UNIJERPS
Unizik Journal of Educational Research and Policy Studies VOL.17 (3);
January-June, 2024 https://unijerps.org
2023 Impact Factor: **6.834;** 2022 *Impact Factor*: 5.156; *2021 Impact Factor: 5.641;*
2020 Impact Factor: 5.355. *Indexing and Impact Factor URL: http://sjifactor.com/passport.php?id=21363*

EXAMINING THE PSYCHOMETRIC PROPERTIES OF WAEC AND NECO MATHEMATICS ITEMS FOR NIGERIAN SECONDARY SCHOOLS USING A 3-PARAMETER LOGISTIC MODEL

Olutoyin Olufunke Fajobi (Ph.D)
Department of Education Guidance and Counselling,
Faculty of Education, Federal University, Oye-Ekiti, Nigeria
E-mail: olutoyin.fajobi@fuoye.edu.ng

ABSTRACT

The quality of examination items can be determined by analysing their psychometric properties. To analyse the levels of difficulty and guessing indices, and the discriminating powers of Mathematics multiple-choice tests, this study examined the psychometric properties of mathematics test items developed by the West Africa Examination Council (WAEC) and National Examination Council (NECO) for Nigerian secondary school students. The study employed the ex-post facto research type of the non-experimental research design. The population of the study comprised public Senior Secondary School three (SSS 3) students in Ekiti State, Nigeria. The sample of the study comprised 1,200 public Senior Secondary School three (SSS 3) students randomly selected from twenty-four (24) Public Senior Secondary Schools in Ekiti State, Nigeria using Multistage Sampling Procedure. Two instruments which consisted of fifty (50) Mathematics objective test items constructed by WAEC and 60 objective test items constructed by NECO were used to collect data for the study. One research question was generated for the study. Bilog MG software statistical analysis was used to generate difficulty levels (b), discrimination powers (a), and guessing (c) parameter estimates on WAEC and NECO items which confirmed the fitness model. The findings of the study showed that WAEC items discriminate better than NECO items. NECO items are more difficult compared to WAEC items; while the guessing parameter of WAEC and NECO are of comparable standard. It is, therefore, recommended that NECO, WAEC, and other national examination bodies adopt the IRT approach to analyse psychometric properties of test items and that WAEC and NECO examinations be used for certifying Senior Secondary School students.

Keywords: Difficulty level; Discriminating Power; Guessing; Indices, IRT, 3-Parameter logistic model; WAEC and NECO Examinations.

Introduction

In Nigeria, the Senior Secondary School Certificate Examinations are administered mainly by the West Africa Examination Council (WAEC) and the National Examination Council (NECO). The Certificates issued by these examination bodies are used for purposes such as admission into higher institutions of learning or employment. However, observations have shown that some students pass very well in NECO examinations, and performed below average in WAEC examinations despite the fact that they have similar syllabi and conduct

examinations for the same sets of students (Aborisade & Fajobi, 2020; Olawole & Olugbemi, 2022). For instance, in 2016 and 2017, Ekiti State came first in the NECO examination in the country which is not so in the case of WAEC even though the students who sat for these examinations were taught by the same teacher and the examination bodies are using similar syllabi (Aborisade & Fajobi, 2020; Amoka, 2020; Olawole & Olotu, 2018).

The variance in the performance of students in Mathematics in both WAEC and NECO was observed; out of 13,617 students registered for WAEC in 2014, 25% of the students passed with 5 credits including Mathematics while 58% passed NECO examination. In 2015 out of 13,690 registered for WAEC 5040 (37%) passed while 11,424 registered in NECO, 627 (52%) passed. In 2016, 11412 registered for WAEC, 5976 (52%) passed while 6431 (65%) passed out of 9850 for NECO. In 2017, out of 11011 students registered for WAEC 3,205 (29%) passed while 10020 students registered for NECO, 7885 (78%) passed which make Ekiti State to be the 1st position in the nation (Aborisade & Fajobi, 2020; Amoka, 2020). With these results, it is evident that students perform better in WAEC than NECO examinations, however, the study of Bulus and Chukwuma (2021) reports otherwise. The psychometric properties of the items developed by these assessment bodies must thus be examined because they may account for the variations in student performance in these examinations (Aborisade & Fajobi, 2020).

