



IRT as Educational Solution in Testing for Building Resilient Future among Students in Delta State

Okagbare, Freedom (PhD)

Department of Guidance and Counselling, Delta State University, Abraka

okagbare.freedom84@gmail.com, freedom-okagbare@delsu.edu.ng

07067484793, 09121831782

Abstract

The study investigated IRT as educational solution in testing for building resilient future among students in Delta State. Two research questions and two hypotheses guided the study at 0.05 of significance. The study adopted ex-post facto design. The sample of this study consists of 600 (300 male and 300 female students) science students drawn from the 3 senatorial districts of Delta State. 2023 WASSCE physics paper 2 questions were used as instrument for data collection to demonstrate the items that differentially function through the item responses. The validity and reliability of the instrument was believed to be valid and reliable since WAEC is a standard examination body across selected English speaking West African Countries. The research questions were answered using L-R model while Independence Chi-square test was used to test the stated hypotheses using SPSS version 23 at 0.05 level of significance. The findings of the study indicated a high percentage of IRT DF in the 2023 WASSCE/SSCE physics multiple choice test response among male and female students, the study also indicated a high percentage of IR DF in the 2023 WASSCE/SSCE physics multiple choice test response among urban and rural students that can distort building of resilient future. In line with findings, it was recommended among others that Test should be constructed to accommodate same responses across groups to make the resilient future possible, test builders should delete and remove tests that are difficult and discriminate for the purpose of resilient society.

Keywords: IRT; DF; Location; Sex



Introduction

The accomplishment of any nation aiming for goal achievement depends on its groundwork. For a nation to attain sustainable future in all areas of life, the foundation of the educational sector must be effectively planned (Akemiyefa, 2024). This could be one of the reasons the United Nations (UN) collaborated with UNESCO to formulate global goals and policies that will enhance the living condition of its citizens in the world. Usman (2017) opined that these goals are to: Eradicate extreme poverty and hunger; Achieve universal primary education, promote gender equality and empower women; Reduce child mortality; Improve maternal health; Combat HIV/AIDS, malaria and other diseases; Ensure environment sustainability and develop a global partnership & development. Odili (2019) noted that the index used in measuring national development is to glance at the country's standing in the United Nations' (UN) seventeen (17) SDGs and these goals are set to be accomplished by 2030.

For this study, the inclusive and equitable quality education and promoting life-long learning opportunities for all which is one of the global goals will be our focus. This target according to Usman (2017) of the goal four as stipulated by SDGs include completion of free, impartial and quality primary and post primary education for all learners that will lead to significant and successful learning result, ensuring access for all men and women with affordable quality industrial, vocational and tertiary education and ensuring all students acquire knowledge and skills needed to promote sustainable development through education for sustainable life way, human rights, gender equality to enhance peace and security. This set goal can only be achieved when meaningful test that measure and tap the constructs of the goal four target through teaching and learning in the classroom situation. Akemiyefa, (2024) stated that in the teaching and learning process, skills, values, knowledge are transmitted from an instructor being the teacher to a learner or receiver and vice versa which are referred to an education. Ruggiero (2020) supported the above points by stating that students are involved in critical thinking, describing critical thinking as a mental readiness that formulate and proffer solution to a problem, making decision, or fulfill a desire to understand. As a result of the importance of education to the realization of goals, every creative government, both local and states have placed a premium on the education of its citizens by test that is covered in higher level of cognitive domain that will enhance their critical thinking skill and problem-solving ability.

However, these comprehensive and interdependent goals that are holistic are hunted and truncated future opportunities. Through the outcome of tests in schools and industries, the attainment of the SDGs objectives still remains a major challenge in Nigeria. It has also been observed that the test used in measuring these goals do not provide equal opportunity for all. This has resulted in a violation of the inclusive and equitable quality education and promoting life-long learning opportunities for all irrespective of location and background.

To develop an instrument that will build a resilient future, Item Response Theory (IRT) must be used. Item Response Theory (IRT) builds on Classical Test Theory (CTT) (Odili, 2021). Okagbare et al (2023) stated that Item Response theory (IRT) is sometimes called modern mental test theory (MMTT), is a type of latent trait model for the design, analysis, and scoring of tests, questionnaires in dichotomized format of right or wrong response that measured educational problems.

