypothesis**

The following research hypothesis is formulated to guide the study:

Ho<sup>1</sup>: There is no significant difference in the items that functioned differentially in the 2021

National Examinations Council Basic Education Certificate Examination multiple choice Mathematics examination paper based on gender.

### **Methodology**

The survey research design was adopted. The target population comprised five thousand one hundred and seventy-eight (5,178) junior secondary school three (JSS 3) students in Federal Unity Colleges in North Central, Nigeria. The sample comprised one thousand two hundred and seventy-five (1,275) Junior Secondary School Three (JSS 3) students drawn through proportionate stratified random sampling. The strata were defined by the six Federal Unity Colleges in North Central, Nigeria. The sample consisted of six hundred and seventy-five (675) male and six hundred female students, ensuring proportional gender representation from each school.

The 2021 NECO (BECE) Mathematics multiple-choice test served as the instrument for data collection. As a standardized national examination, its content validity is established. For this study, its contextual appropriateness was confirmed by two experts in mathematics education.

The instrument was pilot-tested with fifty (50) JSS3 students from a non-participating school to check for administrative clarity. The data obtained was analyzed using Cronbach's formula with the reliability index of 0.78. This reliability was computed on the current study's sample, indicating good internal consistency. Data obtained for the study were analyzed using Binary Logistic Regression (BLR) to detect items functioning differentially by gender. Binary Logistic Regression (BLR) was selected for its robustness in modeling both uniform and non-uniform.

### **Results**

**Hypothesis 1:** There is no significant difference in the items that functioned differentially in the 2021 National Examinations Council Basic Education Certificate Examination multiple choice Mathematics examination paper based on gender.

**Table 1: Binary Logistic Regression of DIF by Gender for 2021 Mathematics NECO (BECE) multiple-choice Items**

Item	B	S.E	Wald	Sig	Exp (B)	95% C.I for Exp (B)		Decision
						Lower	Upper	
1.	.192	.122	2.484	.115	1.212	.954	1.538	NO DIF
2.	-.142	.124	1.311	.252	.867	.680	1.107	NO DIF
3.	.138	.117	1.406	.236	1.148	.914	1.443	NO DIF
4.	.084	.112	.558	.455	1.088	.873	1.356	NO DIF
5.	.026	.113	.051	.821	1.026	.822	1.282	NO DIF
6.	-.052	.115	.210	.647	.949	.758	1.188	NO DIF
7.	.128	.112	1.289	.256	1.136	.912	1.416	NO DIF
8.	-.167	.113	2.179	.140	.846	.678	1.056	NO DIF
9.	-.069	.112	.379	.538	.933	.749	1.163	NO DIF
10.	-.091	.114	.642	.423	.913	.731	1.141	NO DIF
11.	.103	.114	.819	.365	1.109	.887	1.386	NO DIF
12.	.176	.113	2.438	.118	1.192	.956	1.487	NO DIF
13.	.026	.113	.054	.817	1.026	.823	1.280	NO DIF
14.	-.034	.113	.093	.761	.966	.775	1.205	NO DIF
15.	.064	.112	.324	.569	1.066	.855	1.329	NO DIF
16.	.007	.119	.003	.955	1.007	.797	1.271	NO DIF
17.	.211	.116	3.319	.069	1.235	.984	1.551	NO DIF
18.	.074	.112	.428	.513	1.076	.864	1.341	NO DIF
19.	.052	.122	.185	.667	1.054	.830	1.338	NO DIF
20.	.151	.115	1.724	.189	1.163	.928	1.458	NO DIF
21.	.086	.113	.583	.445	1.090	.874	1.360	NO DIF
22.	-.106	.121	.766	.381	.900	.710	1.140	NO DIF
23.	-.120	.113	1.124	.289	.887	.711	1.107	NO DIF
24.	-.158	.113	1.935	.164	.854	.684	1.067	NO DIF
25.	.003	.112	.001	.979	1.003	.805	1.250	NO DIF
26.	-.116	.112	1.074	.300	.890	.714	1.109	NO DIF
27.	.013	.112	.013	.911	1.013	.813	1.262	NO DIF
28.	.006	.112	.003	.958	1.006	.807	1.254	NO DIF
29.	.054	.113	.234	.629	1.056	.847	1.317	NO DIF
30.	-.192	.113	2.905	.088	.825	.662	1.029	NO DIF
31.	.106	.114	.876	.349	1.112	.890	1.390	NO DIF
32.	-.281	.113	6.174	<b>.013*</b>	.755	.605	.942	DIF
33.	-.102	.113	.822	.365	.903	.724	1.126	NO DIF
34.	-.210	.113	3.469	.063	.811	.650	1.011	NO DIF

