

Analysis of Linear Equating of Senior School Certificate Multiple-Choice Examination Papers in Economics

Dorcas Sola Daramola¹, Adekunle Thomas Olutola^{1*}, Mohammed Idris Jimoh²

¹ Department of Social Sciences Education, Faculty of Education, University of Ilorin, Nigeria.

² Department of Educational Foundations, Faculty of Education, Federal University Dutsin-Ma, Katsina State, Nigeria.

*Corresponding Author: aolutola@gmail.com olutolatola@gmail.com

ARTICLE INFO

ABSTRACT

Article history

Received : July, 2021.

Revised : June, 2022.

Accepted : July, 2022.

Keywords

Linear equating.

Economics.

Multiple-Choice Papers.

Unique Items.

Anchor Items.

Background: Test equating methods are statistical tools used to produce exchangeable scores across different test forms. However, people believe that it is impossible to develop multiple forms of tests that have the same psychometric properties.

Contribution: The study therefore analyzed the linear equating of senior school certificate multiple-choice examination papers in Economics.

Method: A Non-equivalent Groups Anchor Test (NEAT) design was adopted for the study. The sample for the study consisted of 1,119 senior secondary three students. The 2009 WAEC, NECO, and NABTEB SSCE multiple-choice Economics papers were adapted and used for data collection with 0.67, 0.64, and 0.60 coefficients of content validity as well as 0.79, 0.76, and 0.70 reliability coefficients. Data collected were analyzed using descriptive statistics, T-scores, and Analysis of Variance.

Results: Findings revealed that there was no difference in examinees' ability, thus equating process could take place. The finding also revealed an F-value of 1.953, which was not significant at the 0.05 alpha level since the p-value (0.142) was greater than 0.05. It showed a score of 41 in WAEC and NECO was equivalent to 42 in NABTEB. It also showed a score of 45 in WAEC was equivalent to 46 and 50 in NECO and NABTEB respectively.

Conclusion: The researchers therefore concluded that the 2009 WAEC and NECO SSCE Economics multiple-choice items tended to be equivalent while that of NABTEB was different. It was recommended among others that a regulatory body could be established to monitor and standardize these examinations conducted in Nigeria.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.
Copyright © 2022 Dorcas Sola Daramola, Adekunle Thomas Olutola,
Mohammed Idris Jimoh.



INTRODUCTION

In the teaching-learning process, learners are faced with a series of instructions. There is the need to constantly gauge the extent to which they respond to instructions, which could be termed students' performance [1]. To ascertain this, there is the need to have a device in doing it. This, therefore, brings about the use of the test. A test is an assessment instrument, tool, technique or method used systematically to measure a sample of behavior by posing a set of questions or to complete certain tasks for the students to react to gauge mastery of a skill or knowledge in the curriculum content [2] [1]. Tests are designed to measure the quality, ability, skill or knowledge of a sample against a given standard [1]. Tests can take the form of standardized and non-standardized.

Standardized tests are tests administered to millions of candidates each year at once. It uses controlled standard procedures for administration and scoring. It is used to evaluate students' learned skills in academic subjects, teachers' and schools' performance, quality of the curriculum, and the educational system. Most of the standardized tests are external to the school environment and have predetermined conditions such as attempting similar questions of the same difficulty level, having the same supervised test conditions, and scoring in the same way. The interpretations are consistent across schools and years [3] [4] and used for making comparisons about students' achievement and promoting accountability of education. The primary purpose of standardized tests is to achieve fairness of assessment of learning outcomes, upon which important decisions are based.

In any standardized test, especially where two or more examining bodies perform similar functions (e.g., the West African Examination Council, the National Examinations Council, and the National Business and Technical Examinations Board (NABTEB) conducting Senior School Certificate Examination), test equating are highly essential. Equating is a process that converts scores on one form of a test to the score scale of another form. It is a family of psychometric approaches to align the numerical format of different measurement instruments [5]. It is a linkage between two or more scores generated from tests. Equating may be viewed as a form of scale aligning in which very strong requirements are placed on the tests being linked. The goal of equating is to produce a linkage between scores on two test forms so that the scores from each test form can be used as if they have come from the same test [6]. It is a technical procedure or process conducted to establish comparable scores on different versions of a test, allowing them to be used interchangeably [7]. It is an important aspect of establishing and maintaining the technical quality of a testing program by directly impacting the validity of assessments [7]. When two tests or items have been successfully equated, educators can validly interpret performance on one test form as having the same substantive meaning compared to the equated score of the other test form [5] [7].

