International Journal of Advanced Education
www.allpaperjournal.com
Online ISSN: 3107-9601
Received: 04-12-2025, Accepted: 15-01-2026, Published: 02-02-2026
Volume 2, Issue 1, 2026, Page No. 78-85

Comparison of item parameters of WAEC, NECO and NABTEB mathematics multiple-choice questions obtained using IRT in Anambra state

Nnedi Somto Kanyo-Egwuonwu, Romy O Okoye
Department of Educational Foundations, Nnamdi Azikiwe University, Awka, Nigeria

Abstract
WAEC, NECO and NABTEB are examination bodies responsible for the award of senior school certificate in Nigeria. The purpose of this study is to compare the means of item parameters of 2022-2023 Mathematics multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB. The study was guided by six research questions and six hypotheses. It employed a descriptive survey design. The population is made up of 57,150 SS 3 students of public secondary schools which includes 22, 421 that registered for WAEC, 31, 613 that registered for NECO and 3,116 that registered for NABTEB. The sample size was 1600 respondents obtained through Multi stage sampling procedure. The instruments for this study consisted of WAEC, NECO and NABTEB 2022 and 2023 Mathematics multiple-choice examination questions. The instruments for the study were administered to the respondents by research assistants who are the Mathematics teachers of the sampled schools. The collected data were subjected to item analysis using MPLUS Version 7.4 software statistical analysis in IRT framework. To answer the research questions, the means of the parameters for each examination body were obtained. To test the hypotheses, the means of the parameters for the three examination bodies were compared using One-Way ANOVA at 0.05 level of significance. The researcher found out that NABTEB sets the most difficult questions, followed by WAEC, while NECO took the least position in 2022 Mathematics. In 2023, NABTEB sets the most difficult, followed by WAEC, while NECO took the least position in Mathematics. NABTEB sets the most discriminating questions, followed by WAEC, while NECO took the least position in 2022 Mathematics. On the contrary, WAEC sets the most discriminating questions, followed by NABTEB, while NECO took the least position in 2023 Mathematics. In the year 2023, NABTEB sets the most guessable questions, followed by NECO, while WAEC took the least position in 2022 Mathematics, while NECO sets the most guessable questions, followed by NABTEB, while WAEC took the least position in 2023 Mathematics. The study also revealed that there was a significant different in difficulty, discrimination and guessing indices of questions in Mathematics within the period under review. Based on the findings, it was recommended among others that WAEC, NECO and NABTEB should improve more on the item analysis of their test items.

Keywords: Item Parameters, Mathematics, Multiple Choice Questions, Item Response Theory, Evaluation

Introduction

Contact sports such as rugby, American football, and ice hockey are among the most physically demanding athletic activities, combining strength, speed, and frequent player-to-player collisions. While they promote fitness and social engagement, they also pose a high risk of acute and chronic injuries. Understanding injury patterns, mechanisms, and prevention strategies is essential for safe participation and effective rehabilitation (Koh *et al*., 2003) [9].

India has a rich tradition of indigenous contact sports that combine physical endurance, strategy, and close physical engagement. Among these, Kabaddi, Kho-Kho, and Wrestling (Kushti) stand out as culturally significant and National examinations bodies in any country organize public examinations which provide uniform standards to all test takers, irrespective of the type or method of instruction they have received. Some of these examination bodies in Nigeria are: the West African Examinations Council (WAEC), National Examinations Council (NECO), Joint Admissions and Matriculation Board (JAMB), National Business and Technical Examination Board (NABTEB) among others. A close look at the functions of these boards reveal that some of them perform similar functions. For instance, WAEC, NECO and NABTEB all conduct senior secondary school leaver's certificate examinations but in the case of NABTEB, the examination is targeted only at secondary school leavers of technical and vocational colleges in Nigeria. Afred (2013), Aborisade and Fajobi (2020) [1] states that, just three examining bodies: WAEC, NECO and NABTEB are responsible for the award of senior school certificate in Nigeria. They conduct parallel or equivalent senior school certificates examinations in the country. They maintain high standards in the development and administration of the examinations, of which performances in the examinations should be good indicators of individuals' standing in any subject area of interest (Bassey & Bassey, 2024; Metibemu, 2016) [9].

Of the five required credit passes in school subjects, Mathematics is known to be the most fundamental and useful tool in the technological advancement of any nation (James & Olufunke, 2020 [6]; Oghenerume, 2025).Furthermore, it has been described as the most useful instrument in commerce, physical sciences, engineering, social sciences, industry, medicine and biological sciences (Bandele & Adewale, 2013; Oghenerume & Egberha, 2024) [4, 8]. Considering the fact that WAEC, NECO and NABTEB have been given similar mandate to conduct the senior school certificate examination; it is expected that these examination bodies should be comparable in various respects. Osoba, Alade and Aletan (2024) [16] claimed that NECO is inferior to WAEC in all standard while Oyiborhoro Odili and Osadebe (2023) [17] concluded that given X grade in NECO Mathematics examination is equal

to X + 1 grade in WAEC examination, meaning that WAEC is more difficult than NECO.

