

ANALYSIS OF DIFFERENTIAL ITEM FUNCTIONING OF 2021 WASSCE BIOLOGY MULTIPLE-CHOICE EXAMINATION AMONG SECONDARY SCHOOL STUDENTS IN NORTH-WEST, NIGERIA

Nuraddeen Yusuf Galadanchi

ygaladanchi2015@gmail.com

Department of Early Childhood Care Education

Yusuf Bala Usman College of Education and Legal Studies Daura

Katsina State

Dr. A. T. Olutola & Dr. S. O. Bada

Department of Educational Psychology and Counselling

Federal University Dutsin-Ma.

Abstract

This study analyse differential item functioning (DIF) of the 2021 West African Senior School Certificate Examination (WASSCE) Biology multiple-choice examination among secondary school students in North-west, Nigeria. The study was guided by two objectives, research questions and hypotheses. A descriptive survey research design was employed. The study population comprised 412, 323 biology students (228,410 males and 183,913 females) from 2,162 public senior secondary schools across the study area. A sample of 1,532 senior secondary school III Biology students was drawn from 18 purposively selected senior secondary schools (12 urban and 6 rural). Data collection involved 50 WASSCE Biology paper 3 items used in the 2021 examination. Reliability index of the instrument was determined using Cronbach alpha at 0.73. Binary logistic regression was used to test the hypotheses at 0.05 level of significance. Findings revealed that the 2021 WASSCE Biology multiple-choice test items contained nine questions (18%) that functioned differentially based on gender and school location, favouring males over their female counterparts and favouring urban school students over their rural school counterparts. It is recommended that West African Examinations Council (WAEC) should constitute an item review committee to scrutinize test items for potential gender and context biases.

Keywords: Assessment, Biology multiple-choice test, DIF analysis.

Introduction

In Nigeria, achievement at any level of education is awarded with certification of those who successfully completed a course of study with good academic records. Educational test result is an important yardstick by which society pronounces the product of its educational system. As such, it is of paramount importance that educational institutions and examination bodies should conduct tests that will establish the desired characteristics of their examinees (Ibrahim & Binta, 2019).

The National Policy on Education (FRN, 2014) points out that a test should measure the knowledge and skills taught in schools with negligible or no bias against students' gender, location, school type and socio-economic background. Similarly, for any examination to be judged reliable and valid the conduct of the examination and the published results should be deemed fair and achieve high level of acceptance by the public. More so, the examination

should ensure that no particular candidate or group of candidates have an unfair advantage over others. Thus, to draw valid conclusions on the results obtained from an achievement test, it is necessary that the test is a valid measurement of what it is intended to measure.

Differential Item Functioning (DIF) occurs when examinees from different groups show different probabilities of success on the item after matching on the underlying ability that the item is intended to measure. In simple terms, DIF arises when two groups of equal ability level students are not equally able to correctly answer an item (Lee, 2012). The psychometric property that typically must hold for scores to be equivalent when compared is acknowledged as measurement invariance, lack of bias or absence of item functioning is intended to be invariant with respect to irrelevant aspects of the test-takers, such as gender, location, ethnicity and socio-economic status (Orluwene & QueenSoap, 2019). If the factor leading to DIF is not part of the construct being tested, then the test is biased. Moreover, a test containing items exhibiting DIF could in turn create inaccurate observed total scores resulting in inaccurate estimation of the focal group's primary latent trait (example, biological concept).

There are two types of DIF, namely uniformed and non-uniformed DIF. Uniform DIF occurs when a group performs better than another group on all ability levels. That is, almost all members of a group outperform almost all members of the other group who are at the same ability levels. In the case of non-uniform DIF, members of one group are favoured up to a level on the ability scale and from that point on the relationship is reversed (Karami, 2012). That is, there is an interaction between grouping and ability level. Although the presence of DIF is a signal that an item may be biased, it does not guarantee that the item is unfair. Rather, the presence of DIF indicates the existence of a latent trait besides the one of primary interest.

