Assessing the differential item functioning of 2018 WASSCE mathematics achievement tests in Lagos State, Nigeria

O. M. Alade, S. Aletan, B. S. Sokenu
Department of Educational Foundations (with Educational Psychology), University of Lagos, Nigeria

Abstract

This study explored the Differential Item Functioning (DIF) of 2018 West Africa Examination Council's Mathematics Objectives Tests Items in Lagos, Nigeria. The research design used for the investigation is a descriptive survey design. The population included all Senior Secondary Three (SS3) students who enrolled for the 2020 West Africa Senior Secondary Certificate Examination (WASSCE) in Lagos State. Multistage sampling procedure was used to select 1334 students from eighteen secondary schools (three schools from each educational district). Three research questions guided the study. The research questions were subjected to item differential functioning analysis using BILOG MG model. Results demonstrated that six items out of the 50 items function differentially in regard to gender. The results also showed that 15 items out of the 50 items work differentially in regard to location, while all the 50 items function differentially with regard to school ownership. The study uncovered that item analysis using item response theory approach isn't adequate to pass judgment on the nature the test, it is necessary that the item bias is also estimated. It was thus recommended that examination bodies and test developers should include item bias during item analysis.

Keynotes: WASSCE, Differential Item Functioning, Mathematics, Multiple Choices, IRT

Introduction

Education is a fundamental mean for national development. It empowers the beneficiaries to work adequately for the good of the society to which they belong. The social and financial advancement of a nation has also been directly linked with students' academic performance (Alade, Aletan & Sokenu, 2018). This is because students' academic accomplishment assumes an important role in producing quality graduates who will become credible instruments for the nation's financial and social advancement.

The outcome of any educational pursuit (secondary education inclusive) is always weighed against the performance of the learners (Alade, Kuku & Osoba, 2017). Teachers' effectiveness is also measured in terms of their students' academic performances. Ali, Jusoff, Ali, Mokhtar, and Azni (2009) opined that scholastic outcome has always been used as yardstick for determining success or otherwise of school years.

The importance of mathematics in the technological development of any nation has been emphasized by various scholars and authors (Agnes, Anthony & Julie, 2009).

Mathematics is a subject that every student at the primary and secondary school levels is expected to offer. Its importance made the Federal Government of Nigeria to make Mathematics a core subject at both primary and secondary education levels (Federal Republic of Nigeria, 2013). In spite of the important role Mathematics plays in everyday life, it has remained one of the subjects students find difficult to pass in Nigerian schools. The pattern of students' performance in Mathematics in Secondary School Certificate Examination particularly before 2012 confirms this affirmation, However, the performance is fluctuating and not stable (see table l). The current Nigerian educational system is by all accounts accomplishing the ideal instructive objectives and destinations as there are perceptible confirmations of progress in students' academics accomplishment particularly in Mathematics as proved by the pattern of results in public examinations since

2013. There is in this manner the need to guarantee that the improvement is continued.

Researcher like Aliyu (2015), identified Mathematics as the bedrock of Science and Technology. Adeyegba in Ojerinde, (2013) saw that there is not really any area of science that doesn't utilize Mathematical ideas to clarify its own ideas, hypotheses or models. In fact, the most dependable advancement of human came through scientific controls including the utilization of numbers. The performance of students in Mathematics at the West African Senior School Certificate Examination (WASSCE) level in Nigeria is gradually becoming encouraging. Though there is a slight improvement, however, the failure rate is still high. For instance, in the last three years 2016, 2017 and 2018 (as shown in table l), 309,526, 160,623 and 220,804 candidates had passes, while 131,755, 44,874 and 77,009 failed the subject respectively. The implication of this is

Table 1: Statistics of Performance in Mathematics between 2008 and 2018 in Nigeria

