International Journal of Education, Management, and Technology

e-ISSN : 3026-5444
p-ISSN : 3026-4200

Index: Harvard, Boston, Sydney University, Copernicus, Lens, Scilit, Semantic, Dimensions, etc

https://doi.org/10.58578/IJEMT.v4i2.9347

Comparison of Item Parameters of WAEC, NECO and NABTEB Mathematics Multiple Choice Questions Obtained Using IRT in Anambra State

Hassan Muhammad
Federal University of Education Kontagora Niger State, Nigeria
profmaths781@fcekg.edu.ng

Article Info:

Submitted:	Revised:	Accepted:	Published:
Mar 28, 2026	Jul 26, 2026	Aug 7, 2026	Aug 12, 2026

Abstract

Nigeria's senior school certificate examinations are administered by the West African Examinations Council (WAEC), National Examinations Council (NECO), and National Business and Technical Examinations Board (NABTEB); however, the psychometric comparability of their Mathematics items requires systematic evaluation. This study compared the mean difficulty, discrimination, and guessing parameters of the Mathematics multiple-choice items administered by the three examination bodies in 2022 and 2023. A descriptive survey design was employed, guided by six research questions and six corresponding hypotheses. The population comprised 57,150 Senior Secondary 3 students from public secondary schools, with reported registrations of 57,150 for WAEC, 44,853 for NECO, and 2,250 for NABTEB. A sample of 1,600 students was selected through a multistage sampling procedure. The instruments consisted of the 2022 and 2023 Mathematics multiple-choice examination items developed by WAEC, NECO, and NABTEB, which were administered by trained research assistants who were Mathematics teachers in the sampled schools. Responses were subjected to item analysis within an item response theory framework using Mplus Version 7.4. Mean parameter estimates were calculated to answer the research questions, while one-way analysis of

IJEMT Journal is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License

variance was conducted at the .05 significance level to test the hypotheses. For both years, NABTEB items had the highest mean difficulty parameters, followed by WAEC and NECO. In 2022, NABTEB items had the highest mean discrimination parameter, followed by WAEC and NECO; in 2023, WAEC ranked highest, followed by NABTEB and NECO. For the guessing parameter, NABTEB ranked highest in 2022, followed by NECO and WAEC, whereas NECO ranked highest in 2023, followed by NABTEB and WAEC. Statistically significant differences were found among the examination bodies in the mean difficulty, discrimination, and guessing parameters during the period examined. These findings demonstrate variation in the psychometric characteristics of the three examinations and underscore the need for WAEC, NECO, and NABTEB to strengthen their item-analysis and test-development procedures.

Keywords: Item Parameters; Item Response Theory; Mathematics Multiple-Choice Questions; Senior School Certificate Examinations; Test Quality

INTRODUCTION

Evaluation is a major educational tool used in the identification of individual talents and skills for placement of students at appropriate learning programmes or vocations. This in effect is the idea behind the establishment of evaluation agencies which also act as examining bodies that maintain a common standard in development and administration of public examinations. Nwankwo and Oguike (2023) reported that these agencies were set up to promote education, coordinate educational programmes, control and monitor the quality of educational institutions.

These agencies 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. Aborisade and Fajobi (2020) 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 certificate 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.

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 (Oghenerume, 2025). Furthermore, it has been described as the most useful instrument in commerce, physical sciences, engineering, social sciences, industry, medicine and biological sciences (Oghenerume, and Egberha, 2024). 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) claimed that NECO is inferior to WAEC in all standard while Oyiborhoro and Osadebe (2023) 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, and Kanzow, 2023). 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. The IRT is useful in constructing short forms used to ensure that items provide adequate precision across the entire range of interest. 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 the persons are invariant across items, and parameters of the items are invariant in different populations of persons (Awopeju and Afolabi, 2023). 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, provide 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 and Oguike, 2023). First, In 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 (e.g., 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 and Fajobi, 2020). 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 later additionally allows for different discrimination or even a pick-answer-at-random parameter. Multidimensional generalizations exist for most of these models. IRT model allow 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 answer options, and only one is correct. It's a type of objective assessment used in various settings, from quizzes to standardized tests. Multiple choice 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. 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.