Item response theory is a modeling technique that tries to describe the relationship between an examinee's test performance and the latent trait underlying the performance (Kanzow, Schmidt & Kanzow, 2023) [7]. The item response theory (IRT) is a paradigm for the design, analysis, and scoring of tests, questionnaires and similar instruments measuring abilities, attitudes or other variables. It is known as latent trait theory, strong true score theory, or modern mental test theory. IRT is applied in diagnostic test to select the most appropriate items for examinees based on individual ability. Item Response Theory looks at the examinee's performance by using item distributions based on the examinee's probability of success on latent variable. Under IRT, parameters of persons are invariant across items, and parameters of the items are invariant in different populations of persons (Awopeju & Afolabi, 2023, Nwankwo, 2023) [3, 10]. It places item difficulty and student performance on the same scale, tells how much information each item or item score level contributes to the test, provides standard error of measurement for each item and statistic of fit of each item to the model. This brings greater flexibility and provides more sophisticated information which allows for the improvement of the reliability of an assessment.

The advantages of IRT model can be achieved only if there is a satisfactory goodness-of-fit between the model and test data (Nwankwo & Oguike, 2023, Oghenerume, 2022 [10, 12]). First, IRT, nonlinear relationship is used. Such a nonlinear function is more general and subsumes a linear relationship. The nonlinear function that relates the probability of observing a certain response to an individual item with the latent trait is called the item response function (IRF) or item characteristic curve (ICC). This function enables flexible specifications of the theoretical relationship between the underlying trait and items, given response format (dichotomous or polytomous), context, or theoretical assumptions about response process (e.g., dimensionality). Additionally, from empirical point of view, the higher flexibility of IRT models allows for a close fit to be achieved between a function and the data (Aborisade & Fajobi, 2020 [1], Ogunbamowo, Adediwura & Diyan, 2024). The most popular function employed is the logistic function, e.g., in the one parameter logistic model or two parameter logistic models. The former assumes constant discrimination of items and unidimensional measurement; the latter additionally allows for different discrimination or even a pick-answer-at-random parameter. Multidimensional generalizations exist for most of these models. IRT model allows assessment of the adequateness by means of statistical goodness-of-fit tests and fit indices.

A multiple-choice question presents a question or statement followed by several response options out of which, and only one is correct. It's a type of objective assessment used in various settings, from quizzes to standardized tests. Multiple-choice test format permits the diagnosis of students' difficulties by analyzing incorrect responses while varying the questions' level of difficulty. Multiple-choice items are also amenable to item analysis, which permits a determination of which items are ambiguous or too difficult (Ogunbamowo, Adediwura & Diyan, 2019). Moreover, multiple-choice format can be easily corrected and it is reliable. Bassey and Bassey (2024) contended that multiple-choice test format, tests only recognition of knowledge, facilitates cheating, severely restricts what can be tested and promotes guessing which may have a considerable but unknown effect on test scores and academic achievement. Again, it is very difficult to write successful items.

A number of empirical enquiries in Nigeria have been advanced to assess the extent of comparability of WAEC, NECO and NABTEB examinations. For example, Mawak, Efomo and Mustapha (2024) [8] compared the psychometric properties of WAEC and NECO Mathematics multiple choice questions and found that the reliability and validity coefficients of the two tests were similar. Empirical studies like Essien *et al* (2024), Osoba, Alade and Aletan (2024) [16]. Oyiborhorho, Odili and Osadebe (2023) [17] on comparative analysis of the item difficulty indices of WAEC, NECO and NABTEB Mathematics questions concluded that NECO produces the most difficult test items followed by NABTEB, while WAEC took least position.

Based on available empirical studies, it appears that the focus of previous research was on establishing the equivalence and comparability of WAEC and NECO Nigeria. The present study focused on comparison of item parameters of WAEC, NECO and NABTEB. From the available empirical studies reviewed, some studies only covered one academic session and with small sample sizes. A research study with a larger sample size and covering the three examination bodies WAEC, NECO and NABTEB a longer period of time is needed, to ensure a more valid conclusion. All these and more prompted the present researcher's interest to further research on the comparison of item parameters of WAEC, NECO and NABTEB Mathematics multiple-choice questions obtained using IRT.