Test equating methods are statistical tools used to produce exchangeable scores across different test forms [7]. The two recognized measurement theories, Classical Test Theory (CTT) and Item Response Theory (IRT), can be used to compute test equating. In CTT, mean scores and standard deviation are used to equate the performance in two forms. The Tucker and Levine Observed Score are the two recognized equating methods. The Tucker estimates the relationship between observed scores on two forms of test scores. Observed-score equating method refers to the transformation of the raw scores of a new test, "X", into the raw scores of an old test, "Y". It is used to ensure that test scores from different test forms are comparable and that the scores can be used interchangeably. The Levine True Score method estimates the relationship between true scores on the two forms. Usually, in any of the above methods,

equipercentile equating comes into play to determine where the relationship could have an equivalent percentile on either form.

In IRT, equating means the process of placing scores from two parallel test forms onto a common score scale. The scores from the two different forms can be compared directly or treated as if they come from the same test form. Horizontal and vertical test equating are recognized. Vertical equating refers to the process of equating tests administered to groups of test-takers with different abilities, in different years of schooling (e.g., Senior Secondary students 1 and Senior Secondary students 2) at a time. Horizontal equating refers to the equating of tests administered to test-takers of different groups (Male and female in Senior Secondary students 1) with similar abilities. However, different tests are used to avoid practice effects. Test equating establishes validity and reliability of the instrument across forms and years, fairness of items, test security, and continuity of the program.

Linear equating is implemented by reflecting the ability level of the students and the spread of scores onto the reference scale scores. It provides a transformation so that scores from two tests will be considered equated if they correspond to equal standard score deviates. It is useful with small samples, and the accuracy of the results is most important near the mean [5]. Five requirements are widely viewed as necessary for a linking to be an equating [8]. Those requirements are:

1. The Equal Construct Requirement: The two tests should both be measured of the same construct (e.g., latent trait, skill, ability).
2. The Equal Reliability Requirement: The two tests should have the same level of reliability.
3. The Symmetry Requirement: The equating transformation for mapping the scores of "Y" to those of "X" should be the opposite of the equating transformation for mapping the scores of "X" to those of "Y".
4. The Equity Requirement: It should be a matter of indifference to an examinee as to which of two tests the examinee takes.
5. The Population Invariance Requirement: The equating function used to link the scores of X and Y should be the same regardless of the choice of (sub) population from which it is derived.

Concerning the best practices, Requirements 1 and 2 mean that the tests need to be built to the same specifications, while Requirement 3 precludes regression methods from being a form of test equating. [9] argues that Requirement 4 implies both Requirements 1 and 2. Requirement 4 is, however, hard to evaluate empirically and its use is primarily theoretical (Hanson, 1991; Lord, 1980). As noted by [8], Requirement 5, which is easy to assess in practice, also can be used to explain why Requirements 1 and 2 are needed. If two tests measure different things or are not equally reliable, then the standard linking methods will not produce results that remain unchanged when a particular transformation is applied to it across certain subpopulations of examinees.

Proposed [10] for the NEAT design, the anchor test has a central role as a proxy of ability because the conditional mean and variances over anchor scores are used to obtain a family of equating transformations. Linear equating is implemented by reflecting the ability level of the students and the spread of scores onto the reference scale scores [11]. To equate scores on the new form to scores on the reference form in a group of test-takers, each score on the new form is to be transformed into the score on the reference form that has the same number of standard deviations above or below the mean of the group. It is called Linear equating because the relationship between the raw scores and the adjusted scores appears on a straight line [12].