Quite 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) 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. Other empirical studies like Essien et al (2024) laid credence to the finding of Osoba, Alade and Aletan (2024). Oyiborhorho, Odili and Osadebe

(2023), did comparative analysis of the item, difficulty indices of WAEC, NECO and NABTEB Mathematics questions and 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 research on establishing the equivalence and comparability of WAEC and NECO at least in Nigeria have always been centered on only the psychometric properties of the tests of the examining bodies. For this reason, one may choose to do the comparability of item parameters obtained using IRT. From the available empirical studies above, some studies only covered one academic session and with small sample sizes. A research study with a larger sample size and covering 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 item Mathematics MCQs set by WAEC, NECO and NABTEB?
2. What are the means of item difficulty indices of 2023 item Mathematics MCQs set by WAEC, NECO and NABTEB?
3. What are the means of item discrimination indices of 2022 item Mathematics MCQs set by WAEC, NECO and NABTEB?
4. What are the means of item discrimination indices of 2023 item Mathematics MCQs set by WAEC, NECO and NABTEB?
5. What are the means of guessing indices of 2022 item Mathematics MCQs set by WAEC, NECO and NABTEB?
6. What are the means of guessing indices of 2023 item Mathematics MCQs set by WAEC, NECO and NABTEB?

The following hypotheses were tested at. 05 level

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

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

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

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

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

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

METHODS

This study was employ a descriptive survey design. The study was undertaken in Anambra State. The population is 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).

The sample size is 1600 respondents. Multi stage sampling procedure was used to compose the sample for the study. The instruments for this study consists of WAEC, NECO and NABTEB 2022 and 2023 Mathematics multiple choice examination questions.

WAEC, NECO and NABTEB

Mathematics multiple choice examination question papers are already validated by examination experts. WAEC, NECO and NABTEB examination question papers are deemed to be reliable having been developed by the public examination bodies. The instruments for the study were administered to the respondents by research assistants who are the Mathematics teachers of the sampled schools.

The responses collected for each year. The OMR was scored manually. The responses were scored 1 for the right answer and zero (0) for wrong answer. 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

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 research question 1, the means of the difficulty parameters for each examination body in the year 2022 were obtained. The results are presented in Table 1.

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

	N	Mean	Std. Deviation	Minimum	Maximum
NABTEB	50	1.43044	1.181741	.059	4.685
WAEC	50	.60378	.996091	.008	5.458
NECO	60	-.91023	3.182059	-11.054	6.558
Total	160	.29436	2.340288	-11.054	6.558

Table 1 shows the means 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 1.43044, WAEC items had mean difficulty of.60378, while NECO items had mean difficulty of- .91023. 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 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB?

To answer research question 2, the means of the difficulty parameters for each examination body in the year 2023 were obtained. The results are presented in Table 2.

Table 2: *Means of Item Difficulty Indices of 2023 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB*

	N	Mean	Std. Deviation	Minimum	Maximum
NABTEB	50	1.49580	1.797364	.173	9.080
WAEC	50	.53968	.816671	.007	4.497
NECO	60	-.19383	2.980085	-11.387	4.639
Total	160	.60326	2.179030	-11.387	9.080

Table 2 shows the mean difficulty indices of WAEC, NECO and NABTEB multiple-choice Mathematics items. The table reveals that in 2023 NABTEB Mathematics items had mean difficulty index of 1.49580, Those in WAEC had mean difficulty index of .53968, while those in NECO had mean difficulty index of- .19383. Hence, in 2023 NABTEB produced the most difficult test items, followed by WAEC, while NECO had the least.