The following research questions guided the study

1. What are the means of item difficulty indices of 2022 Mathematics MCQs set by WAEC, NECO and NABTEB?
2. What are the means of item difficulty indices of 2023 [13] Mathematics MCQs set by WAEC, NECO and NABTEB?
3. What are the means of item discrimination indices of 2022 Mathematics MCQs set by WAEC, NECO and NABTEB?
4. What are the means of item discrimination indices of 2023 [13] Mathematics MCQs set by WAEC, NECO and NABTEB?
5. What are the means of guessing indices of 2022 Mathematics MCQs set by WAEC, NECO and NABTEB?
6. What are the means of guessing indices of 2023 Mathematics MCQs set by WAEC, NECO and NABTEB?

The following hypotheses were tested at.05 level

1. There is no significant difference in the mean difficulty indices of 2022 Mathematics MCQs set by WAEC, NECO and NABTEB.
2. There is no significant difference in the mean difficulty indices of 2023 Mathematics MCQs set by WAEC, NECO and NABTEB.
3. There is no significant difference in the mean discrimination indices of 2022 Mathematics MCQs set by WAEC, NECO and NABTEB.
4. There is no significant difference in the mean discrimination indices of 2023 Mathematics MCQs set by WAEC, NECO and NABTEB.

5. There is no significant difference in the mean guessing indices of 2022 Mathematics MCQs set by WAEC, NECO and NABTEB.
6. There is no significant difference in the mean guessing indices of 2023 Mathematics MCQs set by WAEC, NECO and NABTEB.

Methods

The population was made up of 57,150 SS 3 students of public secondary schools which includes 57,150 that registered for WAEC, 44,853 that registered for NECO and 2,250 that registered for NABTEB (Source: Post-Primary Schools Services Commission, Anambra State, 2025). Similarly, all students that registered for NECO, registered for WAEC and all the students that registered for NABTEB, registered for WAEC 2024 [8]/2025 Mathematics Senior School Certificate Examination in Anambra State.

The sample size is 1600 respondents. Multi stage sampling procedure was used to compose the sample for the study. In the first stage, simple random sampling technique was used to obtain three (3) Education zones out of the six (6) Education zones in Anambra State. In the second stage, two (2) local government areas will be obtained through simple random sampling, from each of the three Education zones, making a total of six (6) local government areas. In the last stage, two schools were obtained through simple random sampling from each of the Local Government Areas, making a total of 12 schools. All the SS 3 students in the 12 schools were used for the study.

The instruments for this study consists of WAEC, NECO and NABTEB 2022 and 2023 Mathematics multiple choice examination questions. The WAEC May/June Mathematics multiple choice test items for each of the years contained 50 items to be completed in one and half hours. Similarly, NECO Mathematics multiple choice examination questions for each of the years contained 50 items to be completed in one and half hours. In contrast, NABTEB June/July Mathematics multiple choice test items for each of the years contained 60 items to be completed in one hour and forty-five minutes. The instruments were collected from the afore-mentioned examination bodies' in Awka zonal offices.

The instruments for the study were administered to the respondents by research assistants who are the Mathematics teachers of the sampled schools. The researcher got a letter of introduction to the principals of sampled secondary schools for the study. Through the help of the principals, the researcher got across to the Mathematics teachers of SS 3 students to ascertain the most convenient time-schedule to administer the Mathematics tests. The administration were done under the supervision of the researcher.

The responses collected for each year. The responses were scored manually. The responses were scored 1 for the right answer and zero (0) for wrong answer. The scored responses were subjected to item analysis using MPLUS Version 7.4 software statistical analysis in IRT framework. The analysis yielded difficulty, discrimination and guessing indices for each examination body, for each of the years under review. For each parameter, the indices for the three examination bodies were compared. To answer the research questions, the means of the parameters for each examination body were obtained. To test the hypotheses, the means of the parameters for the three examination bodies were compared using One-Way ANOVA at 0.05 level of significance. In taking decisions regarding the hypotheses, a null hypothesis was rejected if the probability value (p-value) is less than or equal to significant value of 0.05; if otherwise ($p>0.05$), the null hypothesis was not rejected.

Results and Discussion

This section focuses on the presentation and discussion of findings in line with research questions and hypotheses that guided the study.

Research Question 1. What are the means of item difficulty indices of 2022 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB?

To answer this research question, the mean of the difficulty parameters for each examination body in the year 2022 was obtained. The results are presented in Table 1.

