EXAMINING GENDER AND SCHOOL-TYPE BIAS THROUGH DIFFERENTIAL ITEM FUNCTIONING IN ENGLISH LANGUAGE MOCK SSCE IN GOMBE STATE, NIGERIA

ABUBAKAR, M.[1], OGUNJIMI, M. O.[2], AILERU, O. M.[3] AND SEIDU, Y. H.[5]
[1]Department of Educational Foundations, Faculty of Education, Federal University of Kashere
[2,3 &4]Department of Social Sciences Education, Faculty of Education, University of Ilorin, Ilorin, Nigeria.
abuamirandamira@gmail.com, ogunjimi.mo@unilorin.edu.ng, walechoice1@gmail.com & seiduyaya07@gmail.com

Abstract
This study examined differential item functioning (DIF) in the 2023 English Language Multiple-choice Mock Examination in Gombe State, Nigeria. Ex-Post Facto research design was adopted for the study. The study population comprised 25,173 candidates who sat for the 2023 English Language Mock Examination in public senior secondary schools in Gombe State. A sample size of 756 was selected from 16 public senior secondary schools using multi-stage sampling procedure. Socio-Demographic Inventory (SDI) was used to collect students' demographic details and English Language Paper I Multiple-Choice Test administered by the Gombe State Ministry of Education in the 2023 MOCK SSCE was adopted as instruments for the study. The study employed binary logistic regression analysis to detect DIF. The results revealed that many items exhibited DIF across gender, and school type. The findings highlight the importance of examining DIF in educational tests to ensure fairness and validity. The study recommends that test developers should regularly examine DIF in educational tests and review or replace items exhibiting DIF to prevent bias against certain groups of students.

Keywords: Differential Item Functioning (DIF), Gender, School type, Logistic Regression,

Introduction

Assessment is a vital component of the Nigerian educational system, particularly in measuring students' academic achievement. English Language holds a unique position as both a subject and the primary medium of instruction, making its assessment critical for students' success in Senior School Certificate Examinations (SSCE) such as WAEC, NECO, and NABTEB. However, concerns have been raised about potential biases in test items that may disadvantage certain student subgroups based on gender, location, school type, and subject combination (Abubakar et al., 2024).

English is not only a compulsory subject but also a prerequisite for higher education admission in Nigeria (N P E, 2013). As a result, students' achievement in English Language significantly impacts their academic and career prospects (Umar et al., 2024). The effectiveness of assessments in reflecting true learning outcomes has been questioned, as poor examination results hinder students from advancing to tertiary education (Obinne & Amali, 2014; Amajuoyi, 2015).

Testing, particularly multiple-choice assessments, is a key tool for evaluating students' knowledge and abilities. While tests are intended to be objective, they can sometimes introduce bias, affecting the validity of results (Nworgu, 2015; Moyo & Nenty, 2017). Differential Item Functioning (DIF) is used to identify whether test items advantage or disadvantage specific demographic groups despite having equal ability levels (Jimoh et al., 2023; Umar et al., 2024). DIF can be categorized as uniform (consistent advantage for a group) or non-uniform (performance differences vary based on ability levels) (Abubakar et al., 2024; Bizumic et al., 2023; Walker, 2011).

Two major test theories—Classical Test Theory (CTT) and Item Response Theory (IRT)—are employed to analyze test performance. IRT is particularly effective for DIF analysis as it examines the relationship between students' abilities and test item characteristics (Xue & Chen, 2023). Various statistical methods, including Mantel-Haenszel and logistic regression, are used to detect DIF (Ogbebor & Onuka, 2013 & Magis et al., 2010).

Several socio-demographic factors influence DIF, including gender, school type, and subject combination. Studies suggest that societal norms and cultural expectations contribute to gender disparities in subject performance, with males generally dominating science-related fields (Akinlolu et al., 2023; Wang & Degol, 2017). Additionally, research indicates that private school students tend to outperform their public schools' counterparts, potentially affecting assessment outcomes (Moonpreneur, 2023; Archbishop Murphy High School, 2022).