The fairness is established subsequently if the secondary latent trait that was detected statistically is intentionally related to the primary latent trait. Primary latent trait is a fundamental, directly measured construct that explains individual differences in observed behaviours or responses. For instance, intelligence testing, general cognitive ability is often considered a primary latent trait because it directly explains performance across various cognitive trait. Secondary latent trait is a higher order construct that emerges from the relationships among primary latent traits. It is not directly measured but inferred from the interplay of primary traits. For instance, emotional stability might emerge from the combination of primary traits such as anxiety, depression, and self-esteem. It is possible that the secondary latent traits may be necessary based on the content and test specifications, even if there are differences in performance between the reference individual (the benchmark) and the focal group (the primary group of interest).

DIF analysis is a critical component in psychometrics that examines the probability of correctly responding to or endorsing a test item, conditional on the underlying latent trait or ability being measured. This analysis is crucial for ensuring test fairness and validity across different demographic groups. Several statistical models have been developed to detect DIF, including binary logistic regression, the Mantel-Haenszel procedure, and item response theory (IRT)

models. These models assume that test-takers being compared have approximately equivalent abilities or trait levels.

Binary Logistic regression is a method that uses a logistic regression models to predict the probability of a correct response based on the total test score and group membership. DIF is detected when group membership significantly predicts item performance after controlling for overall ability. Mantel-Haenszel procedure compare the odds of a correct response between two groups across different ability levels. It involves creating contingency tables for each ability level and calculating a common odds ratio across these tables. Furthermore, Item Response Theory (IRT) model compares item characteristics curves between groups. DIF is identified when these curves differ significantly, indicating that the item functions differently for various groups at the same ability levels.

Rural and urban schools are distinguished primarily by their geographical locations, which significantly influence their resources, student demographics, and educational challenges. Rural schools are situated in less populated areas, often facing issues such as limited access to resources, transportation difficulties, and lower socioeconomic status of students (Smith & Lee, 2020). In contrast, urban schools are located in densely populated city areas, typically characterised by diverse student populations, higher student enrolment, and greater access to educational resources, but they may also confront challenges like overcrowding and socioeconomic disparities (John & Patel, 2019).

In addition, the area in which a school is located can affect the educational achievement of a student. For examples, Amuche and Fan (2014) found that out of 60 items in NECO Biology questions, 10 items were biased in relation to school type and eight items were biased in relation to school location. In addition, a study conducted by Madueke and Casmir (2022) showed that out of 50 items in WAEC 2020 May/June multiple choice biology questions, with respect to gender and school location, DIF were discovered in eight items. These items revealed significant DIF between male and female students as well as significant DIF between urban and rural students. There is a pressing need for further research into the identification of DIF in test items used to measure biology achievement. Consequently, this study aims to analyse the presence of DIF in the 2021 West African Senior School Certificate Examination (WASSCE) biology multiple-choice test items among secondary school students in the Northwest region of Nigeria.

Objective of the Study

The objectives of this study was to:

1. Determine the percentage of items in the 2021 Biology WASSCE multiple choice items that functioned differentially on the basis of gender and school location in North-west, Nigeria.

Research Question

The following research questions guided the study:

1. What is the percentage of items in the 2021 Biology WASSCE multiple choice items that functioned differentially on the basis of gender and school location in North-west, Nigeria?

Hypothesis

The following hypothesis was formulated and tested at 0.05 level of significance:

1. There is no significant difference in the percentage of items in the 2021 Biology WASSCE multiple-choice items that functioned differentially on the basis of gender and school location in North-west, Nigeria.

Methodology

In carrying out this study, a survey research design was employed. This design offers flexibility in data collection methods, allowing researchers to tailor its approach to best suit its research questions and target population (De Vaus, 2014). Therefore, in this study, the existing non-manipulated variables are WAEC Biology multiple choice examination items paper conducted in 2021 senior secondary certificate examination (SSCE). From a total population of 412,323 biology students in North-west Nigeria, a sample size of 1,532 was determined using the Research Advisor (2006) guidelines, ensuring a 95% confidence level with a 2.5% margin of error. This sample comprised 981 urban and 551 rural public senior secondary school students from Kaduna, Kano, and Sokoto states. Multi-stage sampling technique was used to select the sample of the study. By utilizing this multi-stage sampling technique, the researcher was able to obtain a representative sample that balanced geographical diversity, urban-rural distribution, and specific selection criteria, while maintaining statistical validity and resource efficiency (Smith, 2025).