There have been studies on linear equating both within and outside Nigeria. For instance is, [13] compared mean equating, linear equating, and equipercentile equating using various degrees of pre-smoothing (including none at all) in samples ranging in size from 25 to 200, using data from only one test. Mean equating is the most accurate of the small sample methods for below-average scores, but the least accurate for above-average scores. Linear equating is more accurate than equipercentile equating for below-average and near average scores, but less accurate for scores more than one standard deviation above the mean. In all test equating studies, according to IRT, alternate forms should be balanced in terms of equivalent test information functions (TIF). To be specific, an examinee who takes Form A should not be more or less advantaged than one who takes Form B or Form C [14]. [6] carries out a study that compares linear equating and Rasch equating. It is conducted from the study that Rasch equating provides essentially the same results as linear equating.

The motivating factor for carrying out this study is that people believe it is impossible to develop multiple forms of tests that have the same psychometric properties. Some stakeholders seem to have the impression that anyone with NECO or NABTEB SSCE results is half-baked and to them, only WASCE is qualitative enough for acceptance. In Nigeria, some studies have indicated doubt about the quality of the Senior Secondary School Certificate Examinations. [15] points out that there are vast differences in the quality of certificate examinations conducted by the various examination bodies. [16, 17, 18, 19, & 20] remarks that the standard of SSCE conducted by NECO is low compared to SSCE conducted by WAEC.

Moreover, some tertiary institutions and employers of labour tend to prefer students with credit passes in the SSCE conducted by WAEC to those conducted by NECO and NABTEB. They believed that the SSCE conducted by WAEC has a higher standard than the SSCE conducted by NECO [21]. Thus, they conduct a study on the conversion of units of WAEC, NECO, and NABTEB for common comparison to establish whether this assertion is true. Therefore, this study statistically analyzes linear equating of WAEC, NECO, and NABTEB Senior School Certificate Multiple-choice Test Items in Economics.

The main objective of this study is to analyze the linear equating of WAEC, NECO, and NABTEB Senior School Certificate Multiple-choice Test Items in Economics. Specifically, the study examined the:

1. form of students' performance on the common items of WAEC, NECO, and NABTEB Senior School Certificate multiple-choice test items in Economics;
2. form of students' performance on the unique items of WAEC, NECO, and NABTEB Senior School Certificate multiple-choice test items in Economics;
3. results of linear equating of WAEC, NECO, and NABTEB Senior School Certificate Economics multiple choice papers with the use of standard score deviates.

Three research questions were raised and answered in the study.

1. What is the form of students' performance on the common items of WAEC, NECO, and NABTEB Senior School Certificate multiple-choice test items in Economics?
2. What is the form of students' performance on the unique items of WAEC, NECO, and NABTEB Senior School Certificate multiple-choice test items in Economics?
3. What are the results of linear equating of WAEC, NECO, and NABTEB Senior School Certificate Economics multiple-choice papers with the use of standard score deviates?

METHOD

The researchers adopt a Non-equivalent Groups Anchor Test (NEAT) design of test score equating for this study because the study is designed to equate WAEC, NECO, and NABTEB

Senior School Certificate multiple-choice test items scores in Economics. Eleven thousand seven hundred and seven Senior Secondary School 1, Senior Secondary School 2 and Senior Secondary School 3 students of Economics in public schools in Kwara State constitute the population for the study. Kwara State is one of the states in Nigeria, located in the North central, Nigeria and its capital is Ilorin. The state has three senatorial districts (Kwara Central, Kwara North and Kwara South). The target population consisted of 3, 359 Senior Secondary School three (SS3) students of Economics.

Multi-stage sampling technique is adopted for the study. The first sampling procedure adopted is stratification of schools to the three Senatorial Districts found in Kwara State. The second sampling procedure adopted is proportionate sampling. Ten per cent (10%) of the total public senior secondary school in each Senatorial District is proportionally selected. Hence, eight (8), seven (7), and fifteen (15) public Senior Secondary Schools are selected in Kwara North, Kwara Central, and Kwara South respectively. This amounts to thirty (30) public schools selected. A purposive sampling technique is also used to select Senior Secondary Schools 3. This is because they are in the final stage of their secondary school programme and are ready to write their external and final examinations. The researchers hope that they ought to have cover much ground of the syllabus. In the Kwara central is 478 (42.7%); Kwara North is 264 (23.6%); while Kwara South has 377 (33.7%) respondents. A total of one thousand hundred and nineteen (1,119) students participate in the study.