Given these considerations, this study seeks to examine the presence of DIF in standardized multiple-choice English Language tests in the Gombe State Mock Examination. The findings will contribute to ensuring fairness and validity in assessments, ultimately enhancing educational equity.

Standardized tests such as the MOCK Senior Secondary Certificate Examination (SSCE) and national exams (WAEC, NECO, NABTEB) should fairly assess students' proficiency without bias. These assessments must be valid and reliable, ensuring all students, regardless of demographic differences, have equal opportunities to demonstrate their abilities. Item Response Theory (IRT) principles, particularly unidimensionality and local independence, should be maintained to ensure that test items assess only English language proficiency without external influences like gender, school type, or socio-economic status.

However, recurring underachievement in English Language has been observed in national and state-level examinations. NECO results and Gombe State's

mock exam performance over the past five years reveal consistently low pass rates, with only 25% of students passing in 2019, 20% in 2020, 32% in 2021, 38% in 2022, and 34% in 2023. This persistent trend raises concerns about potential item bias related to factors such as test dimensionality, gender, school type, location, and socio-economic background.

Although, several studies have examined Differential Item Functioning (DIF) in public and private secondary schools across Nigeria. Research by Ogbogo and Opara (2019), Shanmugan (2018), Onuka (2013), Patrick and Bright (2018), Bakama et al. (2018), and Abba (2014) explored DIF in various subjects and regions.

But, none of these studies specifically analyzed DIF in the 2023 English Language Multiple-Choice Mock Examination in Gombe State. Therefore, this study aims to determine whether specific test items exhibit DIF based on students' characteristics, potentially affecting assessment validity and fairness. Hence, this study aims to detect differential Item Functioning of 2023 English Language Multiple-choice Mock Senior School Certificate Examination in Gombe state through the following objectives:

1. To find out the number of items that functioned differentially between Boys' and Girls' students in the 2023 English Language Multiple-choice Mock Examination.
2. To find out the number of items that functioned differentially between Boarding and Day students in the 2023 English Language Multiple-choice Mock Examination.

Research Questions

The following research questions were raised and answered:

1. What number of items in the 2023 English Language Multiple-choice Mock Examination that functioned differentially between boys' and girls' students?
2. What number of items in the 2023 English Language Multiple-choice Mock Examination that functioned differentially between Boarding and Day students?

Research Hypotheses

Hypothesis One: The test items in 2023 English Language mock SSCE do not significantly function differentially with respect to gender as measured by the item scores of the students.

Hypothesis Two: The test items in 2023 English Language mock SSCE do not significantly function differentially with respect to school type as measured by the item scores of the students.

Methodology

This study adopted an Ex-Post Facto research design, which is suitable for analyzing non-manipulated variables such as school type, school location, gender, and subject combination. The design was used to determine whether items in the

2023 English Language multiple-choice Mock Senior School Certificate Examination (SSCE) in Gombe State functioned differently across student subgroups. The study population comprised 25,000 candidates who sat for the 2023 English Language Mock Examination in public senior secondary schools (14,762 males and 10,238 females). Based on Item Response Theory (IRT) 3-Parameter Logistic Model (3PLM) recommendations, a sample size of 760 guided by the research advisors (2006) was selected. A multi-stage procedure technique was employed. Random and purposive sampling were used to select 16 public senior secondary schools from 7 Local Government Areas (LGAs). Proportionate sampling to determine the percentage of students based on their demographic stratifications. Simple random sampling to ensure unbiased participant selection. Socio-Demographic Inventory (SDI) was used to collect students' demographic details (gender, location, school type, and subject combination), English Language Paper I Multiple-Choice Test: A 60-item, four-option test (A–D), administered by the Gombe State Ministry of Education in the 2023 MOCK SSCE was adopted as instruments for the study. Data was analyzed using SPSS (Version 26) with: Logistic Regression (LR): To detect DIF in test items, analyzing correct (1) vs. incorrect (0) responses. Chi-Square Test: To assess statistical significance of DIF-affected items. Descriptive Statistics (Cross-tabulation): To determine the percentage of students who answered each item correctly across demographic groups (Umar et al, 2024).

