JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Determination of Differential Functioning, Difficulty and Discriminating Indices and Students' Performance in 2023 BECE Multiple-Choice Test

Patience Odu, Lilian Ekwonna
Department of Guidance and Counselling
(Measurement and Evaluation Unit), Faculty of Education, Delta State University, Abraka

ABSTRACT

The study investigated the rate of items response (IR), differential functioning (DF), of students' performance, difficulty and discriminative indices of 2023, Mathematics multiple choice test used in Basic Education Certificate Examination (BECE) in Delta State. Four research questions guided the study which were answered using percentage, difficulty and discriminative indices formula. The instrument was 60 multiple-choice test items adopted from the 2023 BECE in Delta State. The population of the study comprised 140,959 upper basic three students from public secondary schools. The sample size of 1000 upper basic three students were drawn from the population using multi-stage sampling technique. The multi-triangulation research design was utilized. The instrument was administered to the students by the researcher and the researcher assistants. The study discovered the following findings: (1) The incidence of IR, DF issues in 2023 BECE test items was low with only about 5% of items indicating significant concern (2) The items have difficulty indices ranging from 0.31 to 0.55 and the estimated discriminative indices ranged from 0.34 to 0.67 and (3) There were similarity in performance of male and female students. Based on these findings the following recommendations were made among others: teachers should identify the items response, differential functioning in order to establish the psychometric properties of test items designed by them.

ARTICLE INFO
Article History
Received: March, 2026
Received in revised form: June, 2026
Accepted: August, 2026
Published online: September, 2026

KEYWORDS
Item Response (IR), Differential functioning (DF), Item Discriminative Index

INTRODUCTION

Differential Item Functioning is a condition in which examinees from different group shows differential probabilities of endorsing an item after matching on the underlying ability that the item is intended to measure (Issayeva, 2024). DIF exists if an item is relatively difficulty, discriminating or easily guessed for one group than for another (Guo, 2020 & Ereikan, 2020). They further explained that an item displays DIF if the test takers who have an equal amount of the ability to be tested but are members of different population sub-group (e.g. different gender, location or ethnicity) do not have an equal probability of answering item correctly. In essence DIF occurs when people from different groups with the same ability have systematically different responses to test items.

DIF of an item can therefore be understood as a lack of conditional independence between an item responses and group membership (e.g. gender, or ethnicity) giving the same latent ability or trait (Hamza & Husseini, 2023). It has shown that over the years the students' performance in Mathematics has been poor. Conventionally, the performance of students in other school subjects has always been higher than that of Mathematics this have been a topical issue in Nigeria. Odu (2025) observed that from 1991 to 2016, 27.31% of students in Nigeria had credit and above (A1 – C6) while 572.69% has pass and below (D7 – F9) in the May/June

Corresponding author: Lilian Ekwonna
ekwonnalilian@gmail.com
Department of Guidance and Counselling (Measurement and Evaluation Unit), Faculty of Education, Delta State University, Abraka.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

WASSCE in general Mathematics. Johnson et al. (2023) noted that 65.71% pass rate for 2020 and declined to 54.11% pass rate in 2021.

Several research efforts have been made to solve the problems of poor performance in Mathematics. Many of such efforts followed the observation of Rose (2023). Some of these efforts are in the improvement of methods of teaching quality of instructional techniques employed by the teachers and method of presentation before the students in order to improve Mathematics results. Nwachukwu (2023) investigated psycho-social factors like gender, location, attitude, anxiety, confidence and motivation among others in Mathematics achievement. However all these reports did not recognize DIF as a factor of under achievement of students in Junior Secondary School Mathematics. Therefore, the problem of this study is to determine the influence of Items Differential Functioning on teachers' construction of multiple-choice test in Junior Secondary School.

Specifically, this study was designed to answer the following research questions:

1. What is the percentage of IR in DF in 2023 Junior WASSCE Mathematic multiple choice test items among student in Delta State?
2. What is the difficulty index of 2023 Junior WASSCE Mathematics multiple choice test items?
3. What is the discriminative index of the 2023 Junior WASSCE Mathematic multiple choice test items?
4. What is the performance of male and female test takers in WASSCE Mathematics multiple choice test items in 2023?