The researchers adopt 2009 WAEC (Form A), NECO (Form B), and NABTEB (Form C) Economics multiple-choice items for data collection composed into independent and anchor tests. Each test form has unique items shared, a set of twenty common items located at numbers 11-30 in each test form. Test forms A, B, and C contain 30, 40, and 30 WAEC, NECO, and NABTEB multiple-choice items (unique items) respectively and each form also contains 20 multiple-choice items (common items).

The researchers determine the instruments' (Forms A, B and C) content validity with a formula $r_{ts} = 1 - \frac{\sum |d|}{100}$ (r_{ts} is the correlation between test and syllabus contents) where d is the difference (ignoring signs) between corresponding percentage weightings of each of the WAEC, NECO and NABTEB economics topics and the contents of the test. Also, measures of internal consistency with split-half method of estimating reliability is employed. The following Coefficients of content validity 0.67, 0.64, and 0.60 and reliability coefficients of 0.79, 0.76, and 0.70 are respectively obtained for forms A, B, and C.

This shows that the instruments are good and reliable. The researchers use these Economics multiple-choice papers because they are interested in equating the Nigerian Senior Secondary School Certificate Examination (SSCE) Economics multiple-choice papers of the three different examination bodies. Research questions one and two generated are answered using means and standard deviation while research questions three is answered using percentile rank.

RESULTS AND DISCUSSION

Answers to Research Questions

Research Question One: What is the form of students' performance on the common items of WAEC, NECO, and NABTEB Senior School Certificate multiple-choice test items in Economics?

Students' scores in common items of test form A, B, and C are independently summed and their mean and standard deviation are computed.

Table 1. Mean of respondents' pattern of performance in common items

Test Form	Mean	Maximum	Minimum	Standard Deviation
Form A	10.98	17	5	2.44
Form B	11.30	17	5	2.38
Form C	11.28	17	5	2.38

Table 1 reveals 10.98 as the mean performance of test-takers in common items in test form A, while the mean performance of test-takers in common items in test form B and C are 11.30 and 11.28 respectively. Also, 17 and 5 are shown as the test-takers' highest and least scores in common items in the three test forms while standard deviations of 2.44, 2.38, and 2.38 are obtained for forms A, B, and C respectively. This implies that there is no difference in examinees' ability or proficiency in the subject, thus scores equating among WAEC, NECO, and NABTEB multiple-choice items in Economics is upheld. The finding also implies that differences in difficulty of items in the tests are equal since they have the same distribution of common items scores in the three forms.

To ensure that condition for equating using Non-equivalent method is not violated, respondents' scores in common items are subjected to One-way ANOVA to know if there is any significant difference in the test-takers' performance at 0.05 level of significant as demonstrated in table 2.

Table 2. ANOVA summary table of respondents' performance in SSCE common items

Source	Sum of Square	df	Mean Square	F	P-value	Decision
Between Groups	22.812	2	11.406			
Within Groups	6518.348	1,116	5.841	1.953	0.142	NS
Total	6541.160	1,118				

Table 2 reveals an F-value of 1.953, which is not significant at the 0.05 alpha level, since the p-value (0.142) is greater than 0.05 ($0.142 > 0.05$). This therefore, implies that test-takers are not statistically different in proficiency/ability in Economics multiple-choice test items conducted by the WAEC, NECO, and NABTEB in 2009 Senior School Certificate Examination. This is a further confirmation of this test equating process.

This finding is in line with the submission of [6] that test-takers that take different test forms are the same in proficiency, if the test-takers have the same distribution of common items scores on different test forms taken.

Research Question Two: What is the form of students' performance on the unique items of WAEC, NECO, and NABTEB Senior School Certificate multiple-choice test items in Economics?