SUBJECT MATHEMATICS	YEAR	TOTAL ENTRY	TOTAL SAT	TOTAL CREDIT 1-6	TOTAL PASS 7-8	FAIL
	2008	1292890	1268213 (98.09)	726398 (57.28)	302266 (23.83)	218618 (17.24)
	2009	1373009	1348528 (98.22)	634382 (47.04)	344635 (25.56)	315738 (23.41)
	2010	1331374	1306535 (98.13)	548065 (41.95)	363920 (27.85)	355382 (27.20)
	2011	1540141	1508965 (97.98)	608866 (40.35)	474664 (31.46)	421412 (27.93)
	2012	1695878	1658357 (97.79)	838879 (50.58)	478519 (28.86)	298742 (18.01)
	2013	1686990	1656527 (98.19)	1897655 (54.18)	462176 (27.90)	245263 (14.80)
	2014	1655794	1632377 (98.59)	1011608 (61.97)	357555 (21.90)	211941 (12.98)
	2015	1602362	1581420 (98.69)	901845 (57.02)	425628 (26.91)	219759 (13.89)
	2016	1552169	1536643 (98.99)	1056045 (68.72)	309526 (20.14)	131755 (8.57)
	2017	1565106	1550348 (99.05)	1276782 (82.35)	160623 (10.36)	44874 (2.89)
	2018	1576465	1563457 (99.17)	1208457 (77.29)	220804 (14.12)	77009 (4.92)

Source: The West African Examinations Council (WAEC), Test Development Division, Ogba Lagos
(The numbers in the parenthesis represent the percentages ofthe total entry per year that sat for the examination,

that such students will not be able to secure admission into tertiary institutions or get a job. have credit, have passes and failure)

According to House & Telese, (2008), various researches had been undertaken to find ways of improving Mathematics achievement and determine the factors influencing Mathematics' learning and performance. The identified factors include among others motivational orientation, self-esteem/self-efficacy, lack of adequate preparation, shortage of qualified teachers, lack of good school environment and infrastructural facilities. (Aremu & Sokan 2003), students poor attitude towards Mathematics (Bolaji, 2005) and poor teaching methods adopted by teachers (National Mathematics Centre, NMC, 2009). To improve performances, in Mathematics, many interventions have been initiated. Prominent among the interventions are the Lagos Eko Secondary Education project, 2004 — 2017 and the NMC's Mathematics Improvement Programme (MIP) aimed at creating new teaching methodologies to improve students' performance in mathematics. Despite the interventions, the observed gradual performance persisted (as revealed in table l) as evident in the fluctuating result of candidates in WASSCE's Mathematics after the introduction of these interventions.

One of the areas of challenge may be the observation that examiners are often faced with challenges of how to assess students in ways to obtain fair scores by reducing item difficulty especially in Mathematics (Olonode, 2016). As shown by Rover (2005), a fair and equitable test is one which allows all the testees equal opportunity to exhibit the aptitudes and information which they have obtained and which are applicable to the motivation behind the test. The matter of test fairness also brings forth the issue of Differential item functioning.

Differential item functioning (DIF) is a distinction in extent of right responses among respondents of the same ability in two groups. A test that shows DIF is one that is unjustifiable to a subgroup of everybody in which it is being used. DIF happens when two groups (reference group and focal group) that are similar in terms of their relevant

Hence, more efforts are required to reduce the failure rate.

knowledge and skills perform differently in an item (Umoinyang, 2011). It is a risk to the validity of the test and invalidates its interpretation (Pido, 2012). It happens when examinees of a similar capacity don't have equivalent likelihood of answering an item correct (Rover, 2005). This may emerge due to the gender, school ownership and school location of the examinees (Pedrjita, 2009).

Gender describes the social relation between men and women and the way it is socially constructed by society. It interacts with other variables such as class, ethnicity, religion and even school location (Chukwudi, 2019). School location describes the community in which a school is established. A school can be established either in urban or rural areas depending on the interest and vision of the owners. School ownership refers to the financing and the management/administration of the school system. Until the recent time, government had been the sole proprietor of the schools, both in the urban and rural areas, but currently private proprietors, including religious organizations, other nongovernmental organizations and entrepreneurs are involved in the founding and management of schools. Each school evaluates her students' using tests constructed either by the teachers or public examination bodies. Each test should be developed to provide fair and accurate estimate

of the ability of all examinees in the population of the test irrespective of their gender, school location and school ownership. Given the spread of candidates covered by the West Africa Secondary School Certificate Examination (WASSCE) in Lagos State, it is useful to analyze the 2018 WASSCE Mathematics multiple-choice test items for possible differential item functioning. This will guide writers and the examination body to improve the quality of Mathematics test in Lagos State. DIF assessment is a procedure that can be used in test improvement to ensure that the items are not bias (Schnipke, Roussos and Pashley, 2000).

There are two kinds of DIF, specifically uniform and non-uniform differential item functioning.