Examination items can have certain inherent qualities known as psychometric features, which are used to evaluate applicants (Osiesi, 2019). These qualities include test item validity and reliability indices, difficulty indices, discriminating indices, the strength of item alternative distractors, and the power of distractors (Osiesi, 2019). The easier or lesser the difficulty, the higher the value for the item's value. The ability of an item to distinguish between examinees with higher abilities and those with lesser abilities is known as discrimination power, while distractive index shows how the alternative options in an item is plausible so as to know if the student really know the answer or guesses to pick from the options provided. Item analysis, according to Oshkosh (2002), is likely the most crucial method for boosting test efficacy.

According to the National Policy on Education (2014), pupils are required to take mathematics as one of the core topics in secondary school. It makes no difference whether these pupils are enrolled in Science, Commercial, Arts, or Social Science classes at the high school level. There is no doubt about the value of mathematics to both individuals and society. Consequently, in order to live a useful life and contribute to society as a whole, everyone needs to have some knowledge of the subject (Osiesi & Fajobi, 2019; Osiesi & Odinko, 2019).

Mathematics is fundamental to the understanding of science, complexities of modern technology and for several scientific developments (Adegoke, 2013; Osiesi, 2019; Osiesi & Fajobi, 2019). Despite the

importance of Mathematics to every facet of human endeavour, it is dishearten that research and data from the National Examination Bodies such as WAEC and NECO in Nigeria have shown a consistent poor and unstable performance in the subject (Fajobi, 2020), leading to students dreading or showing negative attitudes toward Mathematics. As a result of the importance and contribution of mathematics to national development in terms of science and technology, more attention is being paid by researchers on factors responsible for poor performance of students in this subject, both in national and external (public) examinations, to include lack of instructional materials, shortage of qualifies mathematics teachers for effective teaching, and students' interest in the subject among others (Osiesi & Fajobi, 2019; Olawole & Oniemola, 2013).

Oshkosh (2002) as cited in Fajobi (2020) and Osiesi (2019), the most crucial tool for enhancing test efficacy is by item analysis. It is a methodical approach to raising test and test item quality. Three types of crucial data about the calibre of test items are provided by an item analysis: item difficulty measures how difficult or easy an item is; item discrimination measures how item discriminates between testees who know the item and those who do not know it; while distractor index measure options of an item. The probability of answering an item correctly or achieving a specific response level in relation to a person's ability and the qualities of the item is described by the item response theory (IRT), also known as latent response theory (Aborisade & Fajobi, 2020). The ability range in IRT estimates is typically between -∞ to +∞ theoretically, but typically they range from +3.0 with exceptional abilities, according to Kolawole (2007). Three IRT models—three, two, and one parameter IRT models—are used to characterise the psychometric characteristics of test items that are dichotomously scored (Aborisade & Fajobi, 2020). The number of parameters needed to describe an item's psychometric properties differentiates various models. It is one parameter model (1-PL) if only the difficulty characteristic of the item is estimated; two parameters (2-PL) model if the difficulty and discriminating indices of the item are estimated; while it is three-parameter (3-PL) model entails estimate of the item's difficulty, discriminating ability, and guessing ability indices.

Studies regarding the psychometric properties of Nigerian examination bodies such as WAEC and NECO are few in the extant literature. Aborisade and Fajobi (2020) study ascertained the psychometric properties of the WAEC and NECO Mathematics examination items using the Item Response Theory approach. Findings show that the Mathematics examination items of both examination bodies have similar guessing and difficulty indices, but their discrimination power was different. Similar findings of Alfred (2013), Metibemu (2016), Kolawole (2007), Ogbebor (2017). In contrast, the studies of Adewale (2008), Bamidele and Adewale (2013), and Thomas et al. (2016) report a difference in the guessing and difficulty indices of these examination items, reaffirming that WAEC test items are easier than those in the NECO.

