

*Analysis of Linear Equating System in ... Olutola, A. T., Et. al. (2025)*

## Analysis of Linear Equating in WAEC, NECO, and NABTEB Senior School Certificate Multiple-Choice Economics Examinations in North West, Nigeria

OLUTOLA, Adekunle Thomas<sup>1</sup>  
[olutolatola@gmail.com](mailto:olutolatola@gmail.com)

BAGUDU A. A.<sup>1</sup>  
ENETAIRO, Reuben<sup>2</sup>  
[reubenetairo@gmail.com](mailto:reubenetairo@gmail.com)

<sup>1</sup>Department of Educational Psychology & Counselling, Federal University Dutsin-Ma, Katsina State, Nigeria.

<sup>2</sup>Department of Educational Psychology, Guidance and Counselling, Federal College of Education (T) Bichi, Kano State.

### Abstract

*This paper investigated the analysis of linear equating method on WAEC, NECO, and NABTEB Senior School Certificate Multiple-Choice Economics Examinations in Northwest Nigeria. The “Non-Equivalent Anchor Test (NEAT) group Design” was adopted for this study. The population comprised 404,521 public senior secondary school students of 2023/2024 academic session and a sample of 1,532 SSII Economics students was drawn through multi-stage sampling procedure from the population. The instruments used for data collection were the adapted 2023 WAEC, NECO, and NABTEB SSCE multiple-choice Economics papers with reliability coefficient of 0.75, 0.89, and 0.88 respectively obtained by the researcher using split half method. The research questions were answered using descriptive statistics of mean and standard T-scores. Major findings revealed that (1) there is significant differences in difficulty between the pair of form A-C ( $t=4.473=.000, p < 0.05$ ) and B-C ( $t=4.375=.000, p < 0.05$ ), (2) there was no significant difference in the respondents performance on common items across the three examinations ( $\bar{X}_{WA} = 2.121; \bar{X}_{NE} = 2.114; \bar{X}_{NA} = 2.504$ ), (3) while performance on unique items varied across the test forms ( $\bar{X}_{WA} = 7.30 < \bar{X}_{NE} = 7.40 < \bar{X}_{NA} = 8.34$ ), with form C exhibiting the highest mean performance and form A the least performance (4) linear equating method revealed a positive linear relationship between raw scores and T-scores. The researchers concluded that form C (NABTEB) was distinctly easier, indicating significant differences from other forms (WAEC & NECO) that shown no statistical difference in students' ability or proficiency, that both test items are comparable. It was therefore recommended among others linear equating method for equating scores across these WAEC, NECO, and NABTEB Economics multiple-choice test should be considered.*

**Key Words:** Economics, Multiple-choice Items, Linear Equating, WAEC, NECO, NABTEB

### Introduction

Notably in the Nigerian educational system, different questions are used for Senior School Certificate Examination (SSCE) conducted by various credentialing awarding agency in a standardized examinations and there is no assurance that scores on these test forms are equivalent across the different examination bodies. Hence, different sets of test items might be easier or harder, which produces higher or lower scores across different test form taken by the candidates. Therefore, it is important to examine the quality of a test items over time when a standardized achievement test is administered consecutively with different test forms over several years (Wiberg & Von Davier, 2017). Tests are designed to measure the quality, ability, skill or knowledge of a sample against a given standard (Adom, Mensah, & Dake, 2020). Teachers' carry out a routine evaluation of school learning to achieve various objectives, but this is essentially internal. These internal evaluations go by such names as teacher-made tests, Continuous Assessment, School Based Assessment and local tests. For the conduct of external examinations, However, there are recognized bodies that carry this assignment out for the whole country (Nigeria) and award certificates to candidates at different levels. The National Examinations Council (NECO) and the West African Examinations Council (WAEC) and National Business and Technical Examination Board (NABTEB), are the bodies authorized by the Nigerian law to conduct the Senior School Certificate Examinations (SSCE), General Certificate Examinations (GCE) and other exams (Olutola,2015).