Uniform DIF happens when a group does better than another group on all ability levels. That is, all students from a group perform better than all students from the other group who are at the same ability level. At every ability level, the test item is simpler for one group than it is for another. For instance, the item favours every male respondent regardless of ability. On the other hand, for non-uniform DIF, individuals from a group are favoured up to a level on the ability scale and from that point on, the relationship are switched (Bachman, 2004). That is, there is an association among group and ability level. For instance, the item favours male respondents of low ability while favoring female respondents of higher ability (Bachman, 2004).

Hamzah and Hanan (2008) carried out a study on sex-related DIF analysis of Third International Mathematics and Science study (TIMSS). They found that the averages of male respondents' scores in Mathematics were greater than their females' counterparts. This implies that males did better than female respondents in the Mathematics multiple choice items. In Adedoyin (2010), investigation on gender-biased items in Mathematics examination, he discovered that 5 items were gender-biased out of the 16 test items that fitted the three parameter logistics model (3PL) of IRT statistical analysis.

A study by Madu (2012) on analysis of gender related DIF in Mathematics multiple-choice items showed significant gender differential functioning. This implies that the test contained items that measured different things for male and female examinees with the same Mathematics ability.

Literature provides research reports which suggest that tests could function differentially for subgroups defined by location. Studies by Amuche and Fan (2014) and Mokabi and Adedoyin (2010) have detailed the presence of differential item functioning among urban and rural students. Uruemu and Adams (2013) detected items that have DIF against subgroup of urban and rural school students when they conducted a study on DIF Method as an item bias indicator using logistic regression. Their study revealed that eight items showed DIF out of the sixty items of NECO Economics examination question paper that was examined. Their findings agreed with Tremblay, Ross and Berthelot (2001), who found that rural students enjoy lesser success than urban students. However, their findings disagreed with Lee and McIntire (2001), whose discoveries uncovered that the performance of the urban and rural students has no significant difference.

Uruemu's and Adams' (2013) finding suggests that items utilized in surveying students' ability have components of bias that distraught the rural school students and supported the urban schools' students. Few research studies have been done on the differential item functioning of items due to school ownership. Amuche and Fan (2014) using regression analysis distinguished items that have DIF against private and public-school students. Their investigation revealed that ten items exhibited DIF out of the 60 items in NECO Biology assessment question paper for 2012. Six items favoured private school students, while the public-school students were disadvantaged. Moreover, the remaining 4 items bolstered public schools more than the private schools. Hogan (2010) carried out a study on analysis of location context effects on DIF on NECO Mathematics objective test items using transform item difficulty (TID 45^{0}) procedure.

His discoveries demonstrated that (TID 45^{0}) hailed I l items as functioning significantly different, 3 for urban schools and 8 for rural schools.

In the assessment of students' performance at the SSCE mathematics, two sets of achievement tests are used. They are; Objective test and Essay Test. In the Essay tests, students are free to choose a required number of items to respond to from the list of items in the tests' booklet. This freedom of choice enables students to attempt questions they consider themselves competent enough to answer. This makes students' test scores' comparison a very difficult task. In contrast, examinees are

required to attempt all items contained in the multiple-choice tests, thereby providing a level ground for examinees' test scores' comparison. Multiple choice tests assess students' proficiency without necessarily elongating testing time. More importantly, it is highly objective when it comes to scoring of examinee's responses, (Oguoma, Metibemu & Okoye, 2016).

In this study, therefore, the 2018 Mathematics multiple choice test of WASSCE (SC) was emphasized. The study analyzed the scores from the 2018 WASSCE Mathematics to determine its fairness to students.

This research was anchored on the Item response theory (IRT). IRT lays on two essential proposals: a) the dimensional structure of the test items. b) The local independence assumption that states that an examinee's response to one item doesn't influence his/her responses to some other items in the test. This means that only the testee's ability and characteristics of the test items can influence performance of the testees. This theory is relevant to the study because it is used for construction of items, test calibration and in addressing item bias.