These variances in the literature prompted the current study. In this study, the three-parameter of the WAEC and NECO mathematics item were estimated. The 3-PL was used in this study to obtain difficulty levels, discriminating and guessing estimates on WAEC and NECO items, to ascertain the qualities of the mathematics test items constructed by the examination bodies. Since the level of performance of students could be affected by the quality of test items.

Statement of the problem

Anecdotal records show that students' performance in NECO examinations is somewhat higher than their performance in WAEC examinations despite the similarity in their syllabi. Regarding the psychometric characteristics of the items created by the examining bodies, these tests are anticipated to be of comparable standard. There is a chance that some of the numerous tests created and given by the testing organisations may have technical flaws. Students who take these exams are expected to perform as well on both examinations because they were exposed to the identical course materials and programme. The variation in students' performance may be due to the variations in or quality of the examination's test items. Therefore, it is essential to look into the psychometric characteristics of the items created by these testing organisations. Therefore, this study used a three-parameter logistic model to investigate the characteristics of the mathematics examination items created by the WAEC and NECO examination boards.

Research Question

The following research question was raised for the study:

1. What are the difficulty indices, discriminating power and distractor indices of mathematics examination items constructed by WAEC and NECO?

Research Method

Research Design: The study employed the expost facto research type of the non-experimental design. Population, Sampling Technique and Sample: 11,500 senior secondary school students in their third year (SSS 3) who registered for the April, May/June 2019 Mathematics Senior Secondary School Certificate Examinations of WAEC and NECO in the Nigerian state of Ekiti made up the study's population. Using a multistage sampling technique, a sample of 1,200 students was chosen for the study. To choose the two senatorial districts (Ekiti North and Ekiti Central) for the State, a straightforward random selection approach was utilised. In addition, ten (10) schools from each of the State's three senatorial districts were chosen using a simple random sample technique, and only SSS 3 pupils were chosen using a purposive sampling technique. Instruments and Instrumentation: The two instruments for the study included 60 mathematics objective test items created by NECO in 2017 and fifty (50) mathematics objective test items created by WAEC. 120 students from three secondary schools outside the sampled schools participated in the instruments' trial testing. A reliability coefficient of 0.70 was estimated for the WAEC objective test and a coefficient of

0.75 for the NECO objective test using the Kudar-Richardson formula 20 (KR-20). 1,200 students were given the tests under the same circumstances as the examination bodies. Data Analysis: For each of the WAEC and NECO items, a 3-parameter item characteristic model was created using statistical analysis within the Bilog MG Logistic software.

Presentation of Results

Research question: what are the difficulty indices, discriminating power and distractor indices of mathematics examination items constructed by WAEC and NECO?

Figure I: As illustrated for figure below, a sample of 3-parameter IRT item characteristic curves (ICC) produced for each of the WAEC and NECO items.