LITERATURE REVIEW

The literature review for this study is categorized and organized in the following sub-sections:

1. Overview of 2023 Junior WASSCE Multiple Choice Test Items among Students in Delta State
2. Item Difficulty Index and 2023 Junior WASSCE Multiple Choice Test
3. Item Discriminative Index and 2023 Junior WASSCE Multiple Choice Test
4. Index of DIF among Male and Female Test Takers in Junior WASSCE Mathematic Multiple Choice Test Items.

Overview of 2023 Junior WASSCE Multiple Choice Test Items among Students in Delta State

According to the West African Examination Council (WAEC), out of a total of 1,613, 733 candidates who sat for the 2023 examination, the results of 262,803 candidates are being withheld due to reports of examination malpractices. Additionally, 1,361,608 candidates representing 84.38 percent obtained credit and above in a minimum of five subjects, excluding English Language or Mathematics, from 1991 to 2016, 27.31% of students in Nigeria got credit pass and below (A1-C6) while 72.69% had pass and below (D7-F0) in the May/June WASSCE in general Mathematics (Lawal, 2019).

Nwachukwu (2023) noted a 65.71% pass rate for 2020 and declined to 54.11% pass rate in 2-21. With these worrisome statistics, experts are more inclined to blame the school, majorly the teachers and neglecting the other potential factors like the quality of test instrument used in assessing Mathematics performance in Junior Secondary School. One major way of monitoring the quality and standards of the teaching and learning of Mathematics in schools is through the assessment of the learning outcome of the students which is through a good quality test (National Council of Teachers of Mathematics, 2020).

The types of test used by most teachers is teachers-made test which is constructed by classroom teachers and most often may lack good psychometric properties (Jessa, et al., 2023) Also, many examining bodies like the Delta State Ministry of Education are responsible for the development of Mathematics achievement test at the Junior Secondary Schools level, such tests are used by the examining bodies to test students' performance for the sake of certification.

There are some major challenges confronting these examination bodies. One of

Corresponding author: Lilian Ekwonna
ekwonnalilian@gmail.com
Department of Guidance and Counselling (Measurement and Evaluation Unit), Faculty of Education, Delta State University, Abraka.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

such is the development of test items which will cover the stipulated syllabus and how to objectively access the performance of the candidates. Another challenge with the current (BECE) Mathematics achievement test use by examination bodies is that most of the tests are developed and validated using Classical Test Theory (CTT).

As a result of this, they are faced with some challenges like, poor precision, sample dependency and undue focus on aggregate scores that deny test developer the opportunity of determining how the examiner performed on a test item. This problem can be addressed or overcome with the application of four parameter item response theory (IRT). Therefore, DIF may be an instrument to students' poor performance in Junior Secondary School multiple choice test Mathematics in Delta State (Odu, 2025).

Item Difficulty Index and 2023 Mathematics Multiple Choice Test Used in Basic Education Certificate Examination (BECE)

Item difficulty index (p) measures how easy or hard a question (ranging from 0 to 1, where higher is easier). An ideal test must be of moderate difficulty level. Various methods have been developed by different authors for estimating the difficulty index or facility index. This method includes:

1. P = $\frac{R}{T \text{ or } N}$ where

P = difficulty level, R = number of testee who got the item correctly.
T= total number of testee. The importance of this formula is that it allows the test constructor to estimate the mean value of the total test. However, two weaknesses have been identifies with the methods.

First 'P' is actually a measure of item easiness. The larger the item proportion, the easier the item. Secondly, 'P' is not linearly related to seals of difficulty with equal intervals and as such it cannot be average. To overcome this, 'P' can be transformed to Z-scores

2. The difficulty index 'P' can also be computed using the formula

$$P = \frac{(1-\frac{W}{K-1})}{R+W}$$

Where P = difficulty Index
R = Number of testee answering the item correctly
W = Number of testee answering the item wrongly
K= Number of alternative responses.

This method was used in the computation of difficult index in this test construction (Adopted from Guildford, 1984 in Irighweferhe, 2014).