Statement of the Problem

It is compulsory for a candidate to have a credit pass in Mathematics in WASSCE before such candidate can be offered admission to study any course in Nigeria tertiary institutions. Unfortunately, it is considered very difficult and challenging by most candidates, thus leading to many students' failure to have a credit pass in the subject; however, the situation is gradually improving. Several factors were advanced for the inconsistent scholastic performance experienced in WASSCE's mathematics generally. The means by which Mathematics is presented to learners, the abstract nature of the subject and poor quality of instructional methodology and instructional materials deployed by teachers may account for it. Other observations trail from non-conducive learning environment, poor attitude of learners to Mathematics, an obvious lack of study habit among students and poorly constructed test items.

Some efforts made towards solving the identified problems seem to be yielding little result as students' performance has not shown significant steady improvement (see table l).

These endeavors incorporated the Mathematics improvement ventures which expect to improve Mathematics instructors at the primary, secondary and tertiary levels of training. In spite of these endeavors, accessible insights show that some students still fail to have credit pass in the subject.

This therefore calls for further studies into other factors that may be responsible for these fluctuations in the students' performance in Mathematics. This study was carried to determine the effect of DIF on students' performance in Mathematics in the WASSCE.

Research Questions

The accompanying questions guided the investigation.

l. What extent does the test items in Mathematics multiple-choice paper in the 2018 WASSCE (SC) function differentially due to sex (or gender)?
2. To what degree do the test items in Mathematics-multiple choice paper in the 2018 WASSCE (SC) function differentially due to school location?
3. To what extent do the test items in Mathematics-multiple choice paper in the 2018 WASSCE (SC) function differentially due to school ownership?

Method Research Design

The research design used was a descriptive survey design. The design is suitable because it is capable of studying large and small populations by selecting and studying samples chosen from the population to discover the relative incidence, distribution and interrelations of sociological and psychological variables (Ilogu, 2005).

Population

The populace for the investigation comprised of all the 48,651 SS3 students in both public and private secondary schools in Lagos who enrolled for 2020 WASSCE.

Sampling Techniques

Multistage sampling process was adopted to select 1334 students from eighteen schools, three schools from each of the six educational districts; comprising of public, private, urban and rural areas of Lagos State. Stratified random sampling method was first used to select six local government areas (one per educational district), while simple random sampling technique was later used to pick three schools from each of the educational districts. The Ss3 students in the selected schools were used for this study. The distribution of the respondents is as shown in Table 2.

Table 2: Distribution of Respondent across School Location and School Type

	School Type School		
Private	Public Total location		
Urban	4	5	9 5 9
Rural	5		4 9
Total	9		918

Instrumentation

Research instrument used for the study was the adopted 50 multiple-choice 2018 WASSCE (SC) Mathematics items. Each of the items has options A to D, with one correct option and three distracters. The test content covered the senior secondary school's (SSI-SS3) Mathematics scheme in Lagos State. The items were deemed reliable and valid because they had been validated by the West Africa Examination Council (WAEC).

Data Collection Procedure

The data collection was done between August and September 2019. The researchers spent one day in each of the sampled schools for data collection. Prior to the administration of the instrument, the researchers visited the sampled schools for permission to administer the test to their students and to establish rapport with the Mathematics teacher-s and examinees. The students were counselled to take the test as serious as possible as prizes would be given to the best three students. This was necessary in order to sustain the interest and the participation of the students in the test. The students were given one and half hours (90 minutes) to complete the instrument.

Analysis of Data

After retrieving the instrument from the examinees, the researchers scored the items according to the prepared scoring key. The correct option was scored ,l" while the incorrect option was scored "O". The data was coded into SPSS; passed through cleaning process to remove outliers among the responses. The data was properly edited and saved into a format that can be recognized and accepted by BILOG program to run the analysis.

Data was analyzed using Phase 2 module of BILOG MG. Item with a difference greater than ±0.5 indicates the presence of DIF.

Results

Research Question 1:- What extent does the test items in Mathematics multiple-choice paper in the 2018 WASSCE (SC) function differentially due to sex?