IRT model is used to challenge and quantify the given amount of mental traits inherit on the students. When the amount of mental trait in a student is high, he/she will be able to

get an item in test correctly but if the mental trait is low and do not have knowledge on the test item the examinee will fail the item thing that was given in the test. The introduced mental attribute brought the major differential response among students. The mental attributes are the problems that rest on the students; they are reading ability, writing ability, level of comprehension and communication skills. It is worthy of note that when these attributes are not checked they could affect the mark of a creating a resilient future. One big problem of IRT is the Differential item functioning

Okagbare et al (2023) defined DIF as students' variation of response on test item statement that measure a trait. An instrument that is designed to measure constructs/trait or subject-matter should not differentially function for two students who are of the level ability but from different subgroup of inhabitants. DIF can be classified into two, namely uniform and non-uniform DIF. When DIF moves in the same direction at all skill levels, this is known as uniform DIF. This intends that for uniform DIF, every one of the examinees in a single gathering have higher likelihood of performing preferable on the thing over every one of the examinees in the other gathering. For example, when groups are favoured at the various levels of ability namely high, moderate and low, uniform DIF is said to be present (Okagbare et al, 2023). On the other hand, non-uniform DIF occurs when an item favours one group at certain level of ability and other group at other levels of ability. Non-uniform DIF is displayed if for example, at high level of ability members of group A are favoured while at low ability level, members of group B are favoured

Location is another factor that might impact DIF between students with similar abilities in terms of subject matter. Location is a point of a geographical urban or rural area. The geographic area where the test taker or examinee travels to study and apply physics concepts is referred to as the location of the student. The study of Alordial (2015) also showed that test item differentially function among examinees in urban, semi-urban and rural area.

Ozdemir and Alshamrani (2020) explored IRT-based Differential Item (DI) and Test Functioning (TF) Methods to investigate the fairness of verbal communication tests across male and female learners. Using the IRT-based Lord's chi-square DIF method at the item level and the Mantel-Haenszel/Liu-Agresti differential test function (DTF) method at the test level, the results demonstrate that 6 EPT items display DIF throughout the test. Two of them demonstrate reading comprehension, four to the structural domain, and none to the compositional analysis methodologies. These findings suggest the presence of a content-specific DIF impact. Furthermore, two things/items demonstrate homogeneous DIF, one of which indicates DIF supporting male students and the others supporting female students in a minor to moderate DTF effect connected with sub-domains and the full test implies that DIF effects cancel each item to assure test level fairness.

Brew (2020) investigated DIF of WASSCE in mathematics, English language, integrated science and social studies subjects based on urban and rural metropolis and sex in Southern Ghana Between 2012-2016. The cross-sectional research design was used for the study. One research question and six hypotheses were developed. DIF identification approaches such as Mantel-Haenszel (MH), Logistic-Regression (LR), and IRT were utilized to capture items with DIF. The data revealed a substantial difference in item functioning based on sex (male and female). There was also a considerable difference in item functioning based on geography, as all three approaches found things that functioned differently in each of the five regions studied. There was a high level of items captured with DIF, the Logistic Regression (LR), Mantel-Haenszel, and three parameter latent trait IRT all agreed.

However, there are visible few studies in the existing literature that study DIF of items in examining bodies in Nigeria and overseas. No research work known to the researcher on IRT as educational solution in testing for building resilient future among students in Delta State. As a result, the purpose of this suggested study is to fill the resulting gap. To accomplish the goal of equal opportunity for all citizens in Nigeria that will create a resilient future there is need to address the relevance of IRT with respect to differential items functioning in all fields of studies with focus on physics multiple tests.

The teaching and learning in recent times have been faced with several problems like favorable learning atmosphere and learning behaviour of students. The problem of IR differential functioning may be attributed to difficulty in examination test items. Despite efforts by researchers to improve academic excellence in students, the persisted differential achievement is still found among students. These differential achievements could bring about downward trend of reference and focal groups in a body of knowledge. Responses and its functions and purpose for groups could be weak when IR differentially function. This will of course hinder the goal of building a resilient future for students. When item responses are the same for all students irrespective of background, location, ethnicity, and race resilient future will be achieved. Considering the high rate of hardship, kidnapping, banditry and political tussle in all area of the country or nation there is need for education stakeholders to use IR as educational solution in testing for building resilient future among students in Delta State schools.