Item Discriminative Index

Item discriminative index measures how well a specific item differentiates between high performing students and low performing students. Cheng and Liu (20150 states that discriminative index shows the difference between examinees with high and low achievement. There exist different methods of establishing the discriminative index. Some os the methods are stated below:

1. The Constructing Criterion

In this method discriminative index is obtained by subtracting the number of the testee in the lower group who got the item right from the number of the testee in the upper group who got the item right and divide by half the total number of the testee who responded to the items. It is calculated using the formula:

$$D = \frac{Ru-RL}{\frac{1}{2}T}$$

Where Ru = number of testee in the upper group who got the items right,
RL = number of testee in the low group who got the items right and T = total number of the students who responded to the items.

2. Biserial Correlation Coefficient

The biserial correlation coefficient is an estimate of the Pearson Product Moment Correlation. The use of this method to determine item discriminative index assumes that the relationship between criterion score and the dichotomized item continuum can be represented by a normal bivariate distribution. The student should be normally distributed on item continuum which dichotomized into right and wrong responses.

Corresponding author: Lilian Ekwonna
ekwonnalilian@gmail.com
Department of Guidance and Counselling (Measurement and Evaluation Unit), Faculty of Education, Delta State University, Abraka.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

DIF Index among Male and Female Test Takers in BECE

Most of the studies have reported overall difference between males and females, usually in favour of males especially when cognitive complexity increases and the content changes from Arithematic to Algebra and Geometry (Traub, 2020). On the other hand Thompson (2019) have reported that males were better on word problems and algebra items favoured females. Smith and Brown (2024) have reported that measurement items and items involving solving favoured males. They have also considered the age levels and concluded that gender related DIF becomes larger as age increases. In addition to the gender related studies, there are also studies in the literature concerning race and ethnicity related DIF as well. For example, Johnson and Taylor (2024) have reported that whites had an advantage over African Americans in items with figures applied and realistic problems and word problems. On the other hand, Whites were disadvantaged in Algebra items with a curriculum-like items and abstract items.

METHODOLOGY

The study adopted a multiple triangulation research design. As outlined by Kpolovie and Emekene (2016), this design supports a multiple–method approach to examining the psychometric properties of an instrument. The sample size of 1000 Upper Basic three students from public secondary schools in Delta State were drawn from a population of 140, 959 Upper basic three students using multi-stage sampling technique. The sample size was chosen based on Lord's as cited in Loken & Rolison 2023 recommendation, which states that at least 1000 are needed to accurately estimate a d-parameter.

The primary instrument for the study is the mathematic Achievement Test which has been validated by the Examinations and Standards Office in Delta State. The face and content validity were used to determine the validity of the instrument. Three experts, two in measurement and Evaluation discipline in Delta State University, Abraka and one head of Mathematics teacher in Saint Patrick Asaba, Delta State were used to validate the instrument. The experts' reports revealed that the items covered the basic education curriculum as well as the levels of cognitive domains, thus the instrument was found valid for the study.

The reliability of the instrument was carried out on thirty (30) Junior Secondary School Students in Edo State through a trial test (pilot test) of the Mathematics Achievement Test. 30 copies of the 2023 JSS 3 question papers containing 60 multiple test items each were administered to students in Edo State. The Kuder Richardson -21 was used to compute the reliability index which yield a coefficient of 0.87 using SPSS package. The instrument covered only cognitive domain since it is a multiple-choice test. The instrument was administered to JSS 3 Students (600 male and 400 female) through the help of three research assistants. The four research questions were answered using Jimetrik software, a psychometric tool that supports advanced item response theory (IRT) modelling.

RESULTS

Table 1: Percentage of IRDF in 2023 Junior WASSCE Mathematics Multiple choice test items in Delta State