Results

Table 1: Descriptive Statistics of Groups for 2023 English Language Mock Items

		Descriptive Statistics		
		N	Mean	S.D
Group	All Group	756	27.48	10.95
Gender	Male	409	29.29	10.66
	Female	347	29.43	10.79
School Tupe I	Boarding	431	28.97	10.55
	Day	335	28.39	11.70

Table 1 shows that the descriptive statistics of the examinees in 2023 English Language Mock examination in Gombe State. The gender group has mean of 25.18 and 27.43 as well as standard deviation of 10.71 and10.79 for male and female students respectively. The school type I group has mean of 26.97 and 28.16 as well as standard deviation of 10.55 and 11.46 for boarding and day students respectively. The reliability coefficients of the items used is 0.90 which is above the benchmark of 0.70. This means that the examinees sampled in this study were of equal ability and the data collected conform with unidimensionality assumption of DIF.

Answering the Research Questions

The research questions raised in the study were answered using the logistic regression analysis through the SPSS software.

(*RQ_1*: *What number of items in the 2023 English Language Multiple-choice Mock Examination functioned differentially between Male and Female students?*

***Ho_1*:** *The test items in 2023 English Language mock SSCE do not significantly function differentially with respect to gender as measured by the item scores of the students*).

Table 2 shows the items flagged by Logistic Regression method of detecting gender DIF for the year 2023 English Language multiple choice mock examination. The reference group in this study is the male respondent and the focal group is the female respondent. In the logistic regression method, items whose Odds-ratio value is greater than 1 favour the reference group (male candidates) and those whose Odds-ratio value is less than 1 favour the focal group (female candidates). An item reveals uniform DIF when the significance Odds-ratio is for the group and the item reveals non-uniform DIF when the significant Odds-ratio is for the interaction between the groups and the total score. The classification criteria suggested by Jimoh et al, (2023) was used to categorize the DIF levels. If the difference between R-squared values is below 0.035 then there is said to be negligible DIF (Category A), differences in R-squared values between 0.035 and 0.070 are considered moderate DIF (Category B), and differences in R-squared values above 0.070 are considered large DIF (Category C).

Table 2: Summary of Logistic Regression Method for Detecting DIF in 2023 English Language Multiple Choice Items based on Gender