The results of the analysis are presented as follows:

SIN	DIF Male	Female	Difference	Remark
MTOI	-1.134	-1.018	-0.116	No DIF
MT02	-1.36	— I . 464	0.104	No DIF
MT03	-0.851	-0.717	-0.134	No DIF
MT04	-0.59	-0.748	0.158	No DIF

MT05	-1.24	-1.092	-0.148	No DIF
MT06	-0.562	-0.368	-0.194	No DIF
MT07	-0.601	-0.537	.064	No DIF
MT08	-0.973	-0.912	-0.061	No DIF
MT09	-0.937	-0.7	-0.237	No DIF
MTIO	-1.044	-0.968	-0.076	No DIF
MTH	-0.058	-0.017	-0.041	No DIF
MT12	-0.292	-0.233	-0.059	No DIF
MT 13	-0.498	-0.476	-0.022	No DIF
MT 14	-0.417	-0.448	0.031	No DIF
MT15	0.166	0.246	-0.08	No DIF
MT 16	-0.744	-0.582	-0.162	No DIF
MT 17	-0.566	-0.476	-0.09	No DIF
MT18	-0.629	-0.351	-0.278	No DIF
MT19	-0.065	-0.074	0.009	No DIF
MT20	-0.348	-0.246	-0.102	No DIF
MT21	o. 162	0.229	-0.067	No DIF
MT23	-0.295	-0.319	0.024	No DIF
MT24	-0.554	-0.966	0.412	No DIF
MT25	0.074	0.377	-0.303	No DİF
MT27	-1.86	-1.947	0.087	No DİF

MT28	17.577	21.179	-3.602	DİF	Favoured Female
MT29	-0.184	-0.604	0.42	No DİF	
MT30	0.116	0.246	-0.13	No DİF	
MT31	5.329	1.417	3.912	DİF	Favoured Male
MT32	-ı .214	-1.3	0.086	No DİF	
MT33	-0.196	-().094	-0.102	No DİF	
MT34	0.042	0.325	-0.283	No DİF	
MT35	10.075	5.358	4.717	DİF	Favoured Male
MT36	—O 99	_ 0.394	-0.005	No DİF	
MT37	0.423	1.122	-0.699	DİF	Favoured Female
MT38	0.581	0.452	0.129	No DİF	
MT39	0.331	O. 646	-0.315	No DİF	
MT40	-ı .652	-1.959	0.307	No DİF	
MT42	-0.28	-0.337	0.057	No DİF	
MT43	-2.495	-1.078	-1.417	DİF	Favoured Female
MT44	11.237	11.49	-0.253	No DİF	
MT45	-ı .249	-1.25	0.001	No DİF	
MT46	7.91	9.126	-1.216	DİF	Favoured Female
MT47	-3.169	-3.498	0.329	No DİF	
MT48	-2.537	-2.496	-0.041	No DİF	
MT49	-ı .908	-ı .805	-0.103	No DİF	

MT50	-3.436	-3.06	-0.376	No DİF

Table 3 shows the İRT DİF statistics on examined items' performance with respect to respondents' gender. The result showed that Six (6) items out of 50 have group difficulty differences of +0.5. Out ofthe six (6) items, 2(33.3%) favoured male while 4(66.67%) favoured female.

MTOI	-1.108	-ı .436	0.328	No DİF
MT02	-1.461	-1.782	0.321	No DİF
MT03	-0.785	-1.142	0.357	No DİF
MT04	-0.656	-1.019	0.363	No DİF
MT05	-1.199	-1.514	0.315	No DİF
MT06	-0.478	-0.782	0.304	No DİF
MT07	-0.555	-0.926	0.371	No DİF

Research degree | MT08 | -0.986 | -1 .256 | 0.27 | No DİF | Question 2:- To What do the test items in Mathematics-multiple choice paper in the 2018 WASSCE (SC) function differentially due to school location?
MT09 | -0.824 | -1.165 | 0.341 | No DİF

Table 4: İRT Analysis of DİF with respect to Location on the selected 50 2018 WASSCE (SC) Mathematics items

S/N	Urban	Rural	DİF Difference	Remark	
MTIO	-1.037	-1.351	0.314	No DIF	
MTII	-0.019	-0.333	0.314	No DIF	
MT12	-0.21 1	-0.672	0.461	No DIF	
MT13	—O. 534	-0.704	0.17	No DIF	
MT14	-0.422	-0.76	0.338	No DIF	
MT15	0.294	-0.247	0.541	DIF	Favoured Urban

