ORIGINAL ARTICLE

Comparative Analysis of Item Statistics of WASSCE and NECO SSCE 2023 Data Processing Multiple Choice Tests Using Item Response Theory

Rioborue Alexander Oghenerume[1]*, & Friday Egberha[2]

[1] MEd, Measurement and Evaluation, University of Benin, Benin City, Nigeria.
ORCID: 0009-0008-7808-4810

[2] PhD, Measurement and Evaluation, University of Benin, Benin City, Nigeria.
ORCID: 0009-0002-6754-7943

Conflict of Interest
No conflict of interest is present in the conduction or the reporting of this study.

ABSTRACT

This investigation aimed to comparatively analyze the item statistics of West African Senior School Examination (WASSCE) and National Examinations Council Senior School Certificate Examination (NECO SSCE) 2023 Data Processing multiple choice tests using IRT 3PLM. Three hypothesized questions were expressed. In this study, survey research design was utilized. The population of this work comprised of 10,800 Data Processing SS 3 students in 21 public senior secondary schools in Abuja Municipal Area Council (AMAC). Multistage sampling procedure was used to pick a sample size of 1,080 Data Processing SS 3 students for this study. The instruments used for the current study were WASSCE and NECO SSCE 2023 Data Processing multiple choice tests. The data collected were subjected to IRT 3PLM using jMetrik psychometric software to generate the item parameters. All hypothesized questions were tested for significance utilizing independent samples t-test statistics. The findings of the study indicated no significant differences in the discrimination and difficulty parameters between WASSCE and NECO SSCE 2023 Data Processing multiple choice items. However, a significant difference existed in the guessing parameters between the two. The practical implication highlighted in the findings suggested that employing IRT 3PLM to determine the item statistics before administering the tests could further enhance the quality of Data Processing multiple choice items constructed by WAEC and NECO.

Keywords: Data Processing, Item Statistics, West African Senior School Certificate Examination (WASSCE), National Examinations Council Senior School Certificate Examination (NECO SSCE), Item Response Theory Three Parameter Logistic Model (IRT 3PLM).

*Corresponding Author
Rioborue Alexander Oghenerume, Measurement and Evaluation, University of Benin, Benin City, Nigeria
Email: raoghenerume@gmail.com

INTRODUCTION

As expounded by Oghenerume (2022), Data Processing is a subject offered in Nigerian senior secondary schools, and it covers topics related to the collection, processing, analysis, and presentation of data using various tools and techniques. While it is elective, students who are interested in pursuing careers in fields such as Business Analytics, Financial Accounting, Statistics, Economics, Data Science, and Information Technology often take Data Processing as an addition to their core subjects, which is usually taught alongside other subjects such as Mathematics, Statistics, and Computer Science. To discern the magnitude to which the students have been equipped with the skills and knowledge needed to understand, analyze, and present data appropriately, a test is usually conducted. The test undergoes item statistics computation to assess the quality of each item and identify any that may need revision or elimination. Item statistics are quantitative measures that capture the numerical properties and characteristics of the test items. These quantitative measures encompass various aspects of the test items, including their parameters (Oghenerume, 2022). For an item, to be considered a relevant item, it should satisfy the three parameters: difficulty, discrimination, and guessing. Difficulty is how challenging an item is for test-takers, discrimination indicates an item's adeptness to split amid test-takers with different points of the construct being assessed, and guessing quantifies the possibility that a test-taker with minimal ability would answer an item correctly purely by chance.