Items	Variable	Odds-ratio	Sig.	ETS-C	Items	Variable	Odds-ratio	Sig.	ETS-C
1	Group	1.214	0.186		31	Group	2.873*	0.000	C
	Interaction	0.997	0.382			Interaction	1.040**	0.000	C
2	Group	3.092*	0.000	C	32	Group	2.163*	0.000	B
	Interaction	1.019**	0.000	B		Interaction	1.023**	0.000	B
3	Group	0.701*	0.016	A	33	Group	0.885	0.405	
	Interaction	0.997	0.388			Interaction	0.987**	0.000	A
4	Group	2.253*	0.000	B	34	Group	1.870*	0.000	A
	Interaction	1.019**	0.000	B		Interaction	1.011**	0.005	A
5	Group	2.475*	0.000	B	35	Group	3.658*	0.000	C
	Interaction	1.025**	0.000	B		Interaction	1.021**	0.000	B
6	Group	1.855*	0.000	A	36	Group	2.958*	0.000	C
	Interaction	1.017**	0.000	A		Interaction	1.021**	0.000	B
7	Group	0.685*	0.054	A	37	Group	0.933	0.685	
	Interaction	0.996	0.436			Interaction	0.999	0.893	
8	Group	2.415*	0.000	B	38	Group	3.447*	0.000	C
	Interaction	1.027**	0.000	C		Interaction	1.040**	0.000	C
9	Group	1.796*	0.000	A	39	Group	3.987*	0.000	C
	Interaction	1.008**	0.042	A		Interaction	1.037**	0.000	C
10	Group	2.545*	0.000	B	40	Group	0.431*	0.005	A
	Interaction	1.020**	0.000	B		Interaction	0.952**	0.000	C
11	Group	0.700	0.146		41	Group	3.037*	0.000	C
	Interaction	1.003	0.649			Interaction	1.031**	0.000	C
12	Group	1.830*	0.001	A	42	Group	3.212*	0.000	C
	Interaction	1.016**	0.000	A		Interaction	1.033**	0.000	C
13	Group	1.387	0.085		43	Group	2.391*	0.000	B
	Interaction	1.002	0.737			Interaction	1.023**	0.000	B
14	Group	2.230*	0.000	B	44	Group	2.749*	0.000	B
	Interaction	1.025**	0.000	B		Interaction	1.021**	0.000	B
15	Group	0.545*	0.004	A	45	Group	4.596*	0.000	C
	Interaction	0.972**	0.000	B		Interaction	1.031**	0.000	C
16	Group	2.532*	0.000	B	46	Group	1.660*	0.001	A
	Interaction	1.028**	0.000	C		Interaction	0.999	0.912	
17	Group	0.840	0.237		47	Group	2.140*	0.000	B
	Interaction	0.998	0.556			Interaction	1.028**	0.000	C
18	Group	1.649*	0.004	A	48	Group	2.525*	0.000	B
	Interaction	1.014**	0.002	A		Interaction	1.022**	0.000	B
19	Group	2.317*	0.000	B	49	Group	2.512*	0.000	B
	Interaction	1.011**	0.004	A		Interaction	1.024**	0.000	B
20	Group	2.694*	0.000	C	50	Group	2.257*	0.000	B
	Interaction	1.019**	0.000	B		Interaction	1.019**	0.000	A
21	Group	2.141*	0.000	B	51	Group	1.313	0.100	
	Interaction	1.018**	0.000	A		Interaction	0.997	0.447	
22	Group	3.675*	0.000	C	52	Group	2.118*	0.000	A
	Interaction	1.025**	0.000	B		Interaction	1.014**	0.000	A
23	Group	2.114*	0.000	B	53	Group	2.153*	0.000	B
	Interaction	1.021**	0.000	B		Interaction	1.018**	0.000	B
24	Group	2.159*	0.000	B	54	Group	2.211*	0.000	B
	Interaction	1.018**	0.000	A		Interaction	1.009	0.060	
25	Group	0.611*	0.001	A	55	Group	0.446*	0.000	B
	Interaction	0.986**	0.000	A		Interaction	0.980**	0.000	B
26	Group	2.014*	0.000	B	56	Group	2.699*	0.000	B
	Interaction	1.024**	0.000	B		Interaction	1.023**	0.000	B
27	Group	2.731*	0.000	C	57	Group	3.729*	0.000	C
	Interaction	1.023**	0.000	B		Interaction	1.026**	0.000	C
28	Group	0.968	0.854		58	Group	3.071*	0.000	C

P<0.05, * Uniform DIF, **Non-uniform DIF, ETS-C= ETS Classification

Table 2 reveals that 55 out of 60 items (91.6%) were flagged for gender DIF, meaning the hypothesis one is rejected because the items functioned differently for males and females. 47 items (78.3%) favored male students (reference group). 8

items (13.3%) favored female students (focal group). 5 items exhibited only uniform DIF, while 2 items exhibited only non-uniform DIF. 48 items were flagged for both uniform and non-uniform DIF. Only 5 items (1, 11, 13, 17, 51) were free from DIF and functioned equally for both genders. The high number of moderate and large DIF items suggests that gender bias may exist in the 2023 English Language Mock Examination. If such items are used in decision-making without proper review, they could compromise test fairness and reliability.

Research Question One: What number of items in the 2023 English Language Multiple-choice Mock Examination functioned differentially between Boarding and Day schools?

Hypothesis Two: The test items in 2023 English Language mock SSCE do not significantly function differentially with respect to school type as measured by the item scores of the students.