The initial phase of the sampling process employed a cluster sampling technique to stratify the North-west states into three distinct groups, based on the historical administrative divisions of the region. The first cluster comprised Kaduna and Katsina states, with the former Kaduna state designated as one of the primary study areas. The second cluster consisted of Kano and Jigawa states, with the historical Kano state selected as the second research area. The third cluster encompassed Sokoto, Zamfara, and Kebbi states, with the original Sokoto state chosen as the third focal point for the study. This stratification approach ensured a representative sample across the diverse North-west region, taking into account historical and administrative consideration. The second stage utilized simple random and stratified sampling techniques to select students on the basis of gender in which the researcher utilized mixed school students and students from both urban and rural schools within each educational zone. This approach ensures both randomness and proportional representation of different sub-groups.

The final stage employed purposive sampling to select 12 urban and 6 rural schools, allowing researcher to focus on institutions that met specific criteria relevant to the study's objectives.

The researcher collaborated with biology teachers, who served as research assistants, to establish protocols for administration of the fifty 2021 WASSCE Biology multiple-choice examination questions. This coordination was essential because the biology teachers from the selected secondary schools were responsible for distributing and collecting the completed response of the questions from a sample of 1,532 students. To ensure a high response rate, a systematic approach was implemented to retrieve all administered fifty Multiple-choice biology questions.

The scoring system for the instrument was designed to be straightforward: correct responses were allocated 1 point while incorrect responses were allocated 0 point. The evaluation of the completed instruments was conducted manually by the research team. The data analysis for the study was conducted based on the formulated research hypotheses. Subsequently, the research hypotheses 1 and 2 were tested using binary logistic regression at a significance level of 0.05, employing SPSS version 23.0.

Results

Research Question One: What is the percentage of items in the 2021 Biology WASSCE multiple choice items that functioned differentially on the basis of gender in North-west, Nigeria?

Table1: Binary Logistic Regression analysis of DIF of 2021 biology WASSCE multiple-choice items by gender

Item	B	S.E	Wald	Sig	Exp (B)	95% C.I for Exp (B)		Decision
						Lower	Upper	
1.	-.427	.117	13.270	.000*	.653	.519	.821	DIF
2.	.144	.111	1.658	.198	1.154	.928	1.436	NO.DIF
3.	.252	.116	4.704	.030*	1.286	1.025	1.615	DIF
4.	-.279	.119	5.439	.020*	.757	.599	.957	DIF
5.	.322	.124	6.683	.010*	1.379	1.081	1.760	DIF
6.	-.030	.115	.070	.791	.970	.774	1.215	NO.DIF
7.	-.152	.106	2.062	.151	.859	.698	1.057	NO. DIF
8.	-.138	.112	1.515	.218	.871	.698	1.085	NO. DIF
9.	-.424	.119	12.594	.000*	.655	.518	.827	DIF
10.	.001	.126	.000	.991	1.001	.782	1.282	NO. DIF

11.	.022	.135	.026	.873	1.022	.784	1.332	NO. DIF
12.	-.035	.126	.075	.784	.966	.755	1.237	NO. DIF
13.	.107	.121	.784	.376	1.113	.878	1.410	NO. DIF
14.	.231	.124	3.464	.063	1.260	.988	1.608	NO. DIF
15.	.105	.121	.754	.385	1.111	.876	1.409	NO. DIF
16.	-.244	.116	4.442	.035*	.783	.624	.983	DIF
17.	-.101	.126	.643	.423	.904	.706	1.157	NO. DIF
18.	.209	.115	3.300	.069	1.232	.984	1.543	NO. DIF
19.	-.239	.114	4.423	.035*	.787	.630	.984	DIF
20.	.090	.112	.648	.421	1.094	.879	1.363	NO DIF
21.	-.058	.124	.217	.641	.944	.740	1.204	NO. DIF
22.	-.027	.114	.054	.816	.974	.778	1.216	NO. DIF
23.	.128	.117	1.200	.273	1.137	.904	1.430	NO. DIF
24.	-.030	.127	.055	.081	.971	.757	1.244	NO. DIF
25.	-.015	.114	.017	.896	.985	.788	1.232	NO. DIF
26.	.037	.117	.101	.751	1.038	.825	1.305	NO. DIF
27.	.150	.120	1.561	.211	1.162	.918	1.472	NO. DIF
28.	-.097	.126	.592	.442	.908	.709	1.162	NO. DIF
29.	-.029	.118	.061	.805	.971	.771	1.224	NO. DIF
30.	-.162	.121	1.802	.180	.851	.672	1.077	NO. DIF
31.	-.066	.116	.325	.569	.936	.745	1.175	NO. DIF
32.	-.099	.115	.792	.389	.906	.723	1.135	NO. DIF
33.	-.102	.116	.771	.380	.903	.720	1.133	NO. DIF
34.	-.137	.117	1.359	.244	.872	.693	1.098	NO. DIF
35.	-.303	.123	6.033	.014*	.739	.580	.941	DIF
36.	-.341	.114	8.892	.003*	.711	.568	.890	DIF