In summary, the study showed that in 2022 and 2023, NABTEB had the most difficult items, followed by WAEC and then NECO. The finding agreed with findings of studies carried out by Alfred (2013); Metibemu (2016); James and Olufunke (2020) which says that NABTEB mathematics multiple-choice items were more difficult than WAEC and NECO. However, the study disagreed with Bandele and Adewale (2013) submission that NECO examinations are more difficult than 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 research question 3, the means of the discrimination parameters for each examination body in the year 2022 were obtained. The results are presented in Table 3.

Table 3: *Means of Item Discrimination Indices of 2022 Item Mathematics* Multiple-choice questions (MCQs) *set by WAEC, NECO and NABTEB*

	N	Mean	Std. Deviation	Minimum	Maximum
NABTEB	50	1.82020	2.271097	.004	9.521
WAEC	50	1.17678	.504528	.063	2.475
NECO	60	.64263	.706627	-.270	1.818
Total	160	1.17754	1.446069	-.270	9.521

Table 3 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 1.82020, WAEC items had mean discrimination indices of 1.17678, while NECO multiple-choice Mathematics items had mean discrimination index of.64263. 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 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB?

To answer research question 4, the means of the discrimination parameters for each examination body in the year 2023 were obtained. The results are presented in Table 4.

Table 4: *Means of Item Discrimination Indices of 2023 Mathematics* Multiple-choice questions (MCQs) *set by WAEC, NECO and NABTEB*

	N	Mean	Std. Deviation	Minimum	Maximum
NABTEB	50	.80920	.440415	.009	1.763
WAEC	50	1.16338	.446404	.178	2.038
NECO	60	.75948	.594111	-.167	2.054
Total	160	.90124	.532935	-.167	2.054

Table 4 shows the mean discrimination indices of WAEC, NECO and NABTEB multiple-choice Mathematics items. It reveals that in 2023, 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, the study showed that in 2022, NABTEB had most discriminating test items, followed by WAEC and then NECO, while in 2023 WAEC had the most discriminating items followed by those of NABTEB and then NECO. This result agreed with the findings of Ogunbamowo, Adediwura and Diyan (2019) which showed that the discrimination indices of NABTEB Examination were higher than those of NECO, WAEC Economics items. However, the study disagreed with that of Metibemu (2016); James and Olufunke (2020) who reported that NECO Mathematic multiple-choice items were most discriminating, followed by WAEC and then 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 research question 5, the means of the guessing parameters for each examination body in the year 2022 were obtained.

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

	N	Mean	Std. Deviation	Minimum	Maximum
NABTEB	50	.24510	.164046	.065	.634
WAEC	50	.17010	.104618	.049	.463
NECO	60	.19188	.083209	.047	.461
Total	160	.20171	.123200	.047	.634

Table 5 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.24510, WAEC Mathematics items had mean guessing index of.17010, while NECO had mean guessing index of.19188. Hence, NABTEB produced the most guessable items, followed by NECO, while WAEC produced the least.

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 research question 6, the means of the guessing parameters for each examination body in the year 2023 were obtained. The results are presented in Table 6.

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

	N	Mean	Std. Deviation	Minimum	Maximum
NABTEB	50	.19478	.064347	.089	.402
WAEC	50	.17486	.105105	.041	.472
NECO	60	.20460	.072124	.109	.503
Total	160	.19224	.082252	.041	.503

Table 6 shows the mean guessing indices of WAEC, NECO and NABTEB multiple-choice Mathematics items. The table revealed that NABTEB Mathematics items had mean guessing index of. 19478, WAEC Mathematics items had mean guessing index of. 17486, while NECO multiple-choice Mathematics items had mean guessing index of.20460. Hence, NECO produces the most guessable items, followed by NABTEB, while WAEC had the least.