Table 1: Means of Item Difficulty Indices of 2022 and 2023Mathematics Multiple-Choice Questions (MCQs) set by WAEC, NECO and NABTEB

Year 2022	N	Mean	Std. Deviation	Minimum	Maximum
NABTEB	50	1.430	1.181	.059	4.685
WAEC	50	.603	.996	.008	5.458
NECO	60	-.910	3.182	-11.054	6.558
Year 2023					
NABTEB	50	1.496	1.797	.173	9.080
WAEC	50	.540	.817	.007	4.497
NECO	60	-.194	2.980	-11.387	4.639

Table 1 shows the mean difficulty indices of WAEC, NECO and NABTEB multiple-choice Mathematics items. The table reveals that in 2022, NABTEB Mathematics items had mean difficulty index of 1.430, WAEC items had mean difficulty index of.603, while NECO items had mean difficulty index of -.910. Hence in 2022, NABTEB produced the most difficult test items, followed by WAEC, while NECO had the least.

Research Question 2: What are the means of item difficulty indices of 2023 [13] Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB?

To answer the research question, the mean of the difficulty parameters for each examination body in the year 2023 was obtained. The results are presented in Table 1.

Table 1 shows the mean difficulty indices of WAEC, NECO and NABTEB multiple-choice Mathematics items. The table reveals that in 2023 [13] NABTEB Mathematics items had mean difficulty index of 1.496, Those in WAEC had mean difficulty index of.540, while those in NECO had mean difficulty index of -.194. Hence, in 2023 NABTEB produced the most difficult test items, followed by WAEC, while NECO had the least.

Looking at the result in Table 1, one would conclude that NABTEB had the most difficult items, followed by WAEC and then NECO. It should be noted that this does implies that NABTEB had the best items or that NECO had the worst items. The best one could say is that NABTEB had the most difficult items while NECO had the least difficult. The finding of the present study agrees with those of Alfred (2013), Metibemu (2016), James and Olufunke (2020) [2, 6, 9], who found that NABTEB Mathematics multiple choice items were more difficult than WAEC and NECO items. It however, disagreed with the findings of Bandele and Adewale (2013) [4] the NECO examinations were more difficult than those of WAEC and NABTEB.

Research Question 3: What are the means of item discrimination indices of 2022 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB?
To answer the research question, the mean of the discrimination parameters for each examination body in the year 2022 was obtained. The results are presented in Table 2.

Table 2: Means of Item Discrimination Indices of 2022 and 2023 [13] Item Mathematics multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB

Year 2022	N	Mean	Std. Deviation	Minimum	Maximum
NABTEB	50	1.820	2.271	.004	9.521
WAEC	50	1.177	.504	.063	2.475
NECO	60	.643	.706	-.270	1.818
Year 2023					
NABTEB	50	.809	.440	.009	1.763
WAEC	50	1.163	.446	.178	2.038
NECO	60	.759	.594	-.167	2.054

Table 2 shows the mean discrimination indices of WAEC, NECO and NABTEB Mathematics multiple-choice questions (MCQs). The result reveals that in 2022, NABTEB Mathematics items had mean discrimination index of 1.820, WAEC items had mean discrimination indices of 1.177, while NECO multiple-choice Mathematics items had mean discrimination index of.643. Hence, NABTEB produced the most discriminating test items, followed by WAEC, while NECO had the least index.

Research Question 4: What are the means of item discrimination indices of 2023 [13] Mathematics multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB?
To answer Research Question 4, the mean of the discrimination parameter for each examination body in the year 2023 was obtained. The results are presented in Table 2.
Table 2 shows the mean discrimination indices of WAEC, NECO and NABTEB multiple-choice Mathematics items. It reveals that in 2023 [13], NABTEB Mathematics items had mean discrimination index of.80920, WAEC Mathematics items had mean discrimination index of 1.16338, while NECO multiple-choice Mathematics items had mean discrimination index of.75948. Hence, WAEC produced the most discriminating test items, followed by NABTEB, while NECO had the least.
In summary, Tables 2 show that in 2024, NABTEB had the most discriminating items followed by WAEC and then NECO, while in 2023WAEC had the most discriminating items followed by NABTEB, and then NECO. There was no consistency regarding which examination body had the most discriminating items. A consistent finding however was that NECO had the least discriminating items among the three examination bodies. The results partly agree with the findings of Ogunbamowo, Adediwura and Diyan (2019) which showed that the discrimination indices of NABTEB examination questions were higher than those of NECO and WAEC in Economics. However, the study disagreed with those of Metibemu (2016), James and Olufunke (2020) [6, 9], who reported that NECO Mathematics multiple choice items were most discrimination, followed WAEC and NABTEB.

Research Question 5. What are the means of guessing indices of 2022 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB?
To answer the research question, the mean of the guessing parameter for each examination body in the year 2022 was obtained.