37.	.103	.110	.875	.349	1.108	.893	1.375	NO.DIF
38.	.104	.118	.769	.380	1.109	.880	1.399	NO.DIF
39.	-.034	.113	.087	.768	.967	.774	1.208	NO.DIF
40.	-.098	.121	.648	.421	.907	.715	1.150	NO.DIF
41.	-.062	.115	.292	.589	.904	.705	1.177	NO.DIF
42.	.010	.128	.006	.936	1.010	.785	1.300	NO.DIF
43.	.152	.117	1.690	.194	1.164	.926	1.464	NO.DIF
44.	.062	.135	.209	.648	1.064	.816	1.387	NO.DIF
45.	-.182	.121	2.286	.131	.833	.658	1.056	NO.DIF
46.	-.079	.113	.490	.484	.924	.740	1.154	NO.DIF
47.	.093	.119	.600	.439	1.097	.868	1.386	NO.DIF
48.	.099	.128	.604	.437	1.105	.859	1.420	NO.DIF
49.	-.228	.129	3.140	.076	.796	.619	1.024	NO.DIF
50.	.012	.126	.009	.923	1.012	.791	1.295	NO.DIF

Variables on Gender: *DIF EXIST; Item 1, 3, 4, 5, 9, 16, 19, 35 and 36 only.

Table 1 shows nine items that identified significant DIF in gender of students using binary logistic regression analysis with the aid of SPSS version 23, items 1, 3, 4, 5, 9, 16, 19, 35 and 36 reveals significant differences between male and female students with significant level less than 0.05, (i.e., $P < .05$). This represents 18% of the total 2021 Biology WASSCE multiple choice items while 82% of the items do not differentiate significantly between male and female students.

Hypothesis One: There is no significant difference in the percentage of items in the 2021 Biology WASSCE multiple-choice items that functioned differentially on the basis of gender in North-west, Nigeria.

Table 2: Gender Cross Tabulation in Group Performance of 2021 Biology WASSCE multiple choice Items 1, 3, 4, 5, 9, 16, 19, 35 and 36

Item	Dichotomous Score	Male	Female
1.	Incorrect	719	405

	Correct	219^a	189
3.	Incorrect	638	435
	Correct	300^a	159
4.	Incorrect	721	425
	Correct	217^a	169
5.	Incorrect	684	468
	Correct	254^a	126
9.	Incorrect	732	417
	Correct	205^a	178
16.	Incorrect	694	410
	Correct	244^a	184
19.	Incorrect	674	397
	Correct	265^a	196
35.	Incorrect	744	439
	Correct	194^a	155
36.	Incorrect	692	396
	Correct	246^a	198

a = DIF favoured

In the table 2, 18% of the total 2021 Biology WASSCE multiple choice items indicated DIF while 82% of the items do not differentiate significantly between male and female students. The table 1b reveal that items 1, 3, 4, 5, 9, 16, 19, 35 and 36 which represent 100% of the flagged items that favoured male students thus placing the female students at disadvantaged group. Therefore, there is significant difference in the percentage of items which functioned differentially in the 2021 WASSCE Biology multiple choice items in Northwest based on gender. Thus, the stated hypothesis one is rejected.

Research Question Two: What is the percentage of items in the 2021 Biology WASSCE multiple choice items that functioned differentially on the basis of school location in North-west, Nigeria?

Table 3: Binary Logistic Regression analysis of DIF of 2021 biology WASSCE multiple-choice items by in School Location.