In summary, the study showed that in 2022, WAEC had the best items in term of guessing index, followed by those of NECO and then NABTEB, while in 2023 WAEC had the best items followed by NABTEB, and then NECO. This finding agreed with the findings of James and Olufunke (2020) 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 hypothesis 1, 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 7.

Table 7: 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 7 shows that there is significant difference in the mean difficulty indices of 2022 Mathematics MCQs 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 8.

Table 8: Multiple Comparisons of Difficulty Indices of 2022 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB

		Mean Difference (I-J)			95% Confidence Interval	
(I) EXAM	(J) EXAM		Std. Error	Sig.	Lower Bound	Upper Bound
NABTEB	WAEC	.826660	.426645	.163	-.20578	1.85910
	NECO	2.340673*	.408481	.000	1.35219	3.32916
WAEC	NABTEB	-.826660	.426645	.163	-1.85910	.20578
	NECO	1.514013*	.408481	.001	.52553	2.50250
NECO	NABTEB	-2.340673*	.408481	.000	-3.32916	-1.35219
	WAEC	-1.514013*	.408481	.001	-2.50250	-.52553

The post hoc comparison in Table 8 reveals that there is a significant difference between the difficulty indices of 2022 Mathematics Multiple-choice questions (MCQs) set by NABTEB and NECO (Mean difference=2.340673, SE=.408481, $p<.05$) in favour of NABTEB. Also, there is a significant difference between the difficulty indices of 2022 Mathematics MCQs set by WAEC and NECO (Mean difference =1.514013, SE=.408481, $p<.05$) 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 =.826660, SE=.426645, $p=.165$).

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 2, 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 9.

Table 9: *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 9 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 10

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

		Mean Difference (I-J)			95% Confidence Interval	
(I) EXAM	(J) EXAM		Std. Error	Sig.	Lower Bound	Upper Bound
NABTEB	WAEC	.956120	.415764	.069	-.05057	1.96281
	NECO	1.689627*	.411747	.000	.69266	2.68659
WAEC	NABTEB	-.956120	.415764	.069	-1.96281	.05057
	NECO	.733507	.411747	.231	-.26346	1.73047
NECO	NABTEB	-1.689627*	.411747	.000	-2.68659	-.69266
	WAEC	-.733507	.411747	.231	-1.73047	.26346

A post hoc comparison shows in Table 10 shows that there is a significant difference between the difficulty indices of 2023 Mathematics Multiple-choice questions (MCQs) set by NABTEB and NECO (Mean difference =1.689627, SE=.411747, $p<.05$) in favour of NABTEB. However, there is no significant difference between the difficulty Indices of 2023 Mathematics Multiple-choice questions (MCQs) set by WAEC and NECO (Mean difference =.733507, SE=.411747, p=.231). Also, there is no significant difference between the difficulty Indices of 2022 Mathematics Multiple-choice questions (MCQs) set by NABTEB and WAEC (Mean difference =.956120, SE=.415764, p=.069).

In summary, there was significant difference in the mean difficulty indices of 2022 and 2023 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB. The post hoc comparison revealed that there was a significant difference between the difficulty indices of 2022 Mathematics Multiple-choice questions (MCQs) set by NABTEB and NECO in favour of NABTEB. Also, there was a significant difference

between the difficulty indices of 2022 Mathematics Multiple-choice questions (MCQs) set by WAEC and NECO in favour of WAEC. However, there was no significant difference between the difficulty indices of 2022 Mathematics MCQs set by WAEC and NABTEB. Furthermore, the study shows that there was a significant difference between the difficulty indices of 2023 Mathematics MCQs set by NABTEB and NECO in favour of NABTEB. However, there was no significant difference between the difficulty Indices of 2023 Mathematics Multiple-choice questions (MCQs) set by WAEC and NECO. Also, there was no significant difference between the difficulty Indices of 2022 Mathematics MCQs set by NABTEB and WAEC. The findings agreed with the submission of Thomas et al. (2016) that NECO examinations is significant different from WAEC and NABTEB.