Data Processing tests usually consist of both objective and subjective items that assess the students' knowledge and understanding of the subject. The objective items are usually in the form of multiple choice items that test the students' ability to recall and apply the concepts and principles of Data Processing. The subjective items require the students to write essays or solve problems that demonstrate their analytical ability. Additionally, students are required to undertake practical projects that allow them to apply the skills and knowledge learnt in Data Processing to real life situations (Oghenerume, 2022). The two major bodies responsible for conducting and grading Data Processing tests for Senior Secondary 3 (SS 3) students in Nigeria are the West African Examinations Council (WAEC) and National Examinations Council (NECO). Both WAEC and NECO are in charge of Senior Secondary Certificate Examinations (SSCE). While WAEC conducts the West African Senior School Certificate Examination (WASSCE) for students in Nigeria, Gambia, Ghana, Sierra Leone, and Liberia, NECO is responsible for the students in Nigeria writing the National Examinations Council Senior School Certificate Examination (NECO SSCE). The NECO SSCE is similar to the WASSCE, but it is primarily taken by students in Nigeria.

Literature Review

The theoretical framework for this study is hinged on Item Response Theory (IRT) by Thurstone (1925) and Three-Parameter Logistic Model (3PLM) by Birnbaum (1968). In 1925, Louis Thurstone (1925) propounded Item Response Theory (IRT) and established its conceptual framework in his paper titled, "A Method of Scaling Psychological and Educational Tests." In this influential paper, Thurstone provided a procedure for accurately positioning the Binet scale's items. Cai and Huang (2022) noted that IRT can be viewed minimally as a set of psychometric models for categorical item-level response data. von Davier et al. (2021) described categorical item-level response data as data collected from individuals responding to items with multiple categories or response options which are typically discrete and qualitative, representing different levels or alternatives. Examples include multiple choice options with predefined categories or

categorical response options including strongly agree, agree, disagree, and strongly disagree. Birnbaum (1968) postulated that the Three-Parameter Logistic Model (3PLM) is rooted in IRT. As a result, the possibility of a correct reply to an item depends upon three parameters: slope 'a' (discrimination), threshold 'b' (difficulty), and lower asymptote 'c' (guessing). 3PLM allows items to vary in their adeptness to split amid test-takers of varying proficiency levels and in their difficulty, while also accounting for the possibility of test-takers with extremely minimal proficiency points guessing the correct replies to items. The psychometric formulae of 3PLM describes the possibility that an unintentionally chosen test-taker with competence 'θ' upon 'k' measure will perfectly reply to 'j' item:

$$P(x_j = 1 \mid \theta_k, a_j, b_j, c_j) = c_j + \frac{(1 - c_j)}{1 + e^{-Da_j(\theta_k - b_j)}}$$

In which

xj = reply to j item;

aj = discrimination parameter of j item, characterising the slope;

bj = difficulty of j item, characterising the threshold;

cj = guessing of j item, accounting for the possibility of test-takers with extremely low proficiency points guessing the correct reply;

D = subjective measuring factor usually fixed to 1.7 to imprecise outcomes in a conventional ogive prototype.

Both IRT and 3PLM operate on certain underlying assumptions when applied to the data analytics of test, particularly in WASSCE and NECO SSCE 2023 Data Processing multiple choice tests. de Ayala (2022) and Bock and Gibbons (2021) elucidated these assumptions. Firstly, it is assumed the modelled hidden trait accurately describes the entirety of the hidden competence gap and effectively account for test-taker accomplishment. In numerous instances, countless tests assume the requirement of a single latent ability (unidimensionality). Secondly, it is assumed that all items are independent of one another locally when conditioned on the hidden competence. This implies an examinee's reply to a particular item is not influenced by their reply to other items. Lastly, it is posited that an examinee's reply to a test item can be effectively modelled using the Item Response Function (IRF). Being contingent upon the specific kind of model in IRT, such as 3PLM, IRF expresses the possibility of attaining a certain grade 'X' allowing every level of the hidden ability, denoted as θ. Such possibility association remains neutral of the θ in the distribution within the populace. de Ayala (2022) explained that slope is expressed as a value between ₋2 and +2, where a value below 1 may indicate weaker discrimination, while a value above 1.5 may suggest excessive discrimination. Threshold is represented by values ranging from ₋3 to +3. Difficulty values below 1.5 indicate easier items, whereas values above 2.5 indicate more difficult items. Guessing is expressed as a value between 0 and 1. Bock and Gibbons (2021) affirmed that there is no universally agreed-upon specific numerical value for a 'moderate' slope, threshold, and lower asymptote. However, a commonly referenced guideline suggests a slope value between 0 and 2, a threshold value between ₋3 and +3, and a lower asymptote's from 0 to 1. Essentially, lower asymptote values exceeding 0.35 are judged unacceptable. 5-option and 4-option items are bound to have lower asymptote rates of about 0.20 and 0.25 respectively.