MT16	-O. 702	-0.95	0.248	No DIF	
MT17	-O. 623	-0.673	0.05	No DIF	
MT18	—O. 504	-0.82	0.316	No DIF	
MT19	-O. 027	-0.431	().404	No DIF	
MT20	-0.289	-0.628	0.339	No DIF	
MT21	0.234	-0.18	0.414	No DIF	
MT23	-0.362	-0.484	o. 122	No DIF	
MT24	-0.651	-1.299	0.648	DIF	Favoured Urban
MT25	0.222	-0.091	0.313	No DIF	
MT27	-1.667	-2.882	1.215	DIF	Favoured Urban
MT28	19.11	18.048	1.062	DIF	Favoured Urban
MT29	-0.43	-0.564	o. 134	No DIF	
MT30	o. 154	-0.003	o. 157	No DIF	
MT31	3.92	4.3	-0.38	No DIF	
MT32	-1.35	-1.566	0.216	No DIF	
MT33	-0.119	-0.484	0.365	No DIF	
MT34	0.2	-0.205	0.405	No DIF	
MT35	7.445	10.587	-3.142	DIF	Favoured Rural
MT36	-0.419	-0.647	0.228	No DIF	
MT37	0.812	0.252	0.56	DIF	Favoured Urban
MT38	0.552	0.324	0.228	No DIF	
MT39	0.436	0.303	o. 133	No DIF	
MT40	-1.913	-2.118	0.205	No DIF	
MT42	-0.824	0.53	-1.354	DIF	Favoured Rural
MT43	—5.464	4.7	-10.164	DIF	Favoured Rural
MT44	11.027	15.905	-4.878	DIF	Favoured Rural

Table 4 shows the İRT DİF statistics on examined item performance with respect to respondent location. The results showed that fifteen (15) items out of 50 have group difficulty differences of+0.5. Out of the fifteen (15) items, 9(60%) favoured urban While 6(40%) favoured rural.

Research Question 3:- To what extent do the test items in Mathematics-multiple choice paper in the 2018 WASSCE (SC) function differentially due to school ownership?

Table 5: İRT Analysis of DİF with respect t o School Type on the selected 50 2018 WASSCE (SC) Mathematics items

S/N	Private	Public	Difference		Remark
MTOI	-1.273	-2.439	1.166	DİF	Favoured Private
MT02	-ı .47	-2.765	ı .295	DİF	Favoured Private
MT03	-0.988	-2.188	1.2	DİF	Favoured Private
MT04	-0.806	-2.126	1.32	DİF	Favoured Private
MT05	-1.495	-2.579	I . 084	DİF	Favoured Private
MT06	-0.647	-1.894	ı .247	DİF	Favoured Private
MT07	-0.769	-ı .979	1.21	DİF	Favoured Private
MT08	-1.098	-2.331	ı .233	DİF	Favoured Private
MT09	-o. 986	-2.287	1.301	DİF	FavouredPrivate
MTIO	-1.189	-2.394	ı .205	DİF	Favoured Private
MT11	-0.309	-ı .402	ı .093	DİF	Favoured Private
MT12	-0.445	-ı .681	ı .236	DİF	Favoured Private
MT13	-0.667	-ı .905	1.238	DİF	Favoured Private
MT14	-0.678	-1.815	1.137	DİF	Favoured Private
MT15	-0.048	-1.11	ı .062	DİF	Favour edPrivate
MT16	-0.983	-2.023	1.04	DİF	Favoured Private
MT17	-o. 772	-ı .886	1.114	DİF	Favoured Private
MT18	-0.755	-l .879	1.124	DİF	Favoured Private
MT19	-0.238	-ı .505	ı .267	DİF	Favoured Private
MT20	-0.35	-ı .854	ı .504	DİF	Favoured Private
MT21	0.056	—I .204	1.26	DİF	Favoured Private
MT23	-0.564	-1.7	1.136	DİF	Favoured Private
MT24	-0.833	—2.124	1.291	DİF	Favoured Private

MT25	-0.087	-1.181	I .094	DİF	Favoured Private
MT27	-1.706	-3.316	1.61	DİF	Favoured Private
MT28	15.872	10.983	4.889	DİF	Favoured Private