However, this finding disagreed with findings of studies carried out by Alfred (2013); Metibemu (2016); James and Olufunke (2020) which says that there was no significant difference between the difficulty level of WAEC, NABTEB and NECO multiple-choice items in mathematic..

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 hypothesis 3, 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 11.

Table 11: 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 11 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 12.

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

		Mean			95% Confidence Interval	
(I) EXAM	**(J) EXAM**	**Difference (I-J)**	**Std. Error**	**Sig.**	**Lower Bound**	**Upper Bound**
NABTEB	WAEC	.643420	.273998	.060	-.01963	1.30647
	NECO	1.177567*	.262333	.000	.54275	1.81239
WAEC	NABTEB	-.643420	.273998	.060	-1.30647	.01963
	NECO	.534147	.262333	.130	-.10067	1.16897
NECO	NABTEB	-1.177567*	.262333	.000	-1.81239	-.54275
	WAEC	-.534147	.262333	.130	-1.16897	.10067

A post hoc comparison shows in Table 12 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 =1.177567, SE=.262333, p<.05) in favour of NABTEB. However, there is no significant difference between the discrimination indices of 2023 Mathematics MCQs set by NABTEB and WAEC (Mean difference =.643420, SE=.273998, p=..060).

Also, there is no significant difference between the discrimination indices of 2022 Mathematics MCQs set by NECO and WAEC (Mean difference =-.534147, SE=.262333, 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 hypothesis 4, 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 13.

Table 13: *ANOVA 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 13 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 14.

Table 14: *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	-.354180*	.101069	.002	-.59876	-.10960
	NECO	.049717	.096767	1.000	-.18445	.28388
WAEC	NABTEB	.354180*	.101069	.002	.10960	.59876
	NECO	.403897*	.096767	.000	.16973	.63806
NECO	NABTEB	-.049717	.096767	1.000	-.28388	.18445
	WAEC	-.403897*	.096767	.000	-.63806	-.16973

A post hoc comparison shows in Table 14 reveals that there is a significant difference between the discrimination indices of 2023 Mathematics Multiple-choice questions (MCQs) set by NABTEB and WAEC (Mean difference =.354180, SE=.101069, $p<.05$) 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 =.403897, SE=.096767, $p<.05$) in favour of WAEC. However, there is no significant difference between the discrimination indices of 2023 Mathematics Multiple-choice questions (MCQs) set by NECO and WAEC (Mean difference =.049717 SE=.096767, $p=1.00$).

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 15.

Table 15: 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 15 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 16.

Table 16: Multiple Comparisons Guessing Indices of 2022 Mathematics Multiple-choice questions (MCQs) set by WAEC, NECO and NABTEB.

		Mean Difference (I-J)			95% Confidence Interval	
(I) EXAM	**(J) EXAM**		**Std. Error**	**Sig.**	**Lower Bound**	**Upper Bound**
NABTEB	WAEC	.075000*	.024014	.006	.01689	.13311
	NECO	.053217	.022992	.066	-.00242	.10885
WAEC	NABTEB	-.075000*	.024014	.006	-.13311	-.01689
	NECO	-.021783	.022992	1.000	-.07742	.03385
NECO	NABTEB	-.053217	.022992	.066	-.10885	.00242
	WAEC	.021783	.022992	1.000	-.03385	.07742

A post hoc comparison shows in Table 16 shows that there is a significant difference between the guessing indices of 2022 Mathematics Multiple-choice questions (MCQs) set by NABTEB and WAEC (Mean difference =.075000, SE=.024014, p<.05) in favour of WAEC. However, there is no significant difference between the guessing indices of 2022 Mathematics Multiple-choice questions (MCQs) set by NABTEB and NECO (Mean difference =053217, SE=.022992, p=.066). Also, there is no significant difference between the guessing indices of 2022 Mathematics Multiple-choice questions (MCQs) set by WAEC and NECO (Mean difference =.021783, SE=.022992, p=1.000).