Oghenerume (2022) conducted a study titled "Comparative Analysis of Item Statistics of WASSCE and NECO SSCE

2022 Data Processing Multiple Choice Tests Using Item Response Theory." The study analytically compared the item statistics of WASSCE and NECO SSCE 2022 Data Processing multiple choice tests using IRT 2PLM. Six hypothesized questions were affirmed. In this study, survey research design was utilized. This study's population comprised ten thousand, eight hundred Data Processing SS 3 students in 21 public senior secondary schools in AMAC. Multistage sampling procedure was utilized to pick a sample size of one thousand and eighty Data Processing SS 3 students for this study. The WASSCE and NECO SSCE 2022 Data Processing multiple choice tests were the instruments utilized for the study. The data collected were subjected to IRT 2PLM using Item Response Theory Assistant for Excel (EIRT) to statistically generate the item parameters (slope and threshold). All hypothesized questions were tested for significance utilizing independent samples t-test statistics. The findings of this study revealed a significant difference in the girls' discrimination indices. Significantly, no difference existed in the discrimination indices for boys, as well as in the difficulty levels for boys and girls. Additionally, no significant differences were found in the discrimination indices and difficulty levels between WASSCE and NECO SSCE 2022 Data Processing multiple choice items. Due to these findings, this study has recommended that WAEC and NECO need to harness the IRT 2PLM to further improve the quality of Data Processing multiple choice items.

Aborisade and Fajobi (2020) conducted a study titled "Comparative Analysis of Psychometric Properties of Mathematics Items Constructed by WAEC and NECO in Nigeria Using Item Response Theory Approach." The study explored the comparability of psychometric properties of Mathematics items constructed by the West Africa Examinations Council (WAEC) and National Examinations Council (NECO). This study utilized IRT approach and focused on difficulty index, discrimination power, and distractor parameter. Survey type of descriptive research design was adopted in this study. The population consisted of senior students in secondary schools who took the WAEC and NECO tests in the South Western Nigeria in 2019. A sample size which comprised 1200 senior students in secondary schools, chosen through a multistage sampling process were used. The objective test items constructed by both WAEC and NECO are used as instruments. The findings indicated that the difficulty and guessing indicators of the Mathematics items structured by WAEC and NECO are comparable. However, the discriminating power of the items is not comparable between the two examination bodies. Based on these findings, the study suggested that credentials released by WAEC and NECO can be implemented for the same intentions without sentiments, as the quality of the items developed by them is equivalent.

Ogunbamowo et al. (2019) conducted a study titled "Psychometric Properties of 2017 West African Examination Council and National Examinations Council's Economic Senior School Certificate Examination Items." The study assessed the differences between WAEC and NECO Economics items' discrimination, difficulty, and guessing employing Item Response Theory and Classical Test Theory. The research design employed was descriptive. The population for the study consisted of secondary school students in Osun State, with a sample size of 540 students. The study utilized instruments adopted from the 2017 Economics WAEC and NECO Senior School Certificate Examinations, specifically Economics Achievement Test 1 (EAT 1) and Test 2 (EAT 2). The effects signified there was no significant difference between the discrimination indices of NECO and WAEC Economics items when CTT was used. Moreover, a significant difference was observed in the discrimination indices when IRT was applied. Similarly, while there was no significant difference in the difficulty indices of NECO and WAEC Economics test items with the use of CTT, a significant difference emerged when IRT was employed. In contrast, the difference in the guessing indications of NECO and WAEC Economics items was not significant. In conclusion, this study found a comparability of both tests under CTT, and a dissimilarity under IRT.