Item	B	S.E	Wald	Sig	Exp (B)	95% C.I for Lower	Exp (B) Upper	Decision
1.	.069	.116	.356	.551	1.072	.854	1.345	NO.DIF
2.	.137	.109	1.598	.206	1.147	.927	1.419	NO.DIF
3.	.049	.112	.192	.661	1.050	.843	1.308	NO.DIF
4.	.222	.119	3.511	.061	1.249	.990	1.576	NO.DIF
5.	.185	.119	2.411	.120	1.203	.953	1.520	NO.DIF
6.	-.091	.112	.658	.417	.913	.732	1.138	NO.DIF
7.	.387	.104	13.837	.000*	1.472	1.201	1.805	DIF
8.	.045	.110	.168	.682	1.046	.843	1.299	NO.DIF
9.	-.149	.118	1.581	.209	.862	.684	1.087	NO.DIF
10.	-.116	.123	.895	.344	.890	.699	1.133	NO.DIF
11.	.008	.132	.003	.953	1.008	.778	1.305	NO.DIF
12.	-.145	.123	1.381	.240	.865	.680	1.101	NO.DIF
13.	.223	.118	3.585	.058	1.250	.992	1.576	NO.DIF
14.	.177	.120	2.182	.140	1.194	.944	1.511	NO.DIF
15.	.191	.118	2.615	.106	1.211	.960	1.526	NO.DIF
16.	.071	.114	.390	.532	1.074	.858	1.344	NO.DIF
17.	.032	.124	.065	.799	1.032	.810	1.315	NO.DIF
18.	.123	.111	1.225	.268	1.131	.909	1.406	NO.DIF
19.	.070	.112	.389	.533	1.072	.861	1.336	NO.DIF
20.	-.110	.109	1.024	.312	.896	.723	1.109	NO.DIF
21.	-.206	.121	2.895	.089	.813	.641	1.032	NO.DIF
22.	.212	.112	3.574	.059	1.237	.992	1.541	NO.DIF
23.	.202	.114	3.154	.076	1.224	.979	1.531	NO.DIF
24.	-.158	.124	1.625	.202	.854	.670	1.088	NO.DIF
25.	-.290	.112	6.759	.009*	.748	.601	.931	DIF
26.	-.215	.114	3.573	.059*	.806	.645	1.008	DIF
27.	-.259	.116	4.946	.026*	.742	.614	.970	DIF
28.	-.298	.123	5.835	.016*	.742	.583	.945	DIF
29.	-.074	.115	.414	.520	.929	.741	1.164	NO.DIF
30.	-.061	.119	.268	.604	.940	.745	1.187	NO.DIF
31.	.063	.114	.308	.579	1.065	.852	1.332	NO.DIF
32.	.145	.113	1.652	.199	1.157	.926	1.444	NO.DIF
33.	.099	.114	.756	.385	1.104	.883	1.379	NO.DIF
34.	.033	.115	.082	.775	1.034	.824	1.296	NO.DIF
35.	-.052	.121	.186	.667	.949	.748	1.204	NO.DIF
36.	.127	.113	1.256	.262	1.135	.909	1.417	NO.DIF
37.	-.148	.107	1.916	.166	.862	.699	1.064	NO.DIF

38.	.203	.115	3.077	.079	1.225	.977	1.536	NO.DIF
39.	-.154	.111	1.936	.164	.857	.690	1.065	NO.DIF
40.	.111	.119	.857	.355	1.117	.884	1.411	NO.DIF
41.	.042	.113	.141	.708	1.043	.836	1.301	NO.DIF
42.	.095	.126	.567	.452	1.099	.859	1.406	NO.DIF
43.	-.351	.113	9.595	.002*	.704	.564	.879	DIF
44.	.099	.132	.568	.451	1.105	.853	1.431	NO.DIF
45.	.144	.119	1.459	.227	1.155	.914	1.459	NO.DIF
46.	-.363	.111	10.645	.001*	.696	.559	.865	DIF
47.	-.053	.116	.207	.649	.949	.756	1.191	NO.DIF
48.	-.012	.124	.010	.921	.988	.774	1.260	NO.DIF
49.	.526	.131	16.195	.000*	1.692	1.310	2.186	DIF
50.	.285	.124	5.290	.021*	1.330	1.043	1.695	DIF

Variables on School Location: *DIF EXIST; Item 7, 25, 26, 27, 28, 43, 46, 49 and 50 only

Table 3 shows nine (9) items that identified significant DIF in school location of students, using binary logistic regression analysis with the aid of SPSS version 23, items 7, 25, 26, 27, 28, 43, 46, 49 and 50 reveals significant difference between urban and rural students with significant level less than 0.05 i.e. ($P < .05$). Thus, the stated hypothesis two is rejected.