Table 3 shows the items flagged by Logistic Regression method of detecting DIF for the year 2023 English Language multiple choice mock examination based on boarding and day schools. The reference group in this study is the boarding school students and the focal group is the day school students. In the logistic regression method, items whose Odds-ratio value is greater than 1 favour the reference group (boarding candidates) and those whose Odds-ratio value is less than 1 favour the focal group (day candidates). An item reveals uniform DIF when the significance Odds-ratio is for the group and the item reveals non-uniform DIF when the significant Odds-ratio is for the interaction between the groups and the total score. The classification criteria suggested by Jimoh et al, (2023) was used to categorize the levels of DIF. If the difference between R-squared values is below 0.035 then there is said to be negligible DIF (Category A), differences in R-squared values between 0.035 and 0.070 are considered moderate DIF (Category B), and differences in R-squared values above 0.070 are considered large DIF (Category C).

Table 3: Summary of Logistic Regression Method for Detecting DIF in 2023 English Language Multiple Choice Items based on Boarding and Day Schools

Items	Variable	Odds-ratio	Sig.	ETS-C	Items	Variable	Odds-ratio	Sig.	ETS-C
1	Group	0.284*	0.000	C	31	Group	0.461*	0.000	B
	Interaction	0.980**	0.000	B		Interaction	0.987**	0.000	A
2	Group	0.514*	0.000	A	32	Group	0.632*	0.007	A
	Interaction	0.989**	0.001	A		Interaction	0.987**	0.000	A
3	Group	0.721*	0.026	A	33	Group	0.359*	0.000	A
	Interaction	0.994	0.056			Interaction	0.983**	0.000	A
4	Group	0.497*	0.000	B	34	Group	0.501*	0.000	A
	Interaction	0.990**	0.005	A		Interaction	0.992**	0.013	A
5	Group	0.510*	0.000	B	35	Group	1.081	0.587	
	Interaction	0.987**	0.000	A		Interaction	1.008**	0.016	A
6	Group	0.738*	0.049	A	36	Group	0.581*	0.000	A
	Interaction	0.991**	0.006	A		Interaction	0.994	0.090	
7	Group	1.951*	0.001	A	37	Group	0.335*	0.000	A
	Interaction	1.016**	0.000	A		Interaction	0.978**	0.000	B
8	Group	0.471*	0.000	B	38	Group	0.533*	0.000	C
	Interaction	0.984**	0.000	B		Interaction	0.987**	0.000	B
9	Group	1.878*	0.000	A	39	Group	0.499*	0.000	A
	Interaction	1.014**	0.000	A		Interaction	0.988**	0.001	A
10	Group	0.535*	0.000	A	40	Group	1.160	0.583	
	Interaction	0.993**	0.035	A		Interaction	1.006	0.332	
11	Group	1.634	0.040		41	Group	0.395*	0.000	A
	Interaction	1.015	0.002			Interaction	0.985**	0.000	B
12	Group	2.654*	0.000	B	42	Group	0.464*	0.000	A
	Interaction	1.014**	0.000	A		Interaction	0.985**	0.000	A
13	Group	1.841*	0.001	A	43	Group	0.491*	0.000	A
	Interaction	1.010	0.011			Interaction	0.986**	0.000	B
14	Group	0.453*	0.000	B	44	Group	0.410*	0.000	C
	Interaction	0.985**	0.000	A		Interaction	0.988**	0.000	A
15	Group	1.373	0.117		45	Group	1.340	0.057	
	Interaction	0.998	0.727			Interaction	1.008**	0.013	A
16	Group	0.595*	0.000	A	46	Group	1.495*	0.011	A
	Interaction	0.990**	0.002	A		Interaction	1.007**	0.053	B
17	Group	0.409*	0.000	A	47	Group	0.549*	0.000	A
	Interaction	0.983**	0.000	A		Interaction	0.988**	0.001	C
18	Group	0.639*	0.000	A	48	Group	0.489*	0.000	B
	Interaction	0.996	0.245			Interaction	0.987**	0.000	A
19	Group	0.984	0.911		49	Group	0.561*	0.000	A
	Interaction	1.008**	0.011	A		Interaction	0.987**	0.000	A
20	Group	1.476*	0.009	A	50	Group	0.430*	0.000	A
	Interaction	1.012**	0.001	B		Interaction	0.985**	0.000	A
21	Group	0.521*	0.000	A	51	Group	0.573*	0.000	A
	Interaction	0.987**	0.000	C		Interaction	0.993	0.071	
22	Group	0.423*	0.000	A	52	Group	1.028	0.854	
	Interaction	0.985**	0.000	A		Interaction	1.005	0.156	
23	Group	0.391*	0.000	A	53	Group	1.164	0.303	
	Interaction	0.983**	0.000	A		Interaction	1.003	0.375	
24	Group	0.570*	0.000	A	54	Group	0.777	0.094	
	Interaction	0.992**	0.015	A		Interaction	1.001	0.824	
25	Group	0.369*	0.000	A	55	Group	0.703*	0.040	A
	Interaction	0.983**	0.000	B		Interaction	0.993	0.065	
26	Group	0.489*	0.000	B	56	Group	1.364	0.044	
	Interaction	0.987**	0.000	A		Interaction	1.005	0.177	
27	Group	0.593*	0.000	A	57	Group	1.457	0.109	
	Interaction	0.994**	0.054	A		Interaction	1.006	0.066	
28	Group	1.268	0.185		58	Group	1.007	0.962	