Item	Code	A-Parameter (SE)	B-Parameter (SE)	C-Parameter (SE)	D-Parameter (SE)
MAT1	L4	1.42 (0.57)	3.27 (0.61)	0.12 (0.01)	0.92 (0.05)
MAT 27	L4	2.34 (0.27)	2.23 (0.12)	0.01 (0.00)	0.92 (0.05)
MAT 35	L4	2.08 (0.41)	2.27 (0.17)	0.12 (0.01)	0.92 (0.05)
MAT 3	L4	1.29 (0.54)	5.95 (0.47)	0.11 (0.01)	0.91 (0.05)
MAT 4	L4	0.04 (0.04)	-2.08 (9.77)	0.24 (0.11)	0.91 (0.05)
MAT 5	L4	1.03 (0.60)	5.71 (2.47)	0.29 (0.02)	0.91 (0.05)
MAT 6	L4	0.18 (0.12)	5.85 (1.39)	0.25 (0.12)	0.91 (0.05)
MAT 7	L4	0.82 (0.81)	19.18 (0.62)	0.29 (0.01)	0.91 (0.05)

Corresponding author: Lilian Ekwonna
ekwonnalilian@gmail.com
Department of Guidance and Counselling (Measurement and Evaluation Unit), Faculty of Education, Delta State University, Abraka.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Item	Code	A-Parameter (SE)	B-Parameter (SE)	C-Parameter (SE)	D-Parameter (SE)
MAT 8	L4	0.29 (0.17)	5.27 (2.15)	0.22 (0.09)	0.91 (0.05)
MAT 9	L4	1.30 (0.55)	5.96 (0.44)	0.15 (0.01)	0.91 (0.05)
MAT 10	L4	0.34 (0.12)	1.02 (1.11)	0.23 (0.11)	0.91 (0.05)
MAT 11	L4	0.92 (0.66)	5.02 (2.16)	0.09 (0.02)	0.91 (0.05)
MAT 12	L4	1.10 (0.54)	5.93 (0.63)	0.07 (0.01)	0.91 (0.05)
MAT 13	L4	0.46 (0.18)	3.60 (0.97)	0.16 (0.06)	0.91 (0.05)
MAT 14	L4	1.06 (0.66)	3.85 (1.16)	0.14 (0.02)	0.91 (0.05)
MAT 15	L4	0.47 (0.26)	5.52 (1.89)	0.15 (0.05)	0.91 (0.05)
MAT 16	L4	1.24 (0.58)	5.95 (0.52)	0.30 (0.01)	0.91 (0.05)
MAT 17	L4	1.20 (0.42)	3.75 (0.74)	0.03 (0.01)	0.91 (0.05)
MAT 18	L4	0.98 (0.58)	5.75 (1.98)	0.09 (0.01)	0.91 (0.05)
MAT 19	L4	1.18 (0.59)	5.94 (0.59)	0.19 (0.01)	0.91 (0.05)
MAT 20	L4	1.21 (0.63)	3.24 (0.85)	0.32 (0.02)	0.91 (0.05)
MAT 21	L4	0.72 (0.17)	3.61 (0.63)	0.04 (0.02)	0.91 (0.05)
MAT 22	L4	0.01 (0.01)	-101.00 (131.06)	0.23 (0.11)	0.91 (0.05)
MAT 23	L4	1.86 (0.22)	1.22 (0.11)	0.03 (0.01)	0.91 (0.05)
MAT 24	L4	0.82 (0.81)	19.78 (309.79)	0.30 (0.01)	0.91 (0.05)
MAT 25	L4	1.41 (0.52)	5.96 (0.35)	0.11 (0.01)	0.91 (0.05)
MAT 26	L4	0.82 (0.81)	20.50 (428.06)	0.35 (0.01)	0.91 (0.05)
MAT 30	L4	1.60 (0.29)	2.64 (0.24)	0.02 (0.01)	0.91 (0.05)
MAT 32	L4	0.78 (0.23)	4.19 (0.87)	0.04 (0.01)	0.91 (0.05)
MAT 34	L4	1.17 (0.59)	4.68 (1.58)	0.04 (0.01)	0.91 (0.05)
MAT 36	L4	0.83 (0.81)	3.79 (0.53)	0.11 (0.01)	0.91 (0.05)
MAT 37	L4	0.78 (0.22)	4.19 (0.88)	0.04 (0.01)	0.91 (0.05)
MAT 38	L4	0.47 (0.24)	4.88 (1.72)	0.01 (0.01)	0.91 (0.05)
MAT 39	L4	0.61 (0.17)	3.92 (0.77)	0.03 (0.01)	0.91 (0.05)
MAT 40	L4	1.40 (0.30)	2.91 (0.21)	0.04 (0.01)	0.91 (0.05)
MAT 41	L4	0.98 (0.41)	5.74 (1.26)	0.04 (0.01)	0.91 (0.05)
MAT 42	L4	1.14 (0.56)	5.93 (0.64)	0.05 (0.01)	0.91 (0.05)
MAT 43	L4	0.82 (0.81)	14.59 (9.01)	0.29 (0.01)	0.91 (0.05)
MAT 44	L4	1.83 (0.44)	2.19 (0.22)	0.02 (0.01)	0.91 (0.05)
MAT 45	L4	1.36 (0.58)	5.95 (0.47)	0.07 (0.01)	0.91 (0.05)
MAT 46	L4	0.90 (0.33)	5.51 (1.15)	0.04 (0.01)	0.91 (0.05)
MAT 47	L4	1.14 (0.45)	4.36 (0.61)	0.02 (0.01)	0.91 (0.05)
MAT 48	L4	1.78 (0.25)	1.65 (0.09)	0.04 (0.01)	0.91 (0.05)
MAT 49	L4	0.64 (0.16)	4.18 (0.87)	0.03 (0.01)	0.91 (0.05)
MAT 50	L4	1.35 (0.50)	5.95 (0.51)	0.07 (0.01)	0.91 (0.05)
MAT 51	L4	1.83 (0.47)	1.56 (0.18)	0.01 (0.01)	0.91 (0.05)
MAT 52	L4	0.93 (0.24)	5.36 (0.61)	0.01 (0.01)	0.91 (0.05)
MAT 53	L4	1.16 (0.30)	4.04 (0.45)	0.03 (0.01)	0.91 (0.05)
MAT 54	L4	1.17 (0.36)	4.11 (0.44)	0.03 (0.01)	0.91 (0.05)
MAT 55	L4	1.71 (0.38)	2.42 (0.22)	0.04 (0.01)	0.91 (0.05)
MAT 56	L4	1.02 (0.24)	5.38 (0.59)	0.01 (0.01)	0.91 (0.05)
MAT 57	L4	1.61 (0.41)	2.53 (0.31)	0.03 (0.01)	0.91 (0.05)
MAT 58	L4	1.83 (0.31)	1.93 (0.15)	0.02 (0.01)	0.91 (0.05)
MAT 59	L4	0.91 (0.27)	5.45 (0.73)	0.02 (0.01)	0.91 (0.05)
MAT 60	L4	0.86 (0.26)	5.25 (0.84)	0.02 (0.01)	0.91 (0.05)