The problem of this study is what is the percentage of IR differential function in multiple choice questions used by WASSCE/SSCE for students from urban and rural schools, and male and female student in Delta State?

Research Questions

The stated research questions directed the study.

1. What is the percentage of IR DF in the 2023 WASSCE/SSCE physics multiple choice test items among male and female students in Delta State?
2. What is the percentage of IR DF in the 2023 WASSCE/SSCE physics multiple choice test items among urban and rural students?

Hypotheses

The following hypotheses were tested in the study:

1. There is no significant difference of the percentage of IR DF in the 2023 WASSCE/SSCE physics multiple choice test items among male and female students
2. There is no significant difference of the percentage of IR DF in the 2023 WASSCE/SSCE physics multiple choice test items among urban and rural students

Methods

The study adopted casual-comparative of *expost-facto* design. The researcher considered this design suitable for the study because it facilitates the researcher to statistically decide the effects between the variables. The population used for the investigation is 57,875

SS students across the twenty-five (25) local government of Delta State from 491 public secondary schools in Delta state. The sample of this study consists of 600 (300 male and 300 female students) science students drawn from the 3 senatorial districts of Delta State. The study adopted the multiple stages to get the sample. At the first stage, Delta State was Statified into Northern zone, Central zone and Southern zone. Random Sampling Technique(SRST) was then applied to select three (3) Local Government Area from the three senatorial zones. At the second stage, 2 senior secondary schools in each of the Senatorial zones making a total of 6 senior secondary schools that were drawn for the study. In the final stage, the researcher used SRST by balloting to select hundred (100) students in each of the six (6) senior secondary schools drawn from Delta North, Delta Central and Delta South Senatorial District of Delta State making a total of 600 students used for the study.

The 2023 WASSCE physics paper 2 questions were used as instrument for data collection to demonstrate the items that differentially function through the item responses. The WASSCE test instruments contains 50 multiple choice test questions with one correct response and many distracters of A-D. The instruments were prepared by WAEC. Hence, validation had been taken care of. The WAEC physics content instruments was believed to be valid because WASSCE is a standardized examination taken at equal point in time across regions and some selected English Speaking West African countries. Reliability of the test was established using the Kuder- Richardson method. The instrument was administered to 20 groups of private secondary school students offering physics in certificate classes in Edo state and a value of 0.67 was obtained to ensure internal consistency of the item response.

The test was administered within the limited time as specified by the examination body. The correct response for an item was assigned 1 and incorrect response was assigned 0. The research questions was answered using L-R model while Independence Chi-square test was used to test the stated hypotheses using SPSS version 23 at 0.05 level of significance.

Result

Research Question 1: What is the percentage of IR DF in the 2023 WASSCE/SSCE physics multiple choice test items between male and female students?

Table 1: Logistic Regression Statistics to detect IR DF in the 2023 WASSCE/SSCE physics multiple choice test items between male and female students

Test items	B	SE	Wald	Df	Sig.	Exp (B)	95% C.I lower	Exp B upper	Disfavor group	Percentage of IR DF
1	-.073	.041	3.166	1	.075	.930	.858	1.007	No DIF	88 % (42%
2	.197	.042	22.511	1	.000	1.218	1.123	1.321	Dm	disfavor
3	-1.154	.062	347.213	1	.000	.315	.279	.356	Df	male while
4	.449	.044	103.081	1	.000	1.567	1.436	1.708	Dm	46%
5	-5.774	.707	66.756	1	.000	.003	.001	.012	Df	disfavor
6	-.016	.041	.151	1	.697	.984	.908	1.066	No DIF	female)
7	.289	.042	46.552	1	.000	1.335	1.228	1.450	Dm	
8	.178	.041	18.495	1	.000	1.195	1.102	1.296	Df	
9	.449	.044	103.081	1	.000	1.567	1.436	1.708	Df	
10	.250	.042	35.381	1	.000	1.283	1.182	1.393	Dm	
11	.482	.045	116.416	1	.000	1.620	1.484	1.768	Dm	
12	.336	.043	61.562	1	.000	1.399	1.286	1.521	Dm	
13	.550	.046	143.997	1	.000	1.733	1.584	1.896	Df	