Statement of Problem

Over the years, both WAEC and NECO have not been immune to criticism. Daniel (2005) criticized NECO for its inferior test item quality and the credibility of the grades assigned to candidates. These deficiencies led to some federal Universities rejecting NECO results between 2002 and 2012, as documented by Peter (2012). On the contrary, Ahmed (2014) asserted that NECO test items from 2011 to 2014 exhibited higher standards compared to those of WAEC. Additionally, Dibu-Ojerinde and Faleye (2005) argued that there was no significant difference between NECO and WAEC, while Oghenerume (2022) affirmed that no significant differences were found in the discrimination indices and difficulty levels between WASSCE and NECO SSCE 2022 multiple choice items in Data Processing when a comparison was made. Statistically, the results released by WAEC and NECO (2023) revealed that only 84.38% of candidates who sat for WASSCE 2023 Data Processing test earned credit, while 61.60% of candidates who sat for NECO SSCE 2023 Data Processing test earned credit. Notably, 16.29% of WAEC candidates' results were withheld due to examination malpractice, showing a 6.54% decrease compared to the 22.83% recorded in 2022. Comparatively, Data Processing tests were conducted for the same group of candidates with the expectation that they had been effectively taught by qualified Data Processing teachers who covered the syllabi. Additionally, both WAEC and NECO have similar syllabi for Data Processing and issue senior secondary school certificates with the same goals. Consequently, their test items are expected to be of similar standards, resulting in similar outcomes. Perhaps the underperformance of certain candidates could be linked to the parameters within the multiple choice items of Data Processing tests. Given these issues, this study undertook a comparative analysis of item statistics for the 2023 WASSCE and NECO SSCE Data Processing multiple choice tests. This analysis involved a comparative assessment of the three parameters: discrimination, difficulty, and guessing in the multiple choice items using IRT 3PLM.

Research Questions

1. Is there any significant difference between the difficulty parameters of WASSCE and NECO SSCE 2023 Data Processing multiple choice items?

2. Is there any significant difference between the discrimination parameters of WASSCE and NECO SSCE 2023 Data Processing multiple choice items?

3. Is there any significant difference between the guessing parameters of WASSCE and NECO SSCE 2023 Data Processing multiple choice items?

Hypotheses

1. No significant difference lies between the difficulty parameters of WASSCE and NECO SSCE 2023 Data Processing multiple choice items.

2. No significant difference exists between the discrimination parameters of WASSCE and NECO SSCE 2023 Data Processing multiple choice items.

3. No significant difference occurs between the guessing parameters of WASSCE and NECO SSCE 2023 Data Processing multiple choice items.

METHOD

Research Design

Survey research design was utilized in this investigation because it involved a group of students from which data were obtained, only from a few students considered as emissaries of the all-inclusive group needed. This design is suitable for this study because both WASSCE and NECO SSCE 2023 Data Processing multiple choice test items were employed to obtain data and analyze difficulty, discrimination, and guessing parameters of the items using the IRT 3PLM.

Population, Sample and Sampling Procedure

According to the Federal Capital Territory Secondary Education Board (2023), the population in this study encompassed 10,800 Data Processing SS 3 students representing 5,346 males and 5,454 females in 21 public senior secondary schools in Abuja Municipal Area Council (AMAC). However, statistical population of this study involved 40 Data Processing multiple choice test items of WASSCE and 60 Data Processing multiple choice test items of NECO SSCE. The study sample comprised 1,080 Data Processing SS3 students from four schools in AMAC, representing 10%. They were picked employing a multistage sampling process. During the first stage, the simple random sampling was utilized to pick four schools through a balloting method. During the second's, stratified random sampling was employed to stratify for sex within the four schools in AMAC. At the third stage, a simple random sampling through the method of balloting was implemented to choose 135 boys and 135 girls in each school. From each school, 270 students were sampled, and therefore, a total of 1,080 students were sampled from the four selected schools.