Corresponding author: Lilian Ekwonna
ekwonnalilian@gmail.com
Department of Guidance and Counselling (Measurement and Evaluation Unit), Faculty of Education, Delta State University, Abraka.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Item	Code	A-Parameter (SE)	B-Parameter (SE)	C-Parameter (SE)	D-Parameter (SE)
MAT 31	L4	2.00 (0.34)	1.70 (0.14)	0.10 (0.01)	0.90 (0.05)
MAT 28	L4	0.01 (0.01)	5.71 (2.91)	0.16 (0.06)	0.89 (0.05)
MAT 33	L4	0.66 (0.34)	-2.21 (0.62)	1.00 (0.13)	0.76 (0.05)
MAT 2	L4	0.82 (0.81)	-15.16 (14.45)	0.97 (0.62)	0.76 (0.01)
MAT 29	L4	1.99 (0.45)	1.96 (0.24)	0.35 (0.02)	0.70 (0.62)

Criterion: <0.80

Table 1 shows the percentage of IR DF in 2023 Junior WASSCE mathematic multiple test items in Delta State. The result showed that the incidence of IR DF issues in the 2023 BECE mathematics multiple choice test items in Delta State is low with only about 5% of items (3 out of 60) indicating significant concern. The majority of the items exhibit high IR DF values about 0.80, suggesting that most items function well for high ability students.