Students' scores in unique items of test form A, B, and C are independently converted to a percentage, summed, and are subjected to mean, standard deviation, maximum, minimum, skewness, and kurtosis

Table 3. Descriptive statistics of the pattern of students' performance on unique items

Test Form	Mean	Std.Dev.	Maximum	Minimum	Skewness	Kurtosis
Form A	35.45	5.23	47	24	0.136	-0.822
Form B	36.40	6.44	54	15	0.203	-0.220
Form C	41.54	8.71	56	5	-0.183	4.295

Table 3 shows the performance of the students on the unique items. Respondents in form C with a mean performance of 41.54, skewness of 0-1.83, and kurtosis of 4.295 have the best performance because the skewness value of -1.83 indicates that the mass of the scores is clustered to the right at the high values. Kurtosis value of 4.295 indicates a sharp peak of which the distribution concentrates on the right values. Test-takers in form B have better performance with a mean of 36.40, skewness of 0.203, and kurtosis of -0.220 while the least mean performance of 35.45 is recorded by the respondents in form A with skewness of .136 and kurtosis of -0.822. Positive skewnesses indicate that mass of the scores clusters to the left at the low values. This implies that there is a difference in the pattern of students' performance in unique items across the test forms. Higher mean performance and negative skewness value of students in test form C indicates that it is relatively easy compared to test form A & B.

Research Question Three: What are the results of linear equating of WAEC, NECO, and NABTEB Senior School Certificate Economics multiple choice papers with the use of standard score deviates?

Students' scores in the test form A, B, and C ere independently converted to a percentage and standardized using T-score. The T-scores is used as a scale on which examinees' raw score in each of the test form is compared.

Table 4. Linear Equating of WAEC, NECO, and NABTEB SSCE multiple choice papers

T-Score	Student's raw score in WAEC	Student's raw score in NECO	Student's raw score in NABTEB
-3	-	-	5
1	-	-	-
2	-	-	-
3	-	-	-
4	-	-	-
5	-	-	-
6	-	-	-
7	-	-	-
8	-	-	-
9	-	-	-
10	-	-	-
11	-	-	-
12	-	-	19
13	-	-	20
14	-	-	-
15	-	-	-
16	-	-	-
17	-	29	-
18	-	-	-
19	-	-	-
20	-	-	-

21	-	-	-
22	-	-	-
23	-	-	29
24	-	-	-
25	-	-	-
26	-	-	32
27	-	-	-
28	-	-	-
29	-	-	-
30	-	-	-
31	-	-	-
32	-	-	37
33	-	-	38
34	39	39	39
35	-	-	-
36	40	40	40
37	-	-	-
38	41	41	42
39	-	-	-
40	42	42	-
41	-	-	45
42	-	-	46
43	-	44	47
44	44	-	-
45	-	45	48
46	-	-	49
47	45	46	50
48	-	47	51
49	46	-	52
50	-	48	-
51	47	-	54
52	-	49	-
53	48	-	55
54	-	50	56
55	49	51	57
56	-	-	58
57	50	52	60
58	-	-	61
59	-	53	62
60	-	-	-
61	52	54	-
62	-	55	-
63	-	-	-
64	53	56	65
65	-	-	-
66	-	57	-
67	-	-	-
68	-	58	-
69	-	59	-
70	56	-	-
71	-	60	-
72	-	-	-
73	-	-	-
74	-	-	-

75	-	62	-
76	-	63	-

Deviation in respondents' scores in WAEC, NECO, and NABTEB is noticed in Table 4 at T-scores of 38 and 47, the scores are 41, 41, 42 and 45, 46, and 50 respectively. It is also found that a Score of 45 in WAEC (Form A) is equivalent to 46 and 50 in NECO and NABTEB because they correspond to the same standard score deviate (47). This result can be attributed to a long year of WAEC operation. Thus, its item writers would have gotten better knowledge via experience in item construction. Linear equating also reveals that a score of 53 in WAEC is equivalent to 56 and 65 in NECO and NABTEB respectively because they correspond to the same standard score deviate (64).

This finding corroborates the submission of Alfred (2011) that there is a significant difference in the difficulty level of Economics multiple-choice items conducted by WAEC, NECO, and NABTEB with a mean of 37.30, 35.48, and 30.66 respectively. In the same vein, the finding supports that of [22] that there is a difference in difficulty index of Mathematics examination conducted by WAEC, NECO, NABTEB and JAMB (Joint Admmision Matriculation Board) in Nigeria. The finding is contrary to [23] submission that there is no significant difference in the difficulty level, reliability, and validity coefficients of mathematics items constructed by WAEC, NECO, and NABTEB. The finding does not also substantiate the submission of [24] that WAEC and NECO have the same difficulty indices.