The West African Examinations Council (WAEC), the National Examinations Council (NECO), and the National Business and Technical Examinations Board (NABTEB) are the three main examination bodies that do conduct public examination in Nigeria and they play pivotal roles in administering Senior School Certificate Examinations to candidates. These examination bodies can be referred to as different test forms in test equating. A statistical process called equating can



*Analysis of Linear Equating System in ... Olutola, A. T., Et. al. (2025)*

therefore, be used to adjust for differences in difficulty across alternate forms resulting in comparable score scales and more accurate estimates of ability (Albano, 2011). Ensuring the comparability of scores across these examinations is crucial for fairness and accurate assessment since their syllabi are prepared from the approved curriculum produced by the National Educational Research and Development Council (NERDC) for Senior Secondary Schools, this guides the teaching and learning processes in the classroom. Equating process is used to obtain comparable scores when more than one test forms are used in test administration and the goal of score equating is to allow the scores from both tests to be used interchangeably (Daramola, Owolabi, & Olutola, 2018). Equating various test form is necessary for any testing programme that continually generates new editions of test form and which the interpretation of test scores from these editions having same meaning overtime. Since testing programs nearly always need examinations that measure the same thing, but are composed of different questions (i.e, alternate forms of the same test). Equating process is therefore used to obtain comparable scores when more than one test forms are used in test administration, Daramola et al (2022). Test equating are highly essential because equating ensured a process that can 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 (Kolen & Brennan, 2014). It is a linkage between two or more scores generated from different form of tests. Test equating as defined by Agah (2013), is aimed at putting the scores obtained by students from different forms of a tests on a common scale.

Meanwhile, there are two recognized theories that can be used in equating different forms of test and these include Item Response Theory (IRT) and Classical Test Theory (CTT). The Classical Test Theory (CTT) and Item Response Theory (IRT) are two popular statistical frameworks in educational measurement through which tests can be developed, validated, and finally used for assessing examinees performance. The Classical Test Theory (CTT) focuses on test level, while item response theory or IRT on the other hand, focuses on item level (Kolen & Brennan, 2014), it attempts to model the ability of a test taker and the probability of answering an item correctly based on the pattern of responses to all the items that constitute the test. Its ability parameter estimates are not test dependent and item parameter estimates are not group dependent. It overcomes the weaknesses of CTT with its ability to provide invariant item parameters. In Uysal and Kilmem (2016), opinion test developers attempt to prepare tests as similar as possible in terms of knowledge, skills and content, though, it is impossible to prevent changes in test difficulty and the differentiation of content and statistical qualities from one test format to another. But the variations in difficulty level and other statistical characteristics of different tests scores can be measured and controlled using test equating. However, literatures reviewed that the linear methods of equating is CTT equating approaches. In the context of educational testing, "linear equating" simply means a statistical method used to adjust scores from different versions of a test so that they are comparable. This is crucial when you have multiple test forms designed to measure the same construct, but which may vary slightly in difficulty. Its goal is to find a linear transformation that makes the scores from different forms statistically equivalent. According to Holland and Dorans (2006) as cited by Daramola et al., (2022) stated five requirements that are widely viewed as necessary for a linking to be an equating, these 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.

Therefore, this study investigated the effectiveness of linear methods of equating test scores in WAEC, NECO, and NABTEB Economics Multiple Choice Test Items. This is essential for fair admissions processes, scholarship awards, and other high-stakes decisions based on equivalent test scores obtained by candidates and ultimately enhancing the validity and reliability of test results.



*Analysis of Linear Equating System in ... Olutola, A. T., Et. al. (2025)*

### **Statement of the Problem**

The educational system in Nigeria is characterized by the presence of multiple examination bodies, including the West African Examinations Council (WAEC), the National Examinations Council (NECO), and the National Business and Technical Examinations Board (NABTEB), each offering Senior School Certificate Examinations in Economics just like every other subject area. These examinations serve as crucial determinants of students' academic progression and future opportunities. Moreover, the performances of North West students in achieving five (5) credits and above including English language and Mathematics is contrary to expectations when compared to other states of the federation in Nigeria. Likewise, researchers have different views about the quality of the Senior Secondary School Certificate Examinations in Nigeria because it has been observed that candidates, educational institutions, employers of labour and other end users have preference for certificates of certain examination bodies when compared to others (Olatunji, 2015). Also, Olutola (2011), Olatunji (2007) and Alfred (2011) remarked that the standard of SSCE conducted by NECO is low compared to SSCE conducted by WAEC. Therefore, this study plumb into the qualities of WAEC, NECO and NABTEB (SSCE) test items on Economics Multiple choice papers by applying Linear equating method to analyze the test forms in North West, Nigeria. This allowed for comparability of scores and improve the overall quality and consistency of examinations, ensuring that all students are evaluated fairly and that the results are comparable, regardless of which examining body administers the tests.