Fig 1: WAEC CURVE

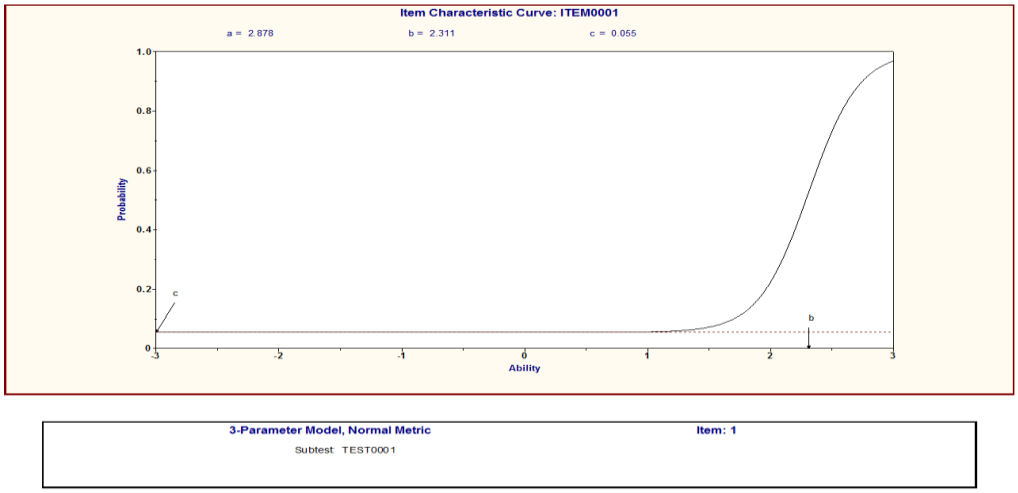


Table 1: Item Parameter Estimates for 50- WAEC Item.

ITEM	*MODEL*	*a*	*B*	*C*
WAEC1	L3	2.88	** 2.31	*** 0.06
WAEC2	L3	* 0.17	** 2.74	*** 0.18
WAEC3	L3	* 0.32	** 2.73	*** 0.08
WAEC4	L3	* 0.09	10.27	*** 0.06
WAEC5	L3	* 0.27	** 1.07	*** 0.09
WAEC6	L3	* 0.07	8.48	*** 0.07
WAEC7	L3	* 0.10	** 3.03	*** 0.09
WAEC8	L3	* 0.56	** 2.10	*** 0.10
WAEC9	L3	* 0.16	8.32	0.43
WAEC10	L3	* 0.08	** 2.81	*** 0.09
WAEC11	L3	* 0.00	0.09	*** 0.20
WAEC12	L3	* 0.07	** 0.68	*** 0.06

WAEC13	L3	* 0.71	** 1.84	*** 0.12
WAEC14	L3	3.67	** 1.00	*** 0.18
WAEC15	L3	* 0.07	** 1.03	*** 0.10
WAEC16	L3	* 0.16	5.24	0.36
WAEC17	L3	* 0.18	** 2.46	*** 0.11
WAEC18	L3	* 0.67	** 1.89	*** 0.05
WAEC19	L3	* 1.75	** 2.04	*** 0.18
WAEC20	L3	* 0.21	7.84	*** 0.14
WAEC21	L3	5.16	** 1.11	*** 0.16
WAEC22	L3	* 0.12	9.61	*** 0.07
WAEC23	L3	* 0.10	** 2.05	*** 0.06
WAEC24	L3	* 0.10	** 0.64	*** 0.08
WAEC25	L3	* 0.72	** 1.66	*** 0.11
WAEC26	L3	3.12	** 1.07	*** 0.23
WAEC27	L3	* 0.06	7.47	*** 0.08
WAEC28	L3	* 0.10	** 1.12	*** 0.08
WAEC29	L3	3.93	** 1.02	*** 0.15
WAEC30	L3	* 0.49	4.17	*** 0.17
WAEC31	L3	* 0.46	** 3.02	*** 0.13
WAEC32	L3	* 0.08	** 0.45	*** 0.07
WAEC33	L3	* 0.10	** 1.90	**** 0.14
WAEC34	L3	3.69	** 0.88	*** 0.12
WAEC35	L3	* 0.11	** 0.63	0.41
WAEC36	L3	* 0.28	4.96	*** 0.14
WAEC37	L3	* 0.07	** 1.84	*** 0.06
WAEC38	L3	* 0.16	** 2.86	*** 0.09
WAEC39	L3	* 0.81	** 1.46	*** 0.12
WAEC40	L3	6.62	** 1.14	*** 0.30
WAEC41	L3	5.65	** 1.05	*** 0.21
WAEC42	L3	* 0.09	10.42	*** 0.06
WAEC43	L3	4.13	** 1.09	*** 0.14
WAEC44	L3	* 0.13	** 0.20	*** 0.10
WAEC45	L3	* 0.59	** 1.62	*** 0.14
WAEC46	L3	* 0.55	** 2.73	*** 0.11
WAEC47	L3	* 0.33	** 2.54	0.40
WAEC48	L3	* 0.43	** 2.06	*** 0.20