P<0.05, * Uniform DIF, **Non-uniform DIF, ETS-C = ETS Classification

Table 3 reveals that Out of 60 items administered, 49 items (81.7%) exhibited DIF based on school type meaning the hypothesis two is also rejected because the items functioned differently between boarding and day schools. 9 items favored

boarding school students (reference group). 40 items favored day school students (focal group). 38 items showed both uniform and non-uniform DIF. 5 items were flagged for only uniform DIF (items: 3, 13, 18, 29, 36, 51, 55). 2 items were flagged for only non-uniform DIF (items: 28, 45). The majority of test items exhibited DIF based on school type, but the DIF levels were mostly negligible to moderate. This suggests that while some bias may exist, it is unlikely to significantly affect test reliability if non-functional items are removed.

Discussion of Findings

The finding that 55 (91.6%) of the items were flagged for gender DIF in the 2023 English Language multiple choice mock examination suggests that there are significant differences in how male and female students respond to the test items. Several possible reasons could explain this finding: Biased item content: Some test items may contain biased language, stereotypes, or cultural references that advantage one gender over the other. Differential instruction: Male and female students may receive different instructional experiences, leading to varying levels of familiarity with the test content. Socio-cultural factors: Societal expectations, gender roles, and cultural norms may influence how male and female students approach the test and respond to specific items. Test-taking strategies: Male and female students may employ different test-taking strategies, such as guessing or time management, which could impact their performance on specific items. Studies that support this finding include Abduaziz (2010), who found that females showed a statistically significant advantage over males on numerical ability, while males showed an advantage over females on spatial and deductive abilities. Barnabas (2012), who discovered that male and female students functioned differentially in 39 items on the Mathematics Achievement Test (MAT). Omorogiuwa (2016), who found that male and female examinees functioned differentially in 17 items (34%) on the National Business and Technical Examinations Board (NABTEB) Mathematics Multiple Choice Test Items. However, some studies contradict this finding: Yousif (2011), who reported that male students significantly outperformed their female counterparts in reading comprehension tests. Pae (2011), who found that items 20 and 68 seemed to be more difficult for females, while items 39, 78, and 86 were more difficult for males. These conflicting results highlight the complexity of the issue and the need for further research to fully understand the factors contributing to gender DIF in educational assessments.

The result of the study revealed that majority of the 2023 English Language mock items were flagged for school type DIF based on boarding and day schools. This study equally detected more negligible and moderate school type DIF items based on boarding and day schools in the English mock examination with more of them in favour of the focal group (day students). The flexibility of access to internet facilities by the day students may probably account for the reason the flagged items in the 2023 English Language mock examination favoured the day students

compared to the boarding students. This finding does not concur with Patrick and Bright (2018) who observed incidence of gender, location, socio-economic, school type and school ownership differential functioning in 2014 BECE social studies multiple choice test and recommended that examination bodies should be mindful of the disparities that exist between the variables. This result is also not in tandem with Enunwah et al (2014) who found that seven of the items and one of the nine topics (variations) were significantly biased against school ownership types.