Table 3: Means of Guessing Indices of 2022 and 2023 Mathematics Multiple-Choice Questions (MCQs) set by WAEC, NECO and NABTEB

Year 2022	N	Mean	Std. Deviation	Minimum	Maximum
NABTEB	50	.245	.164	.065	.634
WAEC	50	.170	.105	.049	.463
NECO	60	.192	.083	.047	.461
Year 2023					
NABTEB	50	.195	.064	.089	.402
WAEC	50	.175	.105	.041	.472
NECO	60	.205	.072	.109	.503

Table 3 shows the mean guessing indices of WAEC, NECO and NABTEB multiple-choice Mathematics items. The table reveals that in 2022, NABTEB Mathematics items had mean guessing index of.245, WAEC Mathematics items had mean guessing index of.170, while NECO had mean guessing index of.192. Hence, NABTEB produced the most guessable items, followed by NECO, while WAEC produced the least. It need to be recalled that an item with smaller guessing index is preferable to one with larger index. In terms of guessing parameter thereafter, WAEC had the best items, followed by NECO, while NABTEB came last.

Research Question 6: What are the means of guessing indices of 2023 Mathematics multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB?
To answer the research question, the mean of the guessing parameter for each examination body in the year 2023 was obtained. The results are presented in Table 3.
Table 3 shows the mean guessing indices of WAEC, NECO and NABTEB multiple-choice Mathematics items. The table reveals that NABTEB Mathematics items had mean guessing index of.195, WAEC Mathematics items had mean guessing index of.175, while NECO multiple-choice Mathematics items had mean guessing index of.205. Hence, NECO produces the most guessable items, followed by NABTEB, while WAEC had the least. In term of guessing parameter therefore, WAEC had the best items followed by NABTEB, and then NECO.
In summary, Table 3 showed that in 2022, WAEC had the best items in terms of guessing index, followed by those of NECO and then NABTEB, while in 2023 WAEC also had the best items followed by NABTEB, and then NECO. This finding agreed with the findings of James and Olufunke (2020) [6] who revealed that WAEC produced the least guessing index of the examination followed by NECO and NABTEB. However, the study disagreed with that of Ogunbamowo *et al*. (2019) that NECO produced the least guessing index of the examination followed by WAEC and NABTEB.

H01: There is no significant difference in the mean difficulty indices of 2022 Mathematics multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB.

To test the hypothesis, the means of the difficulty indices for the three examination bodies in 2022 were compared using One-Way ANOVA test at 0.05 level of significance. The result is presented in Table 4.

Table 4: ANOVA Comparison of Difficulty Indices of 2022 Mathematics Multiple-Choice Questions (MCQs) set by WAEC, NECO and NABTEB.

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	156.384	2	78.192	17.183	.000
Within Groups	714.451	157	4.551		
Total	870.835	159			

Table 4 shows that there is significant difference in the mean difficulty indices of 2022 questions set by WAEC, NECO and NABTEB ($F(2,157) = 17.183$; $p < .05$). A post hoc test with a Bonferroni correction was used after the ANOVA (Analysis of Variance) test to perform multiple pairwise comparisons while controlling for the increased risk of Type I errors. The results are presented in Table 5

Table 5: Multiple Comparisons of Difficulty Indices of 2022 Mathematics multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB

(I) EXAM	(J) EXAM	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
NABTEB	WAEC	.823	.427	.163	-.205	1.859
	NECO	2.341 *	.408	.000	1.352	3.329
WAEC	NABTEB	-.827	.427	.163	-1.859	.205
	NECO	1.514 *	.408	.001	.525	2.502
NECO	NABTEB	-2.341 *	.408	.000	-3.329	-1.352
	WAEC	-1.514 *	.408	.001	-2.502	-.525

The post hoc comparison in Table reveals that there is a significant difference between the difficulty indices of 2022 questions set by NABTEB and NECO (Mean difference=2.341, SE=.408, $p<.05$) in favour of NABTEB. Also, there is a significant difference between the difficulty indices of 2022 questions set by WAEC and NECO (Mean difference =1.514, SE=.408, $p<.05$) in favour of WAEC. However, there is no significant difference between the difficulty indices of 2022 Mathematics Multiple-choice questions (MCQs) set by WAEC and NABTEB (Mean difference = -.827, SE=.427, p =.165).

H02: There is no significant difference in the mean difficulty indices of 2023 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB.

To test the hypothesis, the means of the difficulty indices for the three examination bodies in 2023 were compared using One-Way ANOVA test at 0.05 level of significance. The result is presented in Table 6.

Table 6: *ANOVA Comparison of Difficulty Indices of 2023 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB.*

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	73.072	2	36.536	8.454	.000
Within Groups	643.902	157	4.321		
Total	716.974	159			

Table 6 shows that there is significant difference in the mean difficulty indices of 2023 Mathematics multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB ($F(2,157) = 8.454$; $p < .05$). A post hoc test with a Bonferroni correction was used after the ANOVA (Analysis of Variance) test to perform multiple pairwise comparisons while controlling for the increased risk of Type I errors. The results are presented in Table 6.