Instruments

The instruments utilized in this investigation are the multiple choice items in the WASSCE and NECO SSCE 2023 Data Processing tests. The Data Processing multiple choice items in WASSCE 2023 consist of 40 items, while the NECO SSCE 2023 contains 60 Data Processing multiple choice items. In the WASSCE, the Data Processing multiple choice items are structured with four options 'A' to 'D,' from which Data Processing students are required to select the correct option. In the NECO SSCE, they are structured with five options 'A' to 'E,' and students are required to indicate the correct option. These instruments were deemed to be very valid and highly reliable because WAEC and NECO are reputable for constructing standardized tests and items were still within the three-year limit required before re-establishment of reliability.

Method of Data Collection

The standardized instruments were dispensed directly to the Data Processing SS 3 students in the selected schools in AMAC with the help of the research assistants who are also their Data Processing teachers. The WASSCE multiple choice items were administered first. Upon completion, they were immediately collected. After a day, the NECO SSCE multiple choice items were administered and collected through the same procedure.

Method of Data Analysis

The collected data were subjected to Item Response Theory Three Parameter Logistic Model (IRT 3PLM) utilizing jMetrik psychometric software to analytically generate the item parameters. All hypothesized questions were tested for significance harnessing independent samples t-test statistics in IBM SPSS.

RESULTS AND DISCUSSION

Presentation of Results

Hypothesis One: No significant difference lies between the difficulty parameters of WASSCE and NECO SSCE 2023 Data Processing multiple choice items.

Table 1: Thresholds in WASSCE and NECO SSCE 2023 Data Processing Multiple Choice Items

Data Processing	Items' Number	Mean	Std. Dev.	df	t-value	p-value
WASSCE 2023	40	2.1168	.39156			
				98	1.243	.217
NECO SSCE 2023	60	2.0143	.41154			

α = .05.

Table 1 revealed the number of items in the WASSCE 2023 Data Processing multiple choice test is N=40, while for the NECO SSCE 2023 Data Processing multiple choice test, it was N=60. Their mean values were 2.1168 and 2.0143, while their standard deviations were .39156 and .41154 respectively. The degree of freedom, t-value, and p-value were 98, 1.243, and .217 respectively. Since the p-value was more than the alpha's, the hypothesis which affirms "no significant difference between the difficulty parameters of WASSCE and NECO SSCE 2023 Data Processing multiple choice items," was retained. This implies that neither of the Data Processing multiple choice tests constructed by both examination bodies, WAEC and NECO, was more difficult than the other.

Hypothesis Two: No significant difference exists between the discrimination parameters of WASSCE and NECO SSCE 2023 Data Processing multiple choice items.

Table 2: Slopes in WASSCE and NECO SSCE 2023 Data Processing Multiple Choice Items

Data Processing	Items' Number	Mean	Std. Dev.	df	t-value	p-value
WASSCE 2023	40	1.0785	.36812			
				98	.844	.400
NECO SSCE 2023	60	.9840	.64000			

α = .05.

Table 2 demonstrated that the number of items in the WASSCE 2023 Data Processing multiple choice test was N=40, while for the NECO SSCE 2023 Data Processing multiple choice test, it was N=60. Their mean values were 1.0785 and .9840, while their standard deviations were .36812 and .64000 respectively. The degree of freedom, t-value, and p-value were 98, .844, and .400. Since the p-value was more than the alpha's, the hypothesis which affirms "no significant difference between the discrimination parameters of WASSCE and NECO SSCE 2023 Data Processing multiple choice items," was retained. This implies that neither of the Data Processing multiple choice tests constructed by both examination bodies, WAEC and NECO, was more discriminating than the other.

Hypothesis Three: No significant difference occurs between the guessing parameters of WASSCE and NECO SSCE 2023 Data Processing multiple choice items.