To establish the outcome of Linear Equating the summary of equated scores of WAEC, NECO and NABTEB are presented in Table 5

Table 5. Summary of comparison of Linear Equating of SSCE multiple choice papers

WAEC (FORM A)	NECO (FORM B)	NABTEB (FORM C)
39	39	39
40	40	40
41	41	42
45	46	50
49	51	57
50	52	60
53	56	65

It is shown in table 5 that a score of 39 and 40 in WAEC are also equivalent to 39 and 40 both in NECO and NABTEB, while, the score of 41 in WAEC and NECO is equivalent to 42 in NABTEB. A score of 45 in WAEC is equivalent to 46 and 50 in NECO and NABTEB respectively, and a score of 49 in WAEC is equivalent to 51 and 57 in NECO and NABTEB respectively. A score of 50 in WAEC is equivalent to 52 and 60 in NECO and NABTEB respectively and a score of 53 in WAEC is equivalent to 56 and 65 in NECO and NABTEB respectively.

Findings on Linear equating of WAEC, NECO, and NABTEB Senior School Certificate Economics multiple choice papers with the use of standard score deviates reveal that a Score of 39 in WAEC (Form A) is also equivalent to 39 in NECO (Form B) and NABTEB (Form C) respectively because they correspond to the same standard score deviate (34) in table 4. It is shown on Table 5 that there are closely equivalent scores in WAEC and NECO compared to NABTEB. This can be attributed to a relatively long period of operation of these two Senior School Certificate Examination bodies. This finding supports that of [25,26] who says that there are no significant differences in the difficulty level of WAEC and NECO multiple-choice items in mathematics. The finding is not consistent with [27] submission that NECO is inferior to WAEC in all standards. It also disagrees with the submission of [18] that WAEC SSCE multiple choice

Biology items have more difficult items than NECO SSCE multiple choice Biology items. Finding on difference in NABTEB scores equivalence to WAEC and NECO recorded in this study supports the submission of [22] that there is a difference in difficulty index of Mathematics examination conducted by WAEC, NECO, NABTEB and JAMB (Joint Admission Matriculation Board) in Nigeria.

CONCLUSION

The examinees are not different in proficiency because mean performance of respondents in common items in test form A, B and C is not significantly different. Respondents perform differently in unique items across the test forms. Score of 39 in WAEC is also equivalent to 39 both in NECO and NABTEB, while, score of 41 in WAEC and NECO is equivalent to 42 in NABTEB because they correspond to the same standard score deviate (34). A score of 45 in WAEC is equivalent to 46 and 50 in NECO and NABTEB respectively because they correspond to the same standard score deviate (47), and a score of 53 in WAEC is equivalent to 56 and 65 in NECO and NABTEB respectively because they correspond to the same standard score deviate (64). The researchers therefore conclude that the 2009 WAEC and NECO SSCE Economics multiple-choice items tend to be equivalent while that of NABTEB is different. Equating method with a lower coefficient of variation (Linear equating) should be employed for equating scores. A regulatory body to standardize and monitor examinations conducted by the examining bodies should be established in Nigeria. Private owned examination bodies should be allowed to come on board; the proliferation would lead to a healthy competition which could result in the achievement of standards across the examination bodies.

REFERENCES

- [1] Adom, D., Mensah, J. A. & Dake, D. A. (2020). Test, measurement, and evaluation: Understanding and use of the concepts in education. *International Journal of Evaluation and Research in Education (IJERE)*, 9 (1), pp. 109~119. Available at <http://ijere.iaescore.com>
- [2] Linn, R. L., (2008). *Measurement and assessment in teaching*. India: Pearson Education.
- [3] Allison, M. (2011). *Student standardized testing: current practices in OECD countries and A literature review*. OECD Education Working Paper No. 65.
- [4] Fairman, J., Johnson, A., Mette, I., Wickerd, G. & LaBrie, S. (2018). *A Review of Standardized Testing Practices and Perceptions in Maine*. Published by the Maine Education Policy Research Institute in the Center for Education Policy.
- [5] Kolen, M. J., & Brennan, R. L. (2014). *Test Equating, Scaling, and Linking (3rd ed.)*. Springer.
- [6] Doran, N.J. Moses, T.P. & Eignor, D.R. (2010). *Principle and practices of score equating*. New Jersey: ETC Princeton.
- [7] Davier, A. A. V. (2008). New Results on the Linear Equating Methods for the Non-Equivalent-Groups Design. *Journal of Educational and Behavioral Statistics*, 33 (2), 186–203. Available at <http://jebs.aera.net>.
- [8] Holland, P. W. & Dorans, N. J. (2006). Linking and equating. In R. L. Brennan (Ed.), *Educational measurement (4th ed., pp. 187-220)*. Westport, CT: Praeger.
- [9] Lord, F.M. (1980). *Applications of item response theory to practical testing problems*. New Jersey: Hillsdale.
- [10] Wiberg, M., & van der Linden, W. J. (2011). Local linear observed-score equating. *Journal of Educational Measurement*, 48, 229-254.