WAEC49	L3	* 0.22	** 1.57	*** 0.11
WAEC50	L3	2.88	** 2.31	0.37

NB: * - discriminating *"a"*, ** - difficulty *"b"* and guessing – *"c"*

Fig 2: NECO CURVE

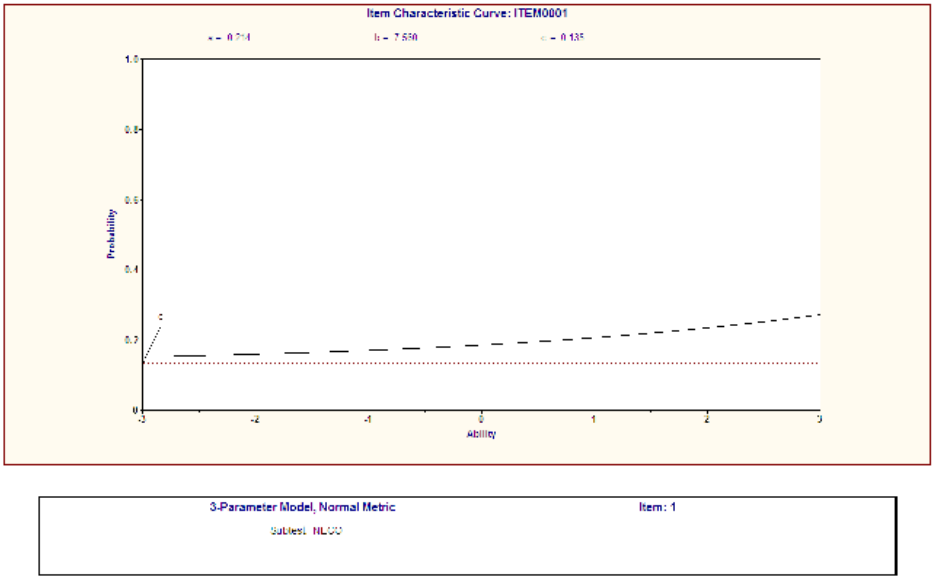


Table 2: Item Parameter Estimates for 60- NECO Items.

ITEM	*MODEL*	*a*	*b*	*C*
NECO1	L3	* 0.21	0.56	*** 0.14
NECO2	L3	* 0.13	0.61	*** 0.07
NECO3	L3	3.51	** 1.75	*** 0.19
NECO4	L3	* 1.66	** 0.98	*** 0.31
NECO5	L3	3.04	** 1.33	*** 0.16
NECO6	L3	* 0.14	** 2.70	*** 0.07
NECO7	L3	* 0.27	** 2.64	*** 0.10
NECO8	L3	* 0.35	2.08	*** 0.08
NECO9	L3	* 0.14	** 0.85	*** 0.08
NECO10	L3	* 0.17	7.99	*** 0.12
NECO11	L3	* 0.22	4.04	*** 0.13
NECO12	L3	* 0.37	** 1.26	*** 0.12
NECO13	L3	* 1.07	** 0.88	*** 0.00
NECO14	L3	* 0.30	2.15	*** 0.09