14	.525	.045	133.726	1	.000	1.691	1.547	1.848	Df
15	.337	.043	61.893	1	.000	1.400	1.288	1.523	Dm
16	.136	.041	10.806	1	.001	1.145	1.056	1.242	Dm
17	.276	.042	42.843	1	.000	1.318	1.213	1.432	Dm
18	.451	.044	103.848	1	.000	1.570	1.439	1.712	Dm
19	-.164	.041	15.662	1	.000	.849	.783	.921	Df
20	.287	.042	46.014	1	.000	1.332	1.226	1.447	Df
21	-.159	.041	14.864	1	.000	.853	.786	.925	Df
22	-.158	.041	14.554	1	.000	.854	.788	.926	Df
23	-.408	.044	87.523	1	.000	.665	.610	.724	Dm
24	.026	.041	.403	1	.526	1.026	.947	1.112	No DIF
25	.239	.042	32.586	1	.000	1.270	1.170	1.379	Dm
26	.189	.042	20.638	1	.000	1.208	1.113	1.310	Df
27	.276	.042	42.843	1	.000	1.318	1.213	1.432	Dm
28	.478	.045	114.815	1	.000	1.614	1.478	1.761	Df
29	.146	.041	12.468	1	.000	1.157	1.067	1.254	Df
30	.218	.042	27.326	1	.000	1.244	1.146	1.350	Df
31	.137	.041	11.075	1	.001	1.147	1.058	1.244	Dm
32	.218	.042	27.326	1	.000	1.244	1.146	1.350	Df
33	.374	.043	75.048	1	.000	1.454	1.336	1.583	Dm
34	.234	.042	31.230	1	.000	1.263	1.164	1.371	Df
35	.218	.042	27.326	1	.000	1.244	1.146	1.350	Dm
36	.159	.041	14.864	1	.000	1.173	1.082	1.272	Df
37	.197	.042	22.511	1	.000	1.218	1.123	1.321	Dm
38	.262	.042	38.780	1	.000	1.299	1.197	1.411	Df
39	.237	.042	32.131	1	.000	1.268	1.168	1.376	Dm
40	.142	.041	11.901	1	.001	1.153	1.063	1.250	Df
41	.332	.043	60.340	1	.000	1.394	1.282	1.515	Df
42	.299	.042	49.845	1	.000	1.349	1.241	1.466	Dm
43	.341	.043	63.416	1	.000	1.406	1.293	1.530	Df
44	-.055	.041	1.769	1	.184	.947	.874	1.026	No DIF
45	.041	.041	1.005	1	.316	1.042	.962	1.129	No DIF
46	-.033	.041	.637	1	.425	.968	.893	1.049	No DIF
47	-.151	.041	13.342	1	.000	.860	.793	.932	Df
48	-.204	.042	24.066	1	.000	.815	.751	.885	Dm
49	-.081	.041	3.935	1	.047	.922	.851	.999	Dm
50	.189	.042	20.638	1	.000	1.208	1.113	1.310	Df

0.05 Df = disfavor female, Dm = disfavor male, Df = degree of freedom

B = slope or regression coefficient called beta which the predicted change in the log odd. It can be negative or positive slope.

S.E = Standard error which involve the computation of statistical significance

Wald = It is the test statistics used to estimate the statistical significant of the parameter following the chi-square distribution with the degree of freedom of 1

Exp (B) = It is called the odd ratio. It is the ratio of probability correct response or incorrect response. For one unit change in predictor, there is exponential change in the probability of the outcome.

95% C.I = It show the relationship between the predictors and the outcome by indicating that the independent variable that have got the significant impact contribution to the choice of the target group in which is the right or wrong in the dependent variable

Table 1 shows the items response that relate to gender of students (male and female), identified by logistic regression method using SPSS version 23. Out of the fifty items in physics multiple choice, 44 responses indicated IRT DF addressing 88%. Therefore, percentage of IRT DF in the 2023 WASSCE/SSCE physics multiple choice test response between male and female students was high. That is 42% of 21 things disfavoured male while 46% of 23 things disfavoured female understudies or examinees.

Research Question Two: What is the percentage of IRT DF in the 2023 WASSCE/SSCE physics multiple choice test items between urban and rural students?