- [11] Petersen, N. S., Cook, L. L., & Stocking, M. L. (1983). IRT versus conventional equating methods: A comparative study of scale stability. *Journal of Educational Statistics*, 8, 137-156.
- [12] Livingston, S. A. (2004). *Equating test scores (without IRT)*. Princeton, NJ: ETS.
- [13] Skaggs, G. (2005). The accuracy of random groups equates to very small samples. *Journal of Educational Measurement*, 42, 309-330.
- [14] Yu, C. H. & Sharon, E. O. P. (2005). *Test Equating by Common Items and Common Subjects: Concepts and Applications*. *Practical Assessment Research & Evaluation*, 10(4). Available online: <http://pareonline.net/getvn.asp?v=10&n=4>
- [15] Oni, J. A. (2001). *Resources and resources utilization as correlates of school academic performance in the secondary pre-vocational education in Ogun State, Nigeria*. Unpublished PhD Thesis, University of Ibadan.
- [16] Eze, N. (2002). School-Based assessment as a predictor of students' performance in Junior School Certificate Examination. *Nigeria Educational Review*, 8 (2), 32-44.
- [17] Olutola, A.T. (2011). *Analysis of item parameters of senior school certificate multiple choice in biology in Ekiti State, Nigeria*. Unpublished PhD Thesis, University of Ilorin, Ilorin.
- [18] Olutola, A.T. (2015). Item Difficulty and Discrimination Indices of Multiple Choice Biology Tests. *Liceo Journal of Higher Education Research*, 11(1), Pp.16-30.
- [19] Olatunji, D.S. (2007). *Effects of some options on psychometric properties of multiple-choice tests in economics*. Unpublished M.Ed Thesis, University of Ilorin, Ilorin.
- [20] Alfred, O. (2010). *Assessment of the Equivalence of 2009 Senior School Certificate Multiple Choice Economics Test Items in Ilorin, Nigeria*. M.Ed Thesis, University of Ilorin. Ilorin.
- [21] Okoro, D. U. (2004). *Quality in education with particular reference to learning achievement*. A paper presented at the National Conference on Education For All. Abuja 20th -23rd August.
- [22] Moyinoluwa, T. D. (2015). *Analysing the Psychometric Properties of Mathematics in Public Examinations in Nigeria*. *Research on Humanities and Social Sciences*. 5 (7), 24-30
- [23] Bandele, S. O., & Adewale, A. E. (2013). Comparative analysis of the item difficulty levels of WAEC, NECO, and NABTEB Mathematics achievement examinations. *Mediterranean Journal of Social Sciences*, 4(2), 761-765.
- [24] 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
- [25] Steve FN, Fadipe RA, Kayode OC. Causes and Consequences of Truancy Among in School Adolescents in Oyo North Senatorial District Nigeria. *SPEKTA (Jurnal Pengabdian Kepada Masyarakat: Teknologi dan Aplikasi)*. 2022 Jun 14;3(1):23-30.
- [26] Kolawole, E. B. (2007). A comparative analysis of psychometric properties of two Nigerian examining bodies for senior secondary schools mathematics. *Research Journal of Applied Sciences*, 2(8): 913-915.
- [27] Adeniran, T. (2000). *NECO certificate is valid for varsity admission*. An Article Published in the Punch Newspaper of Tuesday, 29th February 2000.

This page is intentionally left blank

SPEKTA (Journal of Community Service: Technology and Applications)