Hypothesis Two: There is no significant difference in the percentage of items in the 2021 Biology WASSCE multiple-choice items that functioned differentially on the basis of school location in North-west, Nigeria.

Table 4: School Location Cross Tabulation in Group Performance of 2021 Biology WASSCE multiple choice Items 7, 25, 26, 27, 28, 43, 46, 49 and 50.

Item	Dichotomous Score	Urban	Rural
7.	Incorrect	509	377
	Correct	309^a	337
25.	Incorrect	595	475
	Correct	223	239^a
26.	Incorrect	605	497
	Correct	213	217^a
27.	Incorrect	623	508
	Correct	195	206^a
28.	Incorrect	657	537
	Correct	161	177^a
43.	Incorrect	611	482
	Correct	207	232^a
46.	Incorrect	597	466
	Correct	221	248^a
49.	Incorrect	619	600
	Correct	199 ^a	114

50.	Incorrect	615	572
	Correct	203 ^a	142

a = DIF favoured

In the above table 4, 18% of the total 2021 Biology WASSCE multiple choice items indicated DIF while 82% of the items do not differentiate significantly based on school location. The result further reveals that items 25, 26, 27, 28, 43 and 50, representing 67% of DIF items, which favoured rural students while only items 7, 49 and 50 representing 33% of DIF, favoured the urban students which eventually placed the rural school students at disadvantaged group. Therefore, there is significant difference in the percentage of items which functioned differentially in the 2021 WASSCE Biology multiple choice items in Northwest based on school location.

Discussion of Findings

The result in table 1 and 2 showed that nine items representing 18% in 2021 WASSCE Biology multiple choice items significantly functioned differentially among the students on basis of gender. The above result showed that WASSCE Biology multiple-choice items used in 2021 examinations contain test items that significantly functioned differentially for testees on the basis of gender. Generally, Logistic Regression comparing 2021 WASSCE Biology multiple choice items for male and female flagged nine items with significantly differential items functioning ($p < .05$) DIF. It was revealed that 6 items out of nine items representing 67% in 2021 WASSCE Biology multiple choice items, that displayed DIF favoured male while 3 items out of 9 items representing 33% were in favour of female. The corresponding hypothesis that there is no significant difference between male and female students on the percentage of items which functioned differentially in the 2021 WASSCE Biology multiple-choice items is rejected.

This finding is in line with study of Madueke and Casmir (2022), whose finding revealed that eight items in WAEC 2020 Biology multiple items out of fifty items test were indicated biased based on gender in which the exist DIF items favoured male and placed the female students at disadvantage. This finding disagreed with the findings of Alade, Aletan and Sokenu (2020), which indicated that the biased items favoured females over their male counterparts, ultimately placing males in a disadvantaged group.

The result in table 3 and 4 revealed that a total of nine items representing 18% of 2021 Biology multiple choice items differentially functioned for candidates from urban and rural schools. The above result showed that biology multiple choice test items used in 2021 WASSCE examinations, contain test items that significantly functioned differentially for testees on the basis of school location. Generally, Logistic Regression comparing 2021 WASSCE Biology multiple choice items for male and female flagged nine items with significantly differential items functioning ($p < .05$) DIF.

It was revealed that, six items out of nine items representing 67% in 2021 WASSCE Biology multiple choice items, that displayed DIF favoured rural while 3 items out of 9 items representing 33% were in favour of urban. The corresponding hypothesis that there is no significant difference between male and female students on the percentage of items which functioned differentially in the 2021 WASSCE Biology multiple-choice items is rejected. The finding is in line with the finding of Madueke and Casmir (2022), who reported that the test items of WAEC 2020 Biology multiple choice items revealed significant DIF between urban and rural students for eight items. Thus, flagged biased of the eight items test is therefore placed the urban students at disadvantage group. The finding also agrees with the finding of Bamidele and Abdulwasiu (2020), who reported that eight items were biased against rural and urban students of 2016 WAEC Biology multiple choice items. The finding disagreed with the findings of Alade, Aletan and Sokenu (2020), which found items that favoured rural school students over their counterparts in the urban schools.