NECO15	L3	* 1.12	** 1.33	*** 0.10
NECO16	L3	* 0.44	-0.01	*** 0.19
NECO17	L3	* 0.08	** 1.84	*** 0.27
NECO18	L3	* 0.14	** 0.94	*** 0.05
NECO19	L3	* 0.27	6.61	*** 0.10
NECO20	L3	* 0.20	** 2.93	*** 0.15
NECO21	L3	2.82	** 1.34	*** 0.17
NECO22	L3	* 0.33	6.63	*** 0.29
NECO23	L3	* 0.35	7.09	*** 0.07
NECO24	L3	* 0.23	2.19	*** 0.16
NECO25	L3	3.35	** 1.40	*** 0.23
NECO26	L3	3.50	** 1.42	*** 0.23
NECO27	L3	* 0.27	2.69	*** 0.10
NECO28	L3	* 0.16	2.85	*** 0.12
NECO29	L3	* 0.37	4.93	*** 0.09
NECO30	L3	* 0.29	8.17	*** 0.12
NECO31	L3	2.60	** 1.08	*** 0.23
NECO32	L3	* 0.18	2.56	*** 0.13
NECO33	L3	* 0.22	10.44	*** 0.13
NECO34	L3	* 1.77	** 1.60	*** 0.08
NECO35	L3	* 0.35	2.79	*** 0.09
NECO36	L3	* 0.33	2.83	*** 0.13
NECO37	L3	* 0.29	2.51	*** 0.10
NECO38	L3	* 0.23	** 1.34	*** 0.11
NECO39	L3	* 0.18	** 1.59	*** 0.12
NECO40	L3	* 0.23	2.93	*** 0.13
NECO41	L3	* 0.16	** 0.26	*** 0.11
NECO42	L3	* 0.21	** 2.48	*** 0.12
NECO43	L3	* 0.14	** 2.89	*** 0.23
NECO44	L3	* 0.32	** 2.20	*** 0.08
NECO45	L3	* 0.28	** 2.53	*** 0.11
NECO46	L3	* 0.22	** 2.09	*** 0.11
NECO47	L3	2.79	** 1.08	*** 0.14
NECO48	L3	3.54	** 1.84	*** 0.11
NECO49	L3	* 0.31	2.26	*** 0.13
NECO50	L3	* 0.24	** 2.91	0.50

NECO51	L3	* 0.21	10.02	*** 0.12
NECO52	L3	* 0.11	** 2.50	***0.03
NECO53	L3	* 0.14	** 2.37	*** 0.07
NECO54	L3	* 0.12	2.45	*** 0.10
NECO55	L3	* 0.14	2.88	*** 0.15
NECO56	L3	* 0.13	2.82	*** 0.10
NECO57	L3	* 0.26	** 1.35	*** 0.11
NECO58	L3	* 0.58	** 1.17	*** 0.10
NECO59	L3	* 0.26	** 1.63	*** 0.12
NECO60	L3	* 0.31	6.42	*** 0.14

NB: * - discriminating *"a"*, ** - difficulty *"b"* and guessing – *"c"*

It could be observed from table 1 and 2 that the result of item calibration fitted into 3-PL model of item response theory gave parameter estimates, that is, difficulty 'b', discriminating 'a' and guessing 'c' for all the items of WAEC and NECO. It was observed that for the item discriminator 'a', the value for the WAEC examine ranges between 0.00 and 6.62 while for NECO examination ranges between 0.08 and 3.54. The typical range for discrimination is between 0 and 2. For the WAEC test items, 40 items were within the typical range while NECO had 52 items within the range. Items with discrimination below 0.65 are considered to have low discrimination, item between 0.65 and 1.34 are moderate, 1.35 and 1.69 are considered high while 1.70 and above are very high (Baker, 2001). Thus, 10 items in WAEC and 8 in NECO test were outside the typical range of 0 to 2.

The difficulty parameter *'b'* for WAEC examination items ranges between 0.09 and 10.42 which implied that item 3 is the easiest and item 29 is the most difficult. For NECO examination, the item difficulty ranges between -0.01 and 10.44 which implies that item 16 is the easiest and item 33 the most difficult. Item with lower 'b' value are considered to be easier than item with higher *'b'* value. However the typical range for difficulty is between -3 and +3. Many of the WAEC test items (41 items) falls within this range of difficulty while 9 items fall outside the range.

Also, only 33 items in the NECO test examination items falls within the range. The remaining 27 items have values greater than 3 can be regarded as bad items or need to be raised since their values indicate that they are extremely difficult. For the guessing parameter *'c'*, the WAEC examination ranges between 0.00 and 0.43 while NECO examination ranges between 0.03 and 0.50. Typical, the range of guessing is between 0 and 0.35 is considered acceptable and if otherwise unacceptable (Barker, 2001) WAEC test items (45 items) falls within this range of guessing while 5 items fall outsides the range. Also, only one item (item 1) in the NECO examination falls outside the range.