Table 6: *Multiple Comparisons Difficulty Indices of 2023 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB*

(I) EXAM	(J) EXAM	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
NABTEB	WAEC	.956	.416	.069	-.051	1.962
	NECO	1.689 *	.416	.000	.693	2.687
WAEC	NABTEB	-.956	.416	.069	-1.963	.051
	NECO	.734	.412	.231	-.264	1.730
NECO	NABTEB	-1.689 *	.412	.000	-2.687	-.693
	WAEC	-.734	.412	.231	-1.730	.263

A post hoc comparison in Table 6 shows that there is a significant difference between the difficulty indices of questions set by NABTEB and NECO (Mean difference =1.69, SE=.412, $p<.05$) in favour of NABTEB. However, there is no significant difference between the difficulty Indices of 2023 questions set by WAEC and NECO (Mean difference =.734, SE=.412, p =.231). Also, there is no significant difference between the difficulty Indices of 2023 Mathematics Multiple-choice questions (MCQs) set by NABTEB and WAEC (Mean difference =.956, SE=.416, p =.069).

In summary, therefore, as regards 2022 questions there was significant difference between the difficulty indices of NABTEB and NECO, in favour of NABTEB and also a significant difference between the mean difficulty indices of questions set by WAEC and NECO in favour of WAEC. There was however no difference between the difficulty indices of questions set by WAEC and NABTEB. As regards the 2023 questions, significant difference existed between the difficulty indices of questions set by NABTEB and NECO, in favour of NABTEB. There was therefore no difference between the difficulty indices of questions set by NABTEB and WAEC on one hand, and WAEC and NECO on the other hand.

The above findings disagree with those of Alfred (2013), Metibemu (2016), James and Olufunke (2020) [2, 6, 9] who revealed that there was no significant difference between the difficulty indices of WAEC, NABTEB and NECO multiple choice items in Mathematics.

H03: There is no significant difference in the mean discrimination indices of 2022 Mathematics multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB.

To test the hypothesis, the means of the discrimination indices for the three examination bodies in 2022 were compared using One-Way ANOVA test at 0.05 level of significance. The result is presented in Table 7.

Table 7: ANOVA Comparison of Discrimination Indices of 2022 Mathematics Multiple-Choice Questions (MCQs) set by WAEC, NECO and NABTEB.

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	37.818	2	18.909	10.075	.000
Within Groups	294.669	157	1.877		
Total	332.487	159			

Table 7 shows that there is significant difference in the mean discrimination indices of 2022 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB ($F(2,157) = 10.075$; $p < .05$). A post hoc test with a Bonferroni correction was used after the ANOVA (Analysis of Variance) test to perform multiple pairwise comparisons while controlling for the increased risk of Type I errors. The results are presented in Table 8.

Table 8: Multiple Comparisons of Discrimination Indices of 2022 Mathematics Multiple-Choice Questions (MCQs) set by WAEC, NECO and NABTEB.

(I) EXAM	(J) EXAM	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
NABTEB	WAEC	.643	.274	.060	-.019	1.306
	NECO	1.178 *	.262	.000	.542	1.812
WAEC	NABTEB	-.643	.274	.060	-1.306	.019
	NECO	.534	.262	.130	-.101	1.168
NECO	NABTEB	-1.178 *	.262	.000	-1.812	-.542
	WAEC	-.534	.262	.130	-1.169	.101

A post hoc comparison in Table 8 shows that there is a significant difference between the discrimination indices of 2022 Mathematics Multiple-choice questions (MCQs) set by NABTEB and NECO (Mean difference =1.178, SE=.262, $p<.05$) in favour of NABTEB. However, there is no significant difference between the discrimination indices of 2022 questions set by NABTEB and WAEC (Mean difference =.643, SE=.274, p=.060). Also, there is no significant difference between the discrimination indices of 2022 questions set by NECO and WAEC (Mean difference = -.534, SE=.262, p =.130).

H04: There is no significant difference in the mean discrimination indices of 2023 Mathematics multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB.

To test the hypothesis, the mean of the discrimination indices for the three examination bodies in 2023 were compared using One-Way ANOVA test at 0.05 level of significance. The result is presented in Table 9.

Table 9: ANOVA Comparison Discrimination Indices of 2023 Mathematics Multiple-Choice Questions (MCQs) set by WAEC, NECO and NABTEB.

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	5.065	2	2.533	9.917	.000
Within Groups	40.094	157	.255		
Total	45.159	159			

Table 9 shows that there is significant difference in the mean discrimination indices of 2023 Mathematics multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB ($F(2,159) = 9.917$; $p < .05$). A post hoc test with a Bonferroni correction was used after the ANOVA (Analysis of Variance) test to perform multiple pairwise comparisons while controlling for the increased risk of Type I errors. The results are presented in Table 10.