Table 2: Difficulty Index of Items of the 2023 BECE Mathematics in Delta State

Items	Difficulty Level
1	0.51
2	0.52
3	0.41
4	0.35
5	0.44
6	0.34
7	0.53
8	0.45
9	0.40
10	0.41
11	0.41
12	0.47
13	0.45
14	0.32
15	0.39
16	0.53
17	0.44
18	0.48
19	0.45
20	0.48
21	0.44
22	0.43
23	0.41
24	0.51
25	0.47
26	0.41
27	0.44
28	0.40
29	0.39
30	0.47
31	0.43
32	0.43
33	0.43
34	0.36
35	0.44
36	0.31
37	0.55
38	0.41
39	0.45
40	0.41
41	0.36
42	0.47
43	0.48
44	0.47
45	0.43
46	0.33
47	0.36
48	0.33
49	0.45
50	0.45
51	0.44
52	0.49
53	0.47
54	0.50
55	0.42
56	0.33
57	0.36
58	0.33
59	0.35
60	0.35

The item analysis in Table 2 revealed that 43 items out of the 60 meet the requirement for difficulty index. Therefore, the remaining 17 items should be reviewed.

Corresponding author: Lilian Ekwonna
ekwonnalilian@gmail.com
Department of Guidance and Counselling (Measurement and Evaluation Unit), Faculty of Education, Delta State University, Abraka.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Table 3: Discriminative Index of Items in 2023 BECE Mathematics in Delta State

Items	Discriminative Index
1	0.52
2	0.55
3	0.68
4	0.55
5	0.35
6	0.69
7	0.55
8	0.37
9	0.49
10	0.49
11	0.61
12	0.66
13	0.63
14	0.67
15	0.44
16	0.60
17	0.61
18	0.55
19	0.59
20	0.49
21	0.67
22	0.48
23	0.52
24	0.50
25	0.53
26	0.46
27	0.44
28	0.50
29	0.54
30	0.45
31	0.41
32	0.40
33	0.39
34	0.55
35	0.41
36	0.68
37	0.56
38	0.46
39	0.35
40	0.52
41	0.57
42	0.34
43	0.48
44	0.51
45	0.40
46	0.63
47	0.68
48	0.51
49	0.51
50	0.51
51	0.45
52	0.44
53	0.45
54	0.44
55	0.43
56	0.61
57	0.60
58	0.50
59	0.51
60	0.51

At the end of the items analysis in Table 3, the result indicated that 55 items out of the 60 items met the requirement for discriminative index. Therefore, the remaining five items should be reviewed.

Table 4: Performance of Male and Female Students (Test takers) in BECE Mathematics Multiple Choice test items in 2023

S/N	Scores	Frequency	Male %	Frequency	Female %
1	61-70	29	2.9	24	2.4
2	51-60	121	12.1	85	8.5
3	41-50	190	19	107	10.7
4	31-40	119	11.9	89	8.9
5	21-30	87	8.7	58	5.8
6	11-20	54	5.4	37	3.7
	Total	600	60	400	40

Corresponding author: Lilian Ekwonna
ekwonnalilian@gmail.com
Department of Guidance and Counselling (Measurement and Evaluation Unit), Faculty of Education, Delta State University, Abraka.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Table 4 indicates that 29 male basic Education students representing 2.9 percent of the total sample have scores ranging between 61-70 and 24 female students representing 2.4 percent have the score ranging between 61-70. 121 male (12.1 percent) and 85 female (8.5 percent) scored between 51-60 mark. 190 male students (19 percent) and 107 female (10.7 percent) scored between 41-50 marks. 89 female (8.9 percent) and 119 male (11.9 percent) have scores between 31-40 marks. 87 male students (8.7 percent) scored between 21-30. 54 male student (5.4 percent) and 37 female students (3.7 percent) scored between 11-20 marks. From the above analysis, it can be observed that there are similarity in performance of male and female students in performance of male and female students in the BECE mathematics multiple choice test.

DISCUSSION OF FINDINGS

Incidence of IR, DF in 2023 BECE Mathematics Test Items

The result in Table 1 shows that the incidence of IR. DF issues in the 2023 BECE Mathematics multiple –choice test items in Delta State is low with only about 5% of items indicating significant concern. The finding agrees with Anisal et al. (2022) who found that BECE Mathematics test items are performing well among top performers and that guessing or item ambiguity is minimal. The finding further agreed with Bello and Omisore (2023) who found that the effective performance of most items could reflect the use of curriculum-based design frameworks and alignment with learners' proficiency levels, which has been shown to reduce item bias and increase test validity.