Table 2: Logistic Regression Statistics to detect IR DF in the 2023 WASSCE/SSCE physics multiple choice test items between urban and rural students

Test items	B	SE	Wald	Df	Sig.	Exp (B)	95% C.I lower	Exp B upper	Disfavor group	Percentage of DIF
1	-.031	.039	.628	1	.428	.970	.899	1.046	No DIF	78 % (22% disfavor urban examinee while 56% disfavor rural examinees)
2	.265	.040	44.142	1	.050	1.303	1.205	1.409	No DIF	
3	-.136	.039	12.183	1	.000	.873	.809	.942	Du	
4	.531	.043	150.445	1	.000	1.700	1.562	1.851	Dr	
5	-.479	.042	127.680	1	.061	.619	.570	.673	No DIF	
6	.255	.040	41.198	1	.000	1.291	1.194	1.395	Du	
7	.485	.043	130.055	1	.000	1.624	1.494	1.765	Du	
8	.444	.042	112.272	1	.000	1.559	1.436	1.692	Dr	
9	.747	.048	244.267	1	.000	2.110	1.922	2.317	Dr	
10	.747	.048	244.267	1	.000	2.110	1.922	2.317	Dr	
11	1.040	.056	341.391	1	.000	2.829	2.534	3.159	Dr	
12	1.381	.070	394.269	1	.000	3.980	3.473	4.562	Dr	
13	.763	.048	250.745	1	.000	2.144	1.951	2.357	Du	
14	.540	.043	154.642	1	.000	1.716	1.576	1.869	Du	
15	.422	.042	103.071	1	.000	1.525	1.406	1.655	Du	
16	.249	.040	39.288	1	.000	1.282	1.186	1.386	Du	
17	.295	.040	54.086	1	.000	1.343	1.241	1.453	Du	
18	.451	.042	115.292	1	.000	1.570	1.446	1.704	Du	
19	.049	.039	1.628	1	.202	1.051	.974	1.133	No DIF	
20	.349	.041	73.521	1	.079	1.417	1.309	1.535	No DIF	
21	-.058	.039	2.214	1	.137	.944	.875	1.018	No DIF	
22	-.055	.039	1.990	1	.158	.947	.878	1.021	No DIF	
23	-.326	.040	64.949	1	.000	.722	.667	.782	Dr	
24	.119	.039	9.413	1	.002	1.127	1.044	1.216	Dr	
25	.331	.040	66.746	1	.000	1.392	1.286	1.507	Dr	
26	.052	.039	1.778	1	.182	1.053	.976	1.136	No DIF	
27	.208	.039	27.930	1	.000	1.231	1.140	1.330	Dr	
28	.404	.041	95.532	1	.000	1.498	1.381	1.624	Dr	
29	.094	.039	5.827	1	.016	1.098	1.018	1.185	Dr	
30	.139	.039	12.725	1	.000	1.149	1.065	1.240	Dr	
31	.154	.039	15.605	1	.000	1.167	1.081	1.260	Dr	



32	.145	.039	13.843	1	.010	1.156	1.071	1.248	Dr
33	.285	.040	50.845	1	.000	1.330	1.230	1.439	Dr
34	.183	.039	21.867	1	.000	1.201	1.112	1.297	Dr
35	.186	.039	22.586	1	.000	1.205	1.116	1.301	Dr
36	.153	.039	15.304	1	.000	1.165	1.079	1.258	Dr
37	.138	.039	12.453	1	.000	1.147	1.063	1.239	Dr
38	.258	.040	42.169	1	.000	1.295	1.198	1.400	Dr
39	.231	.040	34.258	1	.000	1.260	1.166	1.362	Dr
40	.432	.042	107.062	1	.000	1.540	1.419	1.671	Du
41	.261	.040	43.150	1	.000	1.299	1.201	1.404	Dr
42	.272	.040	46.668	1	.000	1.313	1.214	1.420	Dr
43	.252	.040	40.238	1	.000	1.286	1.190	1.391	Dr
44	-.058	.039	2.214	1	.117	.924	.875	1.018	No DIF
45	.076	.039	3.807	1	.051	1.079	1.000	1.164	No DIF
46	.044	.039	1.301	1	.254	1.045	.969	1.127	No DIF
47	-.113	.039	8.493	1	.004	.893	.827	.964	Dr
48	-.197	.039	25.189	1	.000	.821	.760	.887	Dr
49	-.150	.039	14.711	1	.000	.861	.798	.929	Du
50	.115	.039	8.719	1	.003	1.122	1.039	1.210	Dr

0.05 DU= disfavor urban DR= disfavor Rural

B = slope or regression coefficient called beta which the predicted change in the log odd. It can be negative or positive slope.