Discussion

The study's finding demonstrates that the two public testing bodies provided test items with comparable psychometric qualities and that neither one was better than the other. The results support those of research by Alfred (2013), Metibemu (2016), Kolawole (2007), Ogbebor (2017), that claim there is no discernible difference between the difficulty levels of WAEC and NECO multiple-choice questions in mathematics. The study does not agree with Adewale (2008), Bamidele and Adewale (2013), or Thomas et al. (2016) who claim that the WAEC test items are easier than those in the NECO examination.

Furthermore, the results of this study contradict those of Abiri (2006) and Olutola (2015), who claim that multiple choice tests with a small number of options—let's say, four—are more difficult than tests with a larger number of options, even though the item discriminations in the aforementioned studies indicated a significant difference. It might be argued that WAEC items contained certain measuring irregularities, such as the opportunity to correct selection, which gives examinees with high ability a distinct advantage. WAEC items also distinguish better between examinees with high and low ability.

The results of this study are consistent with past research by Olatunji (2007), Olutola (2015), Thomas et al., Uchegbue and Ugbe (2016); that is, WAEC test items are more discriminating than those used in NECO tests. However, according to other researchers, such as Alfred (2013), Metibemu (2016), and Ogbebor (2017), there was no statistically significant difference in the slope of test items made by Nigerian public examining bodies like WAEC and NECO, to name a few, in their investigations. This demonstrates that these examination bodies' test questions are similarly discriminant, whereas the variations in their asymptotic parameters were negligible.

Assuming that an examinee's propensity to guess on any given item is solely a quality of the item, which has exactly the same fixed effect on every examinee, is necessary for the items' reliability estimates. But as a researcher, I can attest that the degree to which examinees guess varies. Others test takers make guesses frequently, others occasionally, and some rarely. Since the same group of test takers tried both test items, it was anticipated that their levels of guessing would not vary significantly. According to this study's findings, which concur with those of Alfred (2013), Metibemu (2016), and Ogbebor (2017), there was no discernible variation in the guessing parameter of test items created by public examination organisations in Nigeria.

Conclusion

Based on the study's finding, it can be said that the bulk of the mathematics items created by the two testing organisations (WAEC and NECO) have similar psychometric properties. When WAEC and NECO items were submitted to the item calibration processes of the IRT framework, estimates of all item parameters were provided. As a result, the Mathematics questions created by the WAEC and NECO fit the 3-PL model well.

Recommendations

i. Examination bodies should commit themselves to the construction of test items by using IRT for calibration;
ii. There should be no discrimination against WAEC and NECO examinations since both examinations showed almost the same level of validity/reliability in the construction of their items especially in the area of difficulty indices.

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. DOI: 10.5897/ERR2019.3850

Abiri, O.O. (2006). Elements of Evaluation Measurement and Statistical Techniques in Education. Ilorin: Library and Publication Committee, University of Ilorin, Nigeria.

Adegoke, B. (2013). Primary Intention. *The Australian Journal of World Management.* 16(1), 1837-6304.

Adewale, A. E. (2008). Comparative analysis of the psychometric properties of WAEC, NECO and NABTEB Mathematics achievement examinations. Unpublished Ph.D Thesis, University of Ado-Ekiti, Nigeria.

Alfred, A.O. (2013). *Scoring Tests with Dichotomous and Polytomous items*. Maureen Gras Dean of the Graduate School, University of Georgia.

Amoke, M. K. (2020). Analysis of Students' Performance in Chemistry in the West African Senior School Certificate Examination (WASSCE) and National Examination Council (NECO) from 2015-2018. *International Journal of Research and Analytical Reviews*, 7 (1), 35-49.