Table 5 shows the İRT DİF statistics on examined item

MT29	-0.359	-l .944	1.585	DİF	Favoured Private
MT30	-0.155	-0.973	0.818	DİF	Favoured Private
MT31	2.507	0.822	ı .685	DİF	Favoured Private
MT32	-1.314	-2.524	1.21	DİF	Favoured Private
MT33	-0.324	-1.426	1.102	DİF	Favoured Private
MT34	0.093	—I . 434	ı .527	DİF	Favoured Private
MT35	7.39	2.951	4.439	DİF	Favoured Private
MT36	-0.594	-1.7	1.2	DİF	Favoured Private
MT37	0.443	-0.791	ı .234	DİF	Favoured Private
MT38	0.285	-0.877	1.162	DİF	Favoured Private
MT39	0.084	-0.869	0.953	DİF	Favoured Private
MT40	-1.724	-2.871	1.147	DİF	Favoured Private
MT42	— 12.441	7.798	-20.239	DİF	Favoured public
MT43	-1.37	-3.519	2.149	DİF	Favoured Private
MT44	9.939	7.309	2.63	DİF	FavouredPrivate
MT45	-7.838	-0.902	-6.936	DİF	Favoured public
MT46	6.312	-ı .653	7.965	DİF	Favoured Private
MT47	-14.096	-3.397	-10.699	DİF	Favoured public
MT48	-19.938	-2.592	-17.346	DİF	Favoured public
MT49	-7.712	-2.371	-5.341	DİF	Favoured public
MT50	-10.188	-4.594	-5.594	DİF	Favoured public

performance with respect to respondent school type. The result revealed that all the 50 items indicate group difficulty differences of +0.5. Out ofthe flfty (50) items, 44(88%) favoured private school while 6(12%) favoured public school.

Discussion of Discoveries

The study revealed that there is differential item functioning in some of the mathematics multiple-choice items in the 2018 WASSCE (SC) due to gender. This result showed that six (6) items out of 50 items of the 2018 WASSCE (SC) Mathematics multiple choice tests were biased with respect to gender. Four (4) items were biased against male, while two (2) items were biased against female. The items measured construct having something to do with gender of the students other than their ability in Mathematics. This finding is in agreement with the study of Chukwudi (2019), whose finding showed that 46 out of 60 items of the 2017 BECE Mathematics multiple choice test were biased with respect to gender. 35 items were biased against male, While II items were biased against female.

The study also revealed that there is significant differential item functioning in 2018 WASSCE (SC) Mathematics multiple choice test due to school location for 15 items and there is no significant DİF for 35 items due to location. These showed that 15 out of the 50 items of the 2018 WASSCE (SC) Mathematics multiple choice test functioned differentially among urban students and rural students having the same ability in Mathematics. Six (6) items favoured rural schools While nine (9) items favoured urban schools. This implies that the performance of the examinees on the items do not only depend on their ability in Mathematics, but also on the school location. The finding agrees with Amuche & Fan (2014) reported that the test items of 2002 NECO Mathematics items had elements of bias that disadvantaged the rural school examinees and favours the urban schools exam mees.

Another aspect of the findings revealed that all the fifty (50) items of the 2018 WASSCE (SC) Mathematics multiple choice test were biased with respect to school ownership. 44 items were biased against public schools, while 6 items were biased against private schools. It means that the performance of the examinees on the items do not only depend on their ability in Mathematics, but also on the ownership of their schools. This finding is in congruence with Chukwudi (2019) who affirmed that 38 out of the 60 items of the 2017 BECE Mathematics multiple choice test were biased with respect to school ownership. 31 items were biased against public schools, while 7 items were biased against private schools.

Conclusion

The findings revealed that the performance of testees in the 2018 WASSCE (SC) Mathematics do not only depend on their ability, but also on gender, the schools' locations and their ownership. The study therefore exposed the need for teachers and examiners to be trained by examination bodies and the ministries of education on item writing. This would enable them to produce valid, reliable and bias-free

assessment tools which will ultimately improve the quality of assessment.

Recommendations

- Examination bodies and item writers should adopt the IRT approach to test development and refinement.
- They should ensure that they look at item bias and not item analysis alone to ensure quality items.
- They should also commit themselves to eliminating or reducing gender, location and school type biased items in public examinations.

REFERENCES

Adedoyin, O. O. (2010). Using item response theory approach to detect gender biased items in public examinations: Educational Research and Reviews Academic Journals, 5(7), 385-399.

Agnes, E. Ms, Anthony, N. N. & Julie, E.T. (2009). Analysis of students' performance in junior secondary school Mathematics examination in Bayelsa state in Nigeria. Study Home Comm Sci, 3(2), 131-134.