S.E = Standard Error which involve the computation of statistical significance

Wald = It is the test statistics used to estimate the statistical significant of the parameter following the chi-square distribution with the degree of freedom of 1

Exp (B) = It is called the odd ratio. It is the ratio of probability correct response or incorrect response. For one unit change in predictor, there is exponential change in the probability of the outcome.

95% C.I = It show the relationship between the predictors and the outcome by indicating that the independent variable that have got the significant impact contribution to the choice of the target group in which is the right or wrong in the dependent variable

Table 2 shows the responses that relate urban and rural students identified by logistic regression method using SPSS version 23. Out of fifty items in physics multiple choice, 39 responses indicated IRT DIF addressing 78%. Therefore, percentage IRT DF in the 2023 WASSCE/SSCE physics multiple choice test responses between urban and rural students are high. That means that 22% of 11 item responses disfavoured urban area and 56% of 28 item responses disfavoured rural students or examinees.

Testing of Hypotheses

Hypothesis One: There is no significant difference of the percentage of IR DF in the 2023 WASSCE/SSCE physics multiple choice test responses between male and female students

Table 3: Test of Independent Chi-Square Statistics Showing Percentage of IR DF for Gender

Gender	N	% of Gender-IR DF		X^2	df	Sig.	Decision
		Non-IR DF item	IR DF items				
Male	300	4	21	242.016	1	12.773	Not Significant
Female	300	2	23				
Total	600	6	44				

0.05

Table 3 shows a chi-square value of 242.016 and a P-value of 12.773. Testing at an alpha level of 0.05, the p-value of 12.773 is greater than the alpha level of 0.05. Hence, the null hypothesis is therefore recognised as accepted. As a result, there is no significant difference of the percentage of IRT DF in the 2023 WASSCE/SSCE physics multiple choice test responses between male and female students.

Hypothesis Two: There is no significant difference of the percentage of IR DF in the 2023 WASSCE/SSCE physics multiple choice test responses between urban and rural students

Table 4: Test of Chi-Square Statistics Showing Percentage of IR DF for Urban and Rural Students

Location	N	% of Location- DIF		X^2 value	df	Sig.	Decision
		Non-DIF Items	DIF Items				
Urban	300	5	11	136.344	1	21.822	Not significant
Rural	300	6	28				
Total	600	11	39				

0.05

Table 4 shows a chi-square value of 136.344 and a P-value of 21.822. Using 0.05 to test at an alpha level, the p-value of 21.822 is greater than the alpha level of 0.05. Hence, the null hypothesis is therefore accepted. As a result, there is no significant difference in the percentage of IR DF in the 2023 WASSCE/SSCE physics multiple choice test response between urban and rural students

Discussions

The study is an assessment of DIF in physics multiple test items used by WAEC among senior secondary school students in Delta State. The findings are discussed as follows:

Percentage of IR DF in the 2023 WASSCE/SSCE Physics Multiple Choice Test Responses between Male and Female Students

Data obtained from research question one and its hypothesis revealed a high percentage of IR DF in the 2023 WASSCE/SSCE physics multiple choice test response between male and female students. That is 42% of 21 responses disfavoured male while 46% of 23 responses disfavoured female examinees. This finding therefore implied that, when test



responses are not of equal standing, building a resilient future will be challenging. Accordingly these could be the reasons that impacted the examinees reactions to getting the responses effectively or fruitlessly. The finding of this review is in connection with the investigation of Omorogiuwa and Iro-Aghedo (2016) who analyze DIF in NABTEB science subjects test items and identified high DIF among students' tests in the subjects. The study is also in agreement with Ahmadi and Bazvand (2016) who identified DIF among male and female examinees in national field test.