Table 3: Lower Asymptotes in WASSCE and NECO SSCE 2023 Data Processing Multiple Choice Items

Data Processing	Items' Number	Mean	Std. Dev.	df	t-value	p-value
WASSCE 2023	40	.0787	.11665			
				98	2.899	.005
NECO SSCE 2023	60	.0247	.06961			

α = .05.   p < .05. (Significant).

Table 3 indicated the items' number in the WASSCE 2023 Data Processing multiple choice test was N=40, while for the NECO SSCE 2023 Data Processing multiple choice test was N=60. Their mean values were .0787 and .0247, while their standard deviations were .11665 and .06961 respectively. The degree of freedom was 98, t-value, and p-value were , 2.899, and .005 respectively. Since the p-value was less than the alpha's, the hypothesis which affirms "no significant difference occurs between the guessing parameters of WASSCE and NECO SSCE 2023 Data Processing multiple choice items," was rejected. This implies that the Data Processing multiple choice test constructed by WAEC had more items that were more susceptible to guessing than the corresponding NECO.

Discussion of Findings

Hypothesis one revealed neither of the Data Processing multiple choice tests constructed by both examination bodies, WAEC and NECO, was more difficult than the other. This finding aligns with the finding of Oghenerume (2022), who conducted a comparative analysis of item statistics for the WASSCE and NECO SSCE 2022 Data Processing multiple choice tests using item response theory and found no significant difference in the difficulty levels between the WASSCE and NECO SSCE 2022 Data Processing multiple choice items. It also agrees with the findings of Aborisade and Fajobi (2020), who conducted a comparative analysis of the psychometric properties of Mathematics items constructed by WAEC and NECO in Nigeria using the IRT approach. They did not find any significant difference in the difficulty levels of WASSCE and NECO SSCE Mathematics items; instead, they are comparable. Additionally, this supports the findings of Ogunbamowo et al. (2019), who investigated the psychometric properties of the 2017 West African Examinations Council and National Examinations Council's Economics Senior School Certificate Examination items. They did not find any significant difference amid the difficulty indices of WASSCE and NECO SSCE Economics multiple choice items using CTT. Conversely, it opposes their finding that a significant difference existed amid the difficulty parameter using IRT. This discrepancy may be due to the sample size of 540 they used, as IRT typically relies on a sample size closer to 1,000.

Hypothesis two showed neither of the Data Processing multiple choice tests constructed by both examination bodies, WAEC and NECO, was more discriminating than the other. This finding aligns with the finding of Oghenerume (2022), who conducted a comparative analysis of item statistics for the WASSCE and NECO SSCE 2022 Data Processing multiple choice tests using item response theory and found no significant difference in the discrimination indices between WASSCE and NECO SSCE 2022 Data Processing multiple choice items. This supports the findings of Ogunbamowo et al. (2019), who investigated the psychometric properties of the 2017 West African Examinations Council and National Examinations Council's Economics Senior School Certificate Examination items. They did not find any significant

difference in the discrimination indices of WASSCE and NECO SSCE Economics multiple choice items using CTT. Moreover, it opposes their finding that a significant difference existed in the discrimination parameter of WASSCE and NECO SSCE multiple choice items in Economics using IRT. This discrepancy may be attributed to the sample size of 540 they used, as IRT typically relies on a sample size closer to 1,000. Additionally, it disagrees with the findings of Aborisade and Fajobi (2020), who conducted a comparative analysis of the psychometric properties of Mathematics items constructed by WAEC and NECO in Nigeria using the IRT approach. They found a significant difference in the discrimination powers of WASSCE and NECO SSCE Mathematics items. Perhaps the WAEC 2017 Mathematics objective test with 50 items they used as one of their instruments could be the reason for the disparity, whereas the WASSCE 2023 Data Processing multiple choice test has 40 items.