35.	-.024	.112	.045	.833	.977	.784	1.217	NO DIF
36.	.165	.123	1.786	.181	1.179	.926	1.501	NO DIF
37.	.190	.118	2.572	.109	1.209	.959	1.525	NO DIF
38.	.199	.120	2.740	.098	1.220	.964	1.545	NO DIF
39.	.031	.117	.071	.789	1.032	.821	1.296	NO DIF
40.	-.036	.112	.100	.751	.965	.774	1.203	NO DIF
41.	-.002	.112	.000	.984	.998	.801	1.243	NO DIF
42.	.095	.114	.697	.404	1.099	.880	1.374	NO DIF
43.	.012	.112	.011	.916	1.012	.812	1.261	NO DIF
44.	.130	.112	1.349	.245	1.139	.914	1.420	NO DIF
45.	-.101	.113	.801	.371	.904	.725	1.128	NO DIF
46.	.254	.112	5.095	<b>.024*</b>	1.289	1.034	1.607	DIF
47.	.045	.112	.158	.691	1.046	.839	1.303	NO DIF
48.	.027	.113	.057	.811	1.027	.824	1.281	NO DIF
49.	-.114	.112	1.034	.309	.892	.716	1.112	NO DIF
50.	-.035	.113	.097	.756	.966	.774	1.204	NO DIF
51.	.259	.113	5.263	<b>.022*</b>	1.296	1.038	1.617	DIF
52.	.126	.113	1.249	.264	1.135	.909	1.416	NO DIF
53.	.240	.112	4.548	<b>.033*</b>	1.271	1.020	1.585	DIF
54.	-.116	.113	1.060	.303	.890	.713	1.111	NO DIF
55.	.174	.112	2.386	.122	1.190	.954	1.483	NO DIF
56.	.097	.113	.744	.388	1.102	.883	1.375	NO DIF
57.	.085	.112	.577	.448	1.089	.874	1.357	NO DIF
58.	.126	.112	1.259	.262	1.134	.910	1.414	NO DIF
59.	-.128	.113	1.273	.259	.880	.704	1.099	NO DIF
60.	.050	.112	.196	.658	1.051	.843	1.310	NO DIF

Variables on Gender: \*DIF EXIST; Items 32, 46, 51 and 53 only.

**Table 2: Gender Cross Tabulation in Group Performance of 2021 Mathematics NECO (BECE) multiple choice Items.**

Items	Dichotomous Score	Male	Female
<b>Q32</b>	INCORRECT	327	249
	CORRECT	348	<b>351*</b>
<b>Q46</b>	INCORRECT	306	310
	CORRECT	<b>369*</b>	290
<b>Q51</b>	INCORRECT	283	290
	CORRECT	<b>392*</b>	310
<b>Q53</b>	INCORRECT	305	307
	CORRECT	<b>370*</b>	293

**a = the gender that DIF favoured**

Table 1 and 2 shows four (4) items that identified significant DIF in gender of students using binary logistic regression analysis with the aid of SPSS version 23 which items 32, 46, 51 and 53 reveals significant differences between male and female students with significant level less than 0.05, i.e ( $P < .05$ ). This represents 7% of the total in the 2021 Mathematics NECO BECE multiple choice items while 93% of items do not differentiate significantly between male and female students. The results further reveal that, item Items 46, 51 and 53 favoured male students which placed the female students at disadvantaged group which DIF was found only from item 32. Therefore, there is significant difference in the items that functioned differentially in the 2021 Mathematics NECO (BECE) multiple choice items in North Central Nigeria based on gender. Thus, the null hypothesis one is rejected.

### **Discussion of Findings**

The finding of hypothesis revealed a significant difference in the items that functioned differentially by gender in 2021 NECO (BECE) Mathematics items. A small subset of items (7%) showed statistical DIF, with three items favouring males and one favouring females. This pattern is consistent with studies like Madu (2012) and Abedalaziz (2010), which also reported male advantages in certain mathematics domains, potentially linked to spatial reasoning or context familiarity. However, the failure of these items to retain significance after multiple testing corrections aligns more closely with the findings of Yusuf and Dauda (2022) and Lawal and Salami (2023), who reported no significant gender bias.

### **Conclusion**

The researcher concluded that, while a small number of items in the 2021 NECO (BECE) Mathematics examination initially showed signs of gender-related Differential Item Functioning, these findings were not robust to statistical correction for multiple comparisons. Therefore, the examination as a whole demonstrates a high degree of gender fairness. Nonetheless, the pattern observed in a few items underscores the value of continuous DIF monitoring and qualitative item review to uphold the highest standards of assessment equity.

### **Recommendations**

1. NECO and other examining bodies should initiate specialized training programs for their item writers and test development committees. This training should focus on identifying and mitigating sources of construct-irrelevant variance that may lead to DIF, such as context, phrasing, or illustrative examples that could be unconsciously gendered. The goal is to proactively build fairer tests from the outset, rather than solely relying on post-hoc statistical detection.
1. NECO examining body should establish a qualitative item review panel which will propose the establishment of a permanent, diverse review panel comprising experts in gender studies, mathematics pedagogy, and curriculum development. This panel would conduct a qualitative and sensitivity review of all mathematics items before they are finalized. Their role would be to scrutinize items for potential biases that statistical DIF analysis might only catch after the examination has been administered, thereby adding a crucial layer of preventative quality control.
2. Educational research bodies and examining agencies like NECO commission in-depth qualitative and mixed-methods studies specifically focused on the items flagged for DIF. Instead of just identifying that an item favors one gender, this research should investigate why by using think-aloud protocols and interviews with male and female students. Understanding the cognitive processes and reasoning behind the differential performance is essential for creating meaningful, evidence-based guidelines for future item writing.