Estimated Difficulty Index of BECE Multiple-choice Mathematics Test Items

The results in Table 2 showed that the items have difficulty indices ranging from 0.31 to 0.55. These indices fall within the recommended ranges as Johnson (2023). Lawal (2019) suggested that difficulty level lower than – 2.5 and greater than 2.5 are irrelevant for test. A critical look into the indices showed that 42 items out of the 60 items representing 70% have indices ranging between 0.41 and 0.55 while 18 items out of the 60 items representing 30% have difficulty index ranging from 0.31 to 0.40 which are of very difficult level. From the above analysis, the BECE Mathematics multiple –choice test item covers a wide range spectrum of ability of students.

Estimated Discriminative Indices of BECE Multiple Test Items

The result obtained in Table 3 shows that the estimated discriminative indices range from 0.34 to 0.67. The analysis of discriminative indices shows that 36 items (60%) of the items have a high discriminative index of 0.50 and above. This is in an agreement with Miller and Green (2024) who agreed that items with discriminative index of 0.50 and above discriminate well between the items and have effective discriminative power

Performance of Male and Female Students (Test Takers) in WASSCE Multiple Choice Test Items

The male and female students' performance in the BECE Mathematics tests revealed that there is similarity in their performance. This study aligns with Omoruari (2017), Opara (2016) and Oreheruata (2015). Their findings revealed that individual's abilities and determination to work are responsible for variation in performances of students not necessarily gender. However, Pill (2021) and Robitzac (2022) stated that there are gender differences in performance of students in Mathematics inclusive.

CONCLUSION

The item response (IR) differential functioning in the 2023 BECE Mathematic multiple –choice test items in Delta State is low with only about 5% indicating significant concern. Similarly, the items analysis met the requirements for difficulty and discriminative indices. More so, the male and female students' performance in the test revealed that there is similarity in their performance though with some little variations.

Corresponding author: Lilian Ekwonna
ekwonnalilian@gmail.com
Department of Guidance and Counselling (Measurement and Evaluation Unit), Faculty of Education, Delta State University, Abraka.

RECOMMENDATIONS

The following recommendation were made based on the findings of the study:

1. Teachers should identify the items response differential functioning in order to establish the psychometric properties of test items designed by them
2. Instrument for testing items for examination should follow the standard of difficulty and discriminative indices.
3. Testing instruments should possess or meet the required standard in items of psychometric properties for the instrument to be used in the school system.

REFERENCES

Anisah, N., Suparno, S., & Anwar, Y. (2022). Evaluation of test items using the four-parameter logistic model: A case study of multiple-choice questions. *International Journal of Instruction, 15*(3), 231–246. https://doi.org/10.29333/iji.2022.15313a

Bello, T. A., & Omisore, M. A. (2023). Curriculum alignment and test item performance in standardized assessments: A Nigerian context. *African Journal of Educational Research and Development, 15*(1), 51–66. https://doi.org/10.4314/

Cheng, Y., & Liu, C. (2015). The Effect of Upper and Lower Asymptotes of IRT Models on Computerized Adaptive Testing. *Applied Psychological Measurement,* 39(7), 551–565. https://doi.org/10.1177/0146621615585850

Guo, S., & Ercikan, K. (2020). Test engagement and rapid guessing: Evidence from a large-scale state assessment. *Frontiers in Education,* 8, 1127644. https://doi.org/10.3389/feduc.2023.1127644

Hamza, S. A., & Hussaini, A. (2023). Items Analysis of Mathematics Basic Education Certificate Examination in Junior Secondary Schools before and After Covid-19 Lockdown in Sokoto State. African Journal of Humanities and Contemporary Education Research, 11(1), 243–257. Retrieved from https://publications.afropolitanjournals.com/index.php/ajhcer/article/view/474

Issayeva, L. (2024). *Factor Analysis: Evaluating Dimensionality in Assessment.* SAGE Publishing.