Table 10: Multiple Comparisons Discrimination Indices of 2023 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB.

(I) EXAM	(J) EXAM	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
NABTEB	WAEC	-.354 *	.101	.002	-.599	-.109
	NECO	.049	.097	1.000	-.185	.284
WAEC	NABTEB	.354 *	.101	.002	.109	.599
	NECO	.404 *	.097	.000	.169	.638
NECO	NABTEB	-.049	.097	1.000	-.284	.185
	WAEC	-.404 *	.097	.000	-.638	-.169

A post hoc comparison in Table 10 reveals that there is a significant difference between the discrimination indices of 2023 Mathematics multiple-choice questions (MCQs) set by WAEC and NABTEB (Mean difference =.354, SE=.101, $p<.05$) in favour of WAEC. Also, there is a significant difference between the discrimination indices of 2023 Mathematics MCQs set by WAEC and NECO (Mean difference =.404, SE=.097, $p<.05$) in favour of WAEC. However, there is no significant difference between the discrimination indices of 2023 Mathematics multiple-choice questions (MCQs) set by NABTEB and NECO (Mean difference =.049, SE=.097, p =1.00).

In summary therefore, regarding the discrimination indices of 2022 questions, significant difference only occurred between the questions set by NABTEB and NECO in favour of NABTEB. This implies that the questions set by NABTEB exhibited superiority over those set by NECO. Between NABTEB and WAEC on one hand, there was no exhibition of superiority.

As regards 2023 questions there were significant differences discrimination indices between NABTEB and WAEC, in favour of WAEC and between WAEC and NECO, in favour of WAEC. WAEC therefore showed superiority over each of the other two examination bodies.

H05: There is no significant difference in the mean guessing indices of 2022 Mathematics multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB.

To test hypothesis 5, the means of the guessing indices for the three examination bodies in 2022 were compared using One-Way ANOVA test at 0.05 level of significance. The result is presented in Table 11.

Table 11: ANOVA of Guessing Indices of 2022 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB.

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	.150	2	.075	5.198	.007
Within Groups	2.263	157	.014		
Total	2.413	159			

Table 11 shows that there is significant difference in the mean guessing indices of 2022 Mathematics multiple-choice

questions (MCQs) set by WAEC, NECO and NABTEB ($F(2,159) = 5.198$; $p < .05$). A post hoc test with a Bonferroni correction was used after the ANOVA (Analysis of Variance) test to perform multiple pairwise comparisons while controlling for the increased risk of Type I errors. The results are presented in Table 12.

Table 12: Multiple Comparisons Guessing Indices of 2022 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB.

(I) EXAM	(J) EXAM	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
NABTEB	WAEC	.075 *	.024	.006	.016	.133
	NECO	.053	.022	.066	-.002	.108
WAEC	NABTEB	-.075 *	.024	.006	-.133	-.016
	NECO	-.021	.022	1.000	-.077	.033
NECO	NABTEB	-.053	.022	.066	-.108	.002
	WAEC	.021	.022	1.000	-.033	.077

A post hoc comparison shows in Table 12 shows that in 2022 there was a significant difference only between the guessing indices of questions set by NABTEB and WAEC (Mean Difference= 0.075, SE=0.024, $p<.05$), with NABTEB having higher index, implying that WAEC had better items. There is no significant difference the indices of NABTEB and NECO (Mean Difference= 0.053, SE=0.023, $p>.05$) as well as that between the indices for WAEC and NECO (Mean Difference=0.022, SE=0.023, $p>.05$).

H06: There is no significant difference in the mean guessing indices of 2023 Mathematics multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB.

To test the hypothesis, the means of the guessing indices for the three examination bodies in 2022 were compared using One-Way ANOVA test at 0.05 level of significance. The result is presented in Table 14.

Table 14: ANOVA Comparison of Guessing Indices of 2023 Mathematics Multiple-Choice Questions (MCQs) set by WAEC, NECO and NABTEB.

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	.025	2	.012	1.837	.163
Within Groups	1.051	157	.007		
Total	1.076	159			

Table 14 shows that there is no significant difference in the mean guessing indices of 2022 Mathematics multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB ($F(2,159) = 1.837$; $p < .05$).

In summary therefore, as regards discrimination index, except for 2023 where WAEC showed superiority over the other bodies, that superiority did not exist in 2022. We therefore cannot conclude that any of the exam bodies consistently maintained superiority in terms of discrimination index.

As earlier mentioned, as regards guessing index, apart from WAEC that showed superiority over NABTEB in 2022, there was no examination body that showed consistent superiority. In ability to establish any consistency in the results of the present study could be attributed to the few number of years covered by the study.