Percentage of IR DF in the 2023 WASSCE/SSCE Physics Multiple Choice Test Responses between Urban and Rural Students

Data obtained from research question two and its hypothesis indicated a high percentage of IR DF in the 2023 WASSCE/SSCE physics multiple choice test response between urban and rural students. Therefore, percentage IR DF in the 2023 WASSCE/SSCE physics multiple choice test responses between urban and rural students were high. That means that 30% of the 15 item responses disfavoured students from urban area and 50% of the 25 item responses disfavoured students from rural area. The finding of this study is in relation with study of Felder, Mohr, Dietz and Ward (2014) who found out those urban students enjoy greater success than rural students. On the other hand, the findings of this study disagree with Brew (2020) whose findings revealed a considerable difference in item functioning based on urban and rural area. This implies that items used in assessing or determining ability of student has ingredient of biasness that disadvantaged the rural school examinees and Favors the urban schools' students.

Conclusion

Due to the based foundation of the findings/results, the following conclusions were made. There were high percentage of IR DF in the 2023 WASSCE/SSCE physics multiple choice test response between male and female students, high percentage of IR DF in the 2023 WASSCE/SSCE physics multiple choice test response between urban and rural students that can distort building of resilient future

Recommendations

In respect of the findings/results and conclusion, the following recommendations are made:

1. Test should be constructed to accommodate same responses across groups to make the resilient future possible
2. Test builders should delete and remove tests that are difficult and discriminate for the purpose of resilient society
3. Teachers and other stakeholders in every field of study should be trained on the principles of IRT for the realization of a renewed society.



References

- Akemieyefa, S. (2024). *Application of 2-parameter latent trait theory in development of Mathematics Achievement Test for Senior Secondary Schools in Bayelsa State*. Unpublished Ph.D thesis of Delta State University, Abraka, Nigeria.
- Ahmadi, A., & Bazvand, A. D. (2016). Gender differential item functioning on a national field-specific test: The case of PhD entrance exam of TEFL in Iran. *Iranian Journal of Language Teaching Research*, 4(1), 63-82.
- Alordiah, C. O. (2015). *Comparison of index of Differential Item Functioning Under the methods of Item Response theory and classical test theory in Mathematics*. An unpublished Ph.D thesis of Delta State University, Abraka, Delta State, Nigeria.
- Brew, A. R. (2020). *Differential Item Functioning Of West African Senior School Certificate Examination in Core Subjects in Southern Ghana*. Unpublished PhD Thesis, university of cape coast <https://erl.ucc.edu.gh/jspui>.
- Felder, R. M., Mohr, P. H., Dietz, E. J., & Ward, B. L. (2014). A longitudinal study of engineering student performance and retention: Differences between students from rural and urban backgrounds. *Journal of Engineering Education*, 83(3) 209-217.
- Odili, J. N. (2019). A seminar paper on Advance measurement and evaluation.
- Odili, J. N. (2021). *Overview of Item Response Theory (IRT) and Differential Item Functioning*. Unpublished Ph.D Lecture Note, Delta State University, Abraka, Delta State, Nigeria.
- Odili, J. N. (2021). *Overview of Item Response Theory (IRT) and Differential Item Functioning*. Unpublished Ph.D Lecture Note, Delta State University, Abraka, Delta State, Nigeria.
- Okagbare, F., Ossai, P. A. U. & Osadebe, P. U. (2023). Assessment of Differential Item Functioning in Physics Multiple Choice Items Used By WAEC among Senior Secondary School Students in Delta State. *West Africa Journal of Interdisciplinary Research*, Vol. 1, Issue 2. www.ijaar.org
- Omorogiuwa, K. O., & Iro-Aghedo, E. P. (2016). Determination of differential item functioning by gender in the National Business and Technical Examinations Board (NABTEB) 2015 mathematics multiple choice examination. *International Journal of Educational learning and development*, 4(10), 25-35.
- Ozdemir, B., & Alshamrani, A. H. (2020). Examining the Fairness of Language Test Across Gender with IRT-based Differential Item and Test Functioning Methods. *International Journal of Learning, Teaching and Educational Research*, 19(6), 27-45.
- Ruggiero, V. R. (2020). *Beyond feelings: a guide to critical thinking*. Library of Congress Cataloging-in-Publication, Revised 9th Ed.
- Usman, D. A. (2017). *History of Early Childhood Education in Nigeria: An Overview*. Basic Education in Nigeria. West and Solomon Publishing Coy Ltd, Onitsha