Hypothesis three divulged the Data Processing multiple choice test constructed by WAEC had more items that were more susceptible to the guessing than the corresponding NECO test. This finding is not in alignment with the findings of Aborisade and Fajobi (2020), who conducted a comparative analysis of the psychometric properties of Mathematics items constructed by WAEC and NECO in Nigeria using the IRT approach. They did not find any significant difference in the guessing parameters of WASSCE and NECO SSCE Mathematics items; instead, they are comparable. Perhaps the WAEC 2017 Mathematics objective test with 50 items they used as one of their instruments could be the reason for the disparity, whereas the WASSCE 2023 Data Processing multiple choice test includes 40 items. Additionally, this disagrees with the findings of Ogunbamowo et al. (2019), who investigated the psychometric properties of the 2017 West African Examinations Council (WAEC) and National Examinations Council (NECO) Economics Senior School Certificate Examination items. They did not find any significant difference in the guessing parameters of WASSCE and NECO SSCE multiple choice items in Economics using IRT. Perhaps this discrepancy could possibly be attributed to the sample size of 540 they used, as IRT typically relies on a sample size closer to 1,000.

CONCLUSION AND RECOMMENDATIONS

This study aimed to comparatively analyze the item statistics of WASSCE and NECO SSCE 2023 Data Processing multiple choice tests using the Item Response Theory Three-Parameter Logistic Model (IRT 3PLM). Due to the findings in this study, the conclusion showed no significant differences in the discrimination and difficulty parameters between WASSCE and NECO SSCE 2023 Data Processing multiple choice items. However, a significant difference existed in the guessing parameters between these two. The practical implication highlighted in these findings suggested that employing IRT 3PLM to determine the item statistics before administering the tests could further enhance the quality of Data Processing multiple choice items constructed by WAEC and NECO.

Due to the findings of this study, it is recommended that WAEC and NECO employ the IRT 3PLM to determine the discrimination, difficulty, and guessing parameters of their test items in Data Processing before administering them if they are not presently applying it. This will further enhance the quality of Data Processing multiple choice items they construct.

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.

Ahmed, M. F. (2014). Difficulty index of Mathematics multiple-choice items of West African Examinations Council and National Examinations Council Senior Secondary School Certificate Examinations from 2006–2010. *Journal of ATIP, 13*, 25–31.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick, *Statistical theories of mental test scores* (pp. 397–472). Reading, MA: Addison-Wesley Publishing.

Bock, R. D., & Gibbons, R. D. (2021). *Item response theory*. Wiley.

Cai, L., & Huang, S. (2022). Multidimensional item response theory. In *International Encyclopedia of Education* (4th ed., pp. 72–85). Elsevier.

Daniel, F. (2005). A survey of teachers' and students' opinions on WAEC and NECO SSCE examinations in Mathematics. *Abuja Journal of Education, 6*(1), 32–39.

de Ayala, R. J. (2022). *The theory and practice of item response theory (2nd ed.).* The Guilford Press.

Dibu-Ojerinde, O. O., & Faleye, B. A. (2005). Do they end at the same point? *European Journal of Scientific Research, 11*(3), 449–452.

Federal Capital Territory Secondary Education Board (2023). *Statistics of students offering Data Processing.*

National Examinations Council. (2023). *Statistical report.* https://neco.gov.ng

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.

Ogunbamowo, A. O., Adediwura, A. A., & Diyan, R. O. (2019). 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, 3*(1), 9–18.

Peter, K. (2012). A study of the attitude of some Nigerian science students towards NECO and WAEC. *Journal of Professional Science and Vocational Teachers Association of Nigeria, 12*(1), 15–18.

Thurstone, L. L. (1925). A method of scaling psychological and educational tests. *Journal of Educational Psychology, 16*(7), 433–451. https://doi.org/10.1037/h0073357

von Davier, A. A., Mislevy, R. J., & Hao, J. (Eds.). (2021). *Computational psychometrics: New methodologies for a new generation of digital learning and assessment (1st ed.).* Springer Cham. https://doi.org/10.1007/978-3-030-74394-9

West Africa Examinations Council. (2023). *Statistical report.* https://www.waeconline.org.ng