Alade, O. M.,Aletan, S. & B. S. (2018). Locus of Control and Self-Efficacy as Predictors of Academic Achievement in Senior Secondary School Mathematics in Lagos State, Nigeria. African Journal of Theory and Practice of Educational Research (AJTPER), 5(2), 71-86

Alade, O. M., Kuku, O. O. & OsobaA. (2017). Factors Influencing Candidates' performance In English Language and Mathematics in West African Senior School Certificate Examination in Lagos State, Nigeria. International Journal of Evaluation and Research in Education (IJERE), 3(2), 35-46

Aliyu, R. T. (2015). Development and Validation of Mathematics achievement test using the Rasch Model. International Journal of Education, 2(2), 107-113.

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, Arts and Social Sciences. Retrieved from http://www.iasir.ne

Aremu, O. A., & Sokan, B. O. (2003). A multicausal evaluation of academic performance of Nigerian learners: issues and implications for national development. Journal of Applied Psychology, 34(3), 334-342

Bachman, L. F. (2004). Statistical analysis for language assessment. Cambridge: Cambridge University Press.

Bolaji, C. (2005). A study of factors influencing students' attitude towards mathematics in Certificate Examination in Lagos State, Nigeria. Journal Institute of Advanced Engineering and Science

Chukwudi, D. U., (2019). Differential Item Functioning and Item Analysis of Basic Education Certificate Examination Mathematics Multiple Choice Test Items in Enugu, Nigeria. (Unpublished P. hd thesis), Department of Educational Foundations University of Lagos.

Federal Republic of Nigeria (2013). National Policy on Education. Lagos: NERDC Press.

Hamzah, M. D. & Hanan, A. A. (2008). Sex related differential item functioning (DIF) analysis of TIMSS. Dirasat Educational Sciences, 35, supplement.

Ilogu, G. C. (2005). Educational research and evaluation: A comparison for students. Yaba: Mandate communication Ltd.

Lee, J. & Mclritire, W. (2001). Interstate variation in the mathematics achievement of rural and non- students. Journal of Research in Rural Education, *16*(3),15-21.

Madu, B. C. (2012). Analysis of gender-related differential item functioning in Mathematics multiple choice items. Administered by West African Examination Council (WAEC). Journal of Education and Practice, 3(8), 22-25.

National Mathematics Centre, Abuja (2009). Mathematics improvement Programme. www nmcabuja.org/mathematics_improvement_programme.html. Retrieved 13thAugust, 2019.

Oguoma, C. C., Metibemu, M. A. & Okoye, R. O. (2016). An assessment of the dimensionality of 2014 West African Secondary Certificate Examination Mathematics Objective Test Scores in Imo State, Nigeria. African Journal of Theory and Practice of Educational Assessment, 4(5), 18-33.

Ojerinde, D. (2013). Classical test theory (CTT) versus item-response theory (IRT): An evaluation of comparability of item analysis result. A paper presented at the Institute of Education University of Ibadan 23rd May, 2013 lecture series.

Olonode, P.O. (2016). Equating 2014 senior school certificate mathematics examinations of West African Examinations Council and National Examinations Council in Lagos State. Nigeria. (Unpublished Ph.D. thesis), International center for Educational Evaluation University of Ibadan.

Pedraijita, J. Q. (2009). Using logistic regression to detect biased test items Performance in English Language and Mathematics at West African Senior School. The International of Educational and Psychological Assessment Journal, 33,2.

Pido, S. (2012). Comparison of item analysis results obtained using item response theory and classical test theory approaches. Journal of Educational Assessment in Africa, 7, 192-207.

Rover, C. (2005). 'That's not fair!' Fairness, bias and differential item functioning in language testing. Retrieved from http://www2.hawaii.edu/roever/brownbag.

Schnipke, D. L., Roussos, L. A. & Pashley, P. J. (2000). A comparison of MantelHaenszel differential item functioning parameters. Retrieved from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1a/70/b4 pdf.

Tremblay, S., Ross, N. A. , & Berthelot, J.M. (2001). Individual and community factors affecting grade three academic performance: a multi-level analysis. Education Quarterly Review, 7(4), 2536.

Umoinyang, I. E. (2011). The challenging of removing consistent errors in achievement test using differential item functioning (DIF) detecting methods. Paper presented at the 30th conference of the Association of Educational Assessment in Africa, Nairobi, Kenya.

Uruemu, O. & Adams O. (2013). Differential item functioning method as an item bias indicator. Educational Research, 4(4), 367-373.