Baker, F. B. (2001). *The Basic of Item Response Theory. Test Calibration*. ERIC Clearing House on Assessment and Evaluation. University of Maryland, College Park, MD. 136- 330.

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.

Bulus, C. H. & Chukwuma, N. M. (2021). Comparative Analysis of Students Performance in NECO and WAEC Economics Examination between 2014 and 2018 In Jos Metropolis, Plateau State. *Sapientia Foundation Journal of Education, Sciences and Gender Studies*, 3 (1), 169 – 181.

Fajobi, O. O. (2020). Comparability of Psychometric Properties of Mathematics items constructed by WAEC and NECO for Nigerian Secondary Schools using Item Response Theory. Unpublished Ph.D. Thesis, Institute of Education. Ekiti State University, Ado – Ekiti, Nigeria.

Kolawole, E.B. (2007). A Comparative analysis of psychometric properties of two Nigerian examining bodies for Senior Secondary Schools Mathematics. *International Research Pakistan Journal of Applied Sciences*, 2 (8); 913-915.

Metibemu, M. A. (2016). 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.

Ogbebor, U.C. (2017). Construct of Mock Economics Test for Senior Secondary School Students in Delta State, Nigeria using Classical Test and Item Response Theories. Unpublished Ph.D Thesis. Institute of Education. University of Ibadan.

Olatunji, D.S. (2007). Effects of Number of Options on Psychometric Properties of Multiple Choice Tests in Economics. Unpublished M.Ed Thesis, University of Ilorin, Ilorin.

Olawole, A.O. & Olugbemi, T.A**.** (2022). Organization and Administration of Guidance programme in schools and the implementation of Universal Basic Education.UBE-2022 Book of Readings.

Olawole, A.O. & Olotu, A. (2018). The Relationship between Learning Style and Time-Management on Academic Achievement of Undergraduate Students of University of Lagos. Journal of Counselling & Applied Psychology.

Olawole , A.O. & Oniemola , R.F. (2013). The relationship between procrastination and test anxiety among Distance Learning Institute (DLI) students, University of Lagos: Implication for Counselling. Journal of Business and Educational Policies (International Edition) 9(1) 97-106.

Olutola, A. T. (2015). Item difficulty and discrimination indices of multiple choice Biology tests. *Liceo Journal of Higher Education Research*, 11 (1), 16-31.

Oshkosh, U. W. (2002). *Item Analysis, Testing Services Support*. http://www.uwosh.edu/testing/facultyinfo/itemanalysis.php.

Osiesi, M.P. & Fajobi, O.O (2019). Fostering Mathematics Achievement: The impact of age, gender and self-esteem using the Rosenberg Self-Esteem Scale on Secondary School Students in South West, Nigeria. *West African Journal of Education*, 39, 90-100.

Osiesi, M.P. & Fajobi, O.O. (2019). Mathematics Teachers' Quality: A Sine Qua Non For Enhanced Primary Education in Oyo State, Nigeria. *IOSR Journal of Research and Method in Education*, 9 (6), series II, 1-6.

Osiesi, M.P. & Odinko, M.N. (2019). Influence of Teacher, Pupil Factors and their Knowledge of Statistics on Pupils' Achievement in Mathematics in Osun State, Nigeria. *Journal of Asian and African Social Science and Humanities*, 5 (4), 11-25.

Osiesi, M.P. (2019). Construction, Administration and the Standardisation of Mathematics Multiple Choice Test

Items For Junior Secondary School Two (2) in Oyo State, Nigeria. *British Journal of Education*, 7 (11), 24-40.

Osiesi, M.P. (2019). Evaluation of Teacher and Pupil factors on Pupils Attitude to Mathematics in Osun Staten Nigeria. *FUOYE Journal of Education*, 2 (2), 142-151.

Thomas, E.S., Uchegbue, H.& Ugbe, E. B. (2016). Comparative item analysis of students' WAEC And NECO Mathematics Objectives Test Item Scores. Educational Assessment and Research Network in Africa, volume 3.