Conclusion

With respect to the findings of this study, it could be concluded that the 2023 English Language Multiple-choice Mock Senior School Certificate Examination in Gombe State functioned differentially among various groups of the examinees. More items of the 2023 English Language mock examination were flagged for DIF based on gender, single sex and mixed schools as well as boarding and day schools. On the other hand, less items were flagged for DIF in the 2023 English Language mock examination in Gombe State based on mega and common schools, school location as science and non-science combination. The study equally uncovered more negligible and moderate DIF based on school location, subject combination and school types except with gender which flagged larger DIF items.

Recommendations

Based on the findings of this study, the following recommendations were made:

i. Examination bodies should ensure that test items are free from biased language, stereotypes, or cultural references that may advantage one gender over the other, in order to minimize gender-based Differential Item Functioning (DIF). Educators should provide equal instructional experiences for male and female students to minimize differential instruction. They should also encourage male and female students to employ effective test-taking strategies, such as time management and guessing strategies.

ii. Examination bodies should ensure that test items are relevant and applicable to students from both boarding and day schools, in order to minimize school-type based DIF. Educators should provide equal access to internet facilities for students from both boarding and day schools. They should also encourage students from boarding schools to participate in extracurricular activities that promote interaction with students from day schools. Teachers should be aware of the potential differences in instructional experiences between boarding and day schools.

References

Abba, A. (2014). Analysis of differential item functioning of senior secondary certificate (NECO) English examination in Dawakin Kudu educational zone, Kano State.

Abubakar, M., Ogunjimi, M. O., Muhammad, D., & Abubakar, H. (2024). Differential item functioning method as an item bias indicator for English Language Mock SSCE Examination in Gombe State. *Kashsere Journal of Education*, 7(2). https://www.ajol.info/index.php/kje/issue/view/25717

Akinlolu, A. A., Oludipe, D. O., & Onyeocha, A. O. (2023). Factors influencing students' performance in English language in Nigeria.

Amajuoyi, R. E. (2015). The role of testing and evaluation in educational development.

Bakama, M. A., Malik, S. A., & Chiroma, A. (2018). The impact of pre-national diploma programme on academic performance in northern polytechnics in Nigeria.

Bizumic, B., Reynolds, K. J., Turner, J. C., & Subasic, E. (2023). Avowed versus unavowed prejudice: A study of differential item functioning.

Chang, H. (2013). The impact of socio-cultural factors on girls' education in Nigeria.

Federal Republic of Nigeria (2013). *National policy of education.* Abuja: Federal Government press.

Jimoh, M. I., Daramola, D. S., Oladele, J. I., Sheu, A. L., & Ogunjimi, M. O. (2020). Differential Item Functioning of Mathematics Joint Mock Multiple-Choice Test Items in Kwara State, Nigeria. *Journal of Multidisciplinary Studies, 9(1), 67-92.* http://dx.doi.org/10.7828/jmds.v9i1.1451

Moonpreneur. (2023). Private vs public schools: Which is better?

Obinne, J. I., & Amali, E. O. (2014). The impact of testing and evaluation on students' academic performance.

Ogbebor, P. O., & Onuka, A. O. (2013). *An investigation of differential item functioning in mathematics items using Rasch model.*

Ogbogo, B. E. (2019). An introduction to item response theory.

Ogbogo, B. E., & Opara, J. A. (2019). Differential item functioning in English language test using item response theory for ethnic groups in south-south Nigeria.

Onuka, A. O. (2013). *Differential item functioning method as an item bias indicator in Delta State.*

Onyeocha, A. O., Oludipe, D. O., & Akinlolu, A. A. (2023). *Factors influencing students' performance in English language in Nigeria.*

Premium Times. (2023). NECO releases 2023 SSCE results.

Robitzsch, A. (2023). Item response theory: A review of the literature.

Shanmugan, A. (2018). *Determining gender differential item functioning for mathematics in coeducational school culture.*

Umar, U. Y., Zakka, K., & Abubakar, M. (2024). Differential item functioning of general studies mathematics examination in Gombe State Polytechnic, Bajoga, Nigeria. *Nigerian Journal of Educational Research and Evaluation, 23.*