Conclusion

In conclusion therefore the present study has not established any consistency in the superiority of any of the examination bodies over the bodies in terms of their item parameters.

Recommendations

Based on the findings of the study, the following recommendations are made:

(i) Certificates issued by WAEC, NABTEB and NECO could be used freely for same purposes without any discrimination since the items of the examinations are of comparable standard.

(ii) NABTEB and NECO should improve their quality of item options so that students will not be able to guess their test items correctly.

(iii) WAEC items parameters are of high quality as revealed in the study. Based on this the examination body should continue to maintain her high standard known for over the years.

(iv) NABTEB items were too difficult as revealed in the study. Based on this, the examination body should revisit most of their test items to reduce level of difficulty.

(v) Government should establish a joint examination evaluation team to ensure that examination items constructed by different examination bodies are of the same standard.

(vi) It is recommended that a similar study is carried out covering more numbers of years.

References

1. Aborisade OJ, Fajobi OO. Comparative analysis of psychometric properties of Mathematics items constructed by WAEC and NECO in Nigeria using item response theory approach. Educational Research and Reviews,2020:15(1):1-7.
2. Alfred AO. Scoring tests with dichotomous and polychromous items. Maureen Grass Dean of Graduate School University of Georgia,2013.
3. Awopeju AO, Afolabi ER. Comparative analysis of classical test theory and item response theory based item parameter estimates of senior school certificate mathematics examination. European Scientific Journal,2023:12(2):263-284.
4. Bandele SO, Adewale AE. Comparative analysis of the reliability and validity coefficient of WAEC, NECO and NABTEB constructed mathematics examination. Journal of Educational and Social Research,2013:3(2):761-764.
5. Bassey AO, Bassey TO. Measurement models and the psychometric properties of 2016 and 2017 NECO Mathematics Instrument in Calabar Education Zone of Cross River State. Journal of Advances in Education and Philosophy,2024:6(6):311-319.
6. James AO, Olufunke FO. Comparative Analysis of Psychometric Properties of Mathematics items constructed by WAEC and NECO in Nigeria using Item Response Theory approach. Educational Research and Reviews,2020:15(1):1-7.
7. Kanzow AF, Schmidt D, Kanzow P. Scoring single-response multiple-choice items: Scoping review and comparison of different scoring methods. JMIR Medical Education,2023:9:e47092. doi:10.2196/47092.
8. Mawak KL, Efoma KC, Mustapha OO. The item difficulty and discrimination indices of senior school

certificate examination (SSCE) multiple choice Chemistry tests of West African Examinations Council (WAEC) and National Examinations Council (NECO) using classical test theory in Benue State. International Journal of Asian Social Sciences,2024:3(4):130-135.
9. Metibemu MA. Comparison of classical test theory and item response theory in the development and scoring of senior secondary school physics tests in Ondo State. Unpublished Ph.D Thesis. Institute of Education. University of Ibadan, 2016.
10. Nwankwo AB, Oguike CE. Examining the psychometric properties of WAEC and NECO mathematics items for Nigerian secondary schools using a 3-parameter logistic model. Unizik Journal of Educational Research and Policy Studies,2023:1(1):1-15.
11. Oghenerume RA, Egberha F. Comparative analysis of item statistics of WASSCE and NECO SSCE 2023 Data Processing multiple choice tests using item response theory. International Journal of Educational Researchers,2024:15(1):58-67.
12. Oghenerume RA. Comparative analysis of item statistics of WASSCE and NECO SSCE 2022 Data Processing multiple choice tests using item response theory. Unpublished Master's Thesis. University of Benin,2022.
13. Oghenerume RL. Item Statistics Disparity between 2023 WASSCE and NECO SSCE Mathematics Large-Scale Assessments. International Journal of Educational Researchers,2025:16(1):01-14.
14. Ogunbamowo AD, Adediwura OP, Diyan GO. Psychometric Properties of 2017 West African Examination Council and National Examinations Council's Economic Senior School Certificate Examination Items. Journal of Advances in Education and Philosophy,2024:3(1):9-18.
15. Ogunbamowo AO, Adediwura AA, Diyan RO. Psychometric properties of 2017 West African Examinations Council and National Examinations Council's Economic Senior School Certificate Examination items. Journal of Advances in Education and Philosophy,2019:3(1):9-18.
16. Osoba AV, Alade JN, Aletan PU. Application of item response theory in the validation of basic science test of Delta State Basic Education Certificate Examination. International Journal of Research in Education and Sustainable Development,2024:3(7):1-13.
17. Oyiborhoro PE, Odili JN, Osadebe PU. Assessment of concurrent validity of Delta State BECE Basic Science multiple choice items using item response theory. African Journal of Educational Assessment,2023:19(2):55-67.