Jessa, M. O., Odili, J. N., & Osadebe, P. U. (2023). Development of Social Studies Aptitude Test for Testing Critical Thinking Skills: Implication for the Achievement of Education for Sustainable Development (ESD). *Canadian Journal of Educational and Social Studies,* 3(4), 99-119. DOI: 10.53103/cjess.v3i4.163

Johnson, L., & Taylor, A. (2024). Convergent Evidence of Construct Validity: New Perspectives. *Educational Research Review,* 19(2), 112-126. https://doi.org/10.1016/j.edurev.2024.100212

Johnson, L., Smith, R., & Lee, R. (2023). Exploring Latent Space in Educational Assessments. *Educational and Psychological Measurement,* 82(2), 75-89. https://doi.org/10.1177/00131644221111456

Kose, S. (2023). Evaluating model-data fit in IRT: A stepwise approach. *Educational Research Review,* 18(2), 123-135. https://doi.org/10.1016/j.edurev.2023.100123

Kpolovie, P.J. and Emekene, C.O. (2016). Item Response Theory Validation of Advanced Progressive Matrices in Nigeria. British Journal of Psychology Research 4(1) 1-32

Lawal, A. (2019). Assessing Senior Secondary School Students' Mathematical Proficiency as Related to Gender and Performance in Mathematics in Nigeria. *International Journal of Research in*

Corresponding author: Lilian Ekwonna
ekwonnalilian@gmail.com
Department of Guidance and Counselling (Measurement and Evaluation Unit), Faculty of Education, Delta State University, Abraka.

Education and Science (IJRES), 3(2), 488-502.

Miller, J., & Green, T. (2024). Understanding the Standard Error of Measurement in Psychological Assessments. *Psychological Assessment,* 39(1), 55-70. https://doi.org/10.1037/pas0000486

Miller, J., & Thompson, R. (2024). Understanding Face Validity: Implications for Test Design. *Psychological Assessment,* 36(4), 330-345. https://doi.org/10.1037/pas0000462

National Council of Teachers of Mathematics. (2020). Catalyzing change in high school mathematics: Initiating the conversation on a pathway to equity and excellence. https://www.nctm.org/Standards-and-Positions/Principles-and-Standards/

Nwachukwu, T. (2023). Secondary School Students Self-Regulated Learning Skill as Predictor of Mathematics Achievement in Imo State Nigeria: Focus on Gender. *Brillo Journal*, 2(2), 94-107.

Odu, P. (2025). Evaluation of the implementation of Junior Secondary School Mathematics Curriculum for Future Resilient Society Renewal in Delta State. *Irish Journal of Education Practice*; 8, Issue 04, 43-53.

Omoruan, B. E. (2017). *Construction, validation and application of mathematics language mastery achievement test for junior secondary school students in Delta and Edo States, Nigeria.* Unpublished Doctoral Thesis, University of Port Harcourt, Port Harcourt.

Opara, I.M. (2016). *Test construction and measurement: Concept and applications.* Owerri: Career Publishers.

Orheruata, M. U. (2015). Item Parameter Drift in Certificate Examinations and it's implication on Decision Making. *African Journal of Theory and Practice of Educational assessment. Educational Assessment* 2, 98-105.

Pill, J. (2021). Content Validation and Its Role in Ensuring Test Accuracy. *Educational Research Review,* 16(1), 45-62. https://doi.org/10.1016/j.edurev.2021.100109

Robitzsch, A. (2022). Four-Parameter Guessing Model and Related Item Response Models. *Mathematical and Computational Applications,* 27(6), 95. https://doi.org/10.3390/mca27060095

Smith, R., & Brown, T. (2024). The Role of SEM in Test Reliability. *Educational and Psychological Measurement,* 88(1), 90-105. https://doi.org/10.1177/00131644231123349

Thompson, N. J. (2019). *Item response theory: Modelling psychometric data* (2nd ed.). Routledge.

Traub, R. E. (2020). *Psychometric theory and applications.* Springer.

Corresponding author: Lilian Ekwonna
ekwonnalilian@gmail.com
Department of Guidance and Counselling (Measurement and Evaluation Unit), Faculty of Education, Delta State University, Abraka.