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 hypothesis 6, 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 17.

Table 17: *ANOVA 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 17 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 ($F(2,159) = 1.837$; $p < .05$).

CONCLUSION

The result of the study shows that there is significant difference between the 2022-2023 NABTEB, WAEC and NECO Mathematics multiple choice items with WAEC producing the best test items, followed by NABTEB and then those of NECO.

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 distracter 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 sbody 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.

REFERENCES

Aborisade, O. J., & Fajobi, O. O. (2020). Comparative analysis of psychometric properties of mathematics items constructed by WAEC and NECO in Nigeria using item response theory approach. *Educational Research and Reviews, 15*(1), 1–7. https://doi.org/10.5897/ERR2019.3850

Alfred, A. O. (2013). *Scoring tests with dichotomous and polychromous items*. University of Georgia.

Asuquo, U. B., Joseph, E., & Bassey, H. E. (2022). 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, 6*(6), 311–319. https://doi.org/10.36348/jaep.2022.v06i06.001

Awopeju, O. A., & Afolabi, E. R. I. (2016). Comparative analysis of classical test theory and item response theory-based item parameter estimates of senior school certificate mathematics examination. *European Scientific Journal, 12*(28), 263–284. https://doi.org/10.19044/esj.2016.v12n28p263

Bandele, S. O., & Adewale, A. E. (2013). Comparative analysis of the reliability and validity coefficients of WAEC, NECO and NABTEB constructed mathematics examination. *Journal of Educational and Social Research, 3*(2), 397–402. https://doi.org/10.5901/jesr.2013.v3n2p397

Fajobi, O. O. (2024). 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, 17*(3), 302–314. https://unijerps.org/index.php/unijerps/article/view/694

Kanzow, A. F., Schmidt, D., & Kanzow, P. (2023). Scoring single-response multiple-choice items: Scoping review and comparison of different scoring methods. *JMIR Medical Education, 9*, Article e44084. https://doi.org/10.2196/44084

Metibemu, M. A. (2016). *Comparison of classical test theory and item response theory frameworks in the development and equating of physics achievement tests in Ondo State, Nigeria* [Unpublished doctoral dissertation]. University of Ibadan.

National Bureau of Statistics. (2025). *Statistics of entries and results, 2004–2012*.

National Business and Technical Examinations Board. (2020). *No. 6 on the grading system for NBC/NTC (ordinary level)* [Information bulletin].

Nwankwo, O. O. (2023). The item difficulty and discrimination indices of Senior School Certificate Examination (SSCE) multiple-choice chemistry tests of the West African Examinations Council (WAEC) and National Examinations Council (NECO) using classical test theory in Benue State. *International Journal of Asian Social Sciences, 3*(4), 130–135.

Oghenerume, R. A. (2022). *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.

Oghenerume, R. A. (2025). Item statistics disparity between 2023 WASSCE and NECO SSCE mathematics large-scale assessments. *International Journal of Educational Researchers, 16*(1), 1–14. https://doi.org/10.29329/ijer.2025.1295.1

Oghenerume, R. A., & Egberha, F. (2024). 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, 15*(1), 58–67. https://doi.org/10.29329/ijer.2024.652.4

Ogunbamowo, A. O., Adediwura, A. A., & Diyan, R. O. (2019). 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, 3*(1), 9–18. https://doi.org/10.21276/jaep.2019.3.1.2

Oyiborhoro, A. V., Odili, J. N., & Osadebe, P. U. (2023). 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, 3*(7), 1–13. https://www.openjournals.ijaar.org/index.php/ijresd/article/view/27

Oyiborhoro, P. E., Odili, J. N., & Osadebe, P. U. (2023). Assessment of concurrent validity of Delta State BECE basic science multiple-choice items using item response theory. *African Journal of Educational Assessment, 19*(2), 55–67.