Conclusion

It is important to note that while the current study and the previous research agree on the existence of DIF based on gender and school location, there may be differences in the specific items identified or the magnitude of the bias observed. The current study provides a more detailed list of items exhibiting DIF, which may offer a more comprehensive understanding of the issue compared to the earlier studies. This level of detail could potentially provide educators and test developers with more specific areas to address when designing assessments that are fair across gender and school location.

Recommendations

Based on the findings of the analysis of Differential Item Functioning (DIF) in WASSCE Biology multiple-choice examinations among secondary school students in North West Nigeria, the following recommendations were made:

1. The West African Examinations Council (WAEC) should establish a gender-balanced item review committee to scrutinize test items for potential gender bias before finishing the examination.
2. WAEC, in collaboration with the Nigeria Educational Research and Development Council (NERDC), should develop location-specific item banks that account for the differences in rural and urban educational experiences.

References

- Alade, O. M., Aletan, A., & Sokenu, B. S. (2020). Assessing the differential item functioning of 2018 WASSCE mathematics achievement tests in Lagos, Nigeria. *AJBSDR*, 2(2), 8-20.
- Alokan, F.B. (2010). Influence of sex and location on relationship between student problems and academic performance. *The Social Sciences (TSS)*, 5(4), 340 – 345

- Amuche, C. I. & Fan, A. F. (2014). An assessment of item bias using differential item functioning technique in NECO Biology conducted examinations in Taraba State Nigeria. *American International Journal of Research in Humanities, Art, and Social Sciences*, 95-100.
- Bamidele, A. F. & Abdulwasiu, A. R. (2020). Differential Item Functioning of 2015/2016 biology multiple choice items of Osun State joint promotion Examination. *Alhikma Journal of Education (AJE)*.7(1), 312-321.
- De Vaus, D. (2014). *Surveys in social research* (6th ed.). Routledge.
- FRN (2014). *National Policy on Education*. Federal Ministry of Education, Abuja.
- Holland, R., & Wainer, H. (1993). *Differential Item Functioning*. Routledge.
- Ibrahim, S. & Binta, A. (2019). Identification of school type of differential item functioning of 2010 senior secondary certificate (NECO) English language examination among students of Dawakin Kudu Education Zone, Kano State, Nigeria. *International Journal of Advanced Educational Research*, 4(3), 31-35
- Karami, H. (2012). An Introduction to Differential Item Functioning. *The International Journal of Educational and Psychological Assessment*, 11(2). 59-65.
- Khalid, M. N., & Glas, C. A. W. (2014). A scale purification procedure for evaluation of differential item functioning. *Journal of Measurement*, 50, 186-197.
- Lee, H. Y. (2012). Evaluation of two types of Differential Item Functioning in factor mixture models with binary outcomes. *The University of Texas at Austin educational and psychological measurement* 74(5), 831-858
- Madueke, U. O., & Casmir, N. E. (2022). Differential Item Functioning of WAEC senior secondary certificate examination Biology multiple choices. *BIJESS*, 9(1), 31-48.
- Obinne, A.D.E., & Amali, A.O. (2014). Differential Item Functioning: The Implication for Educational Testing in Nigeria. *International Review of Social Sciences and Humanities* 7(1), 52-65.
- Orluwene, G. W. & Queensoap, M. (2019). Use of Mantel Haenszel Differential Item Functioning in Detecting Item Bias in a Chemistry Achievement Test in four ethnic groups in Nigeria. *International Journal of Current Research*, 11(3), 2665-2670.
- Osadebe, P. U. & Agbure, B. (2019). Assessment of differential item functioning in social studies multiple choice questions in basic education certificate examination. *European Journal of Education Studies*, 6(8), 96-109
- Smith, J. (2025). Utilizing multi-stage sampling for representative and efficient research design. *Journal of Research Methods*, 15(3), 123-145.
- The Research Advisors (2006). *Sample size table*. <http://www.research.advisor.com>.