### **Objectives of the Study**

The main aim of this study was the analysis of linear equating of test forms across WAEC, NECO and NABTEB Senior School Certificate Examination (SSCE) of Multiple-Choice Test Items on Economics in North - West, Nigeria.

Specifically, the study sought to find out the:

1. form of student's performance on WAEC, NECO and NABTEB Economics multiple choice papers in North West, Nigeria.
2. pattern of students' performance on the common items of WAEC, NECO and NABTEB Senior School Certificate Multiple Choice papers in Economics in North West, Nigeria.
3. pattern of students' performance on the unique items of WAEC, NECO and NABTEB Senior School Certificate Multiple Choice papers in Economics in North West, Nigeria.
4. results of linear equating of WAEC, NECO and NABTEB Senior School Certificate Multiple Choice Test Items in Economics papers using standard score deviates in North West, Nigeria.

### **Research Questions**

This study provided answers to the following research questions:

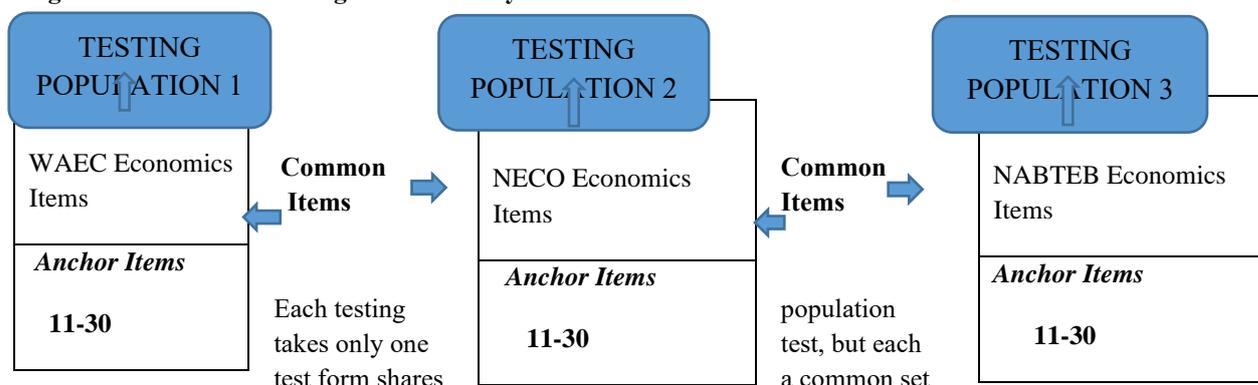
1. What is the form of student's performance with regards to WAEC, NECO and NABTEB test form in Economics multiple choice papers in North West, Nigeria?
2. What is the pattern of student's performance on the common items of WAEC, NECO and NABTEB Senior School Certificate Multiple Choice papers in Economics in North West, Nigeria?
3. What is the pattern of student's performance on the unique items of WAEC, NECO and NABTEB Senior School Certificate Multiple Choice papers in Economics in North West, Nigeria?
4. What are the results of linear equating of WAEC, NECO and NABTEB Senior School Certificate Multiple Choice Test Items in Economics papers using standard score deviates in North West, Nigeria?

### **Methodology**

The researcher adopted Non-Equivalent Groups Anchor Test (NEAT) design for this study, also known as Common-items Non-Equivalent Group (CINEG) design, is said to be the most flexible tool available for equating tests (Kolen & Brennan, 2014). Shown below is an Anchor test design:

*Analysis of Linear Equating System in ... Olutola, A. T., Et. al. (2025)*

**Figure 1: Anchor Test Design for the Study**



of test questions (*anchor items*).

**Source:** Researcher schema plan (2024)

The population of the study comprises of four hundred and four thousand, five hundred and twenty-one (404,521) public senior secondary school students across the seven states that made up North West geopolitical Zone, Nigeria. Among these students, there are two hundred and thirty-four thousand, five hundred and fifty-two (234,552) males and one hundred and sixty-nine thousand, nine hundred and nine (169,969) females. The sample size of one thousand, five hundred and thirty-two (1,532) Senior Secondary School II (SS2) Economics students were derived from a target population of two hundred and sixteen thousand and one (216,001) Senior Secondary School II (SS2) students in eighteen (18) sample schools from three (3) selected old states such as Kaduna, Kano and Sokoto States respectively in North West geopolitical Zone, Nigeria, by adopting multistage sampling techniques.

The researcher determined the content validity and as well the reliability coefficient of the instruments and obtained Coefficients of content validity 0.71, 0.68 and 0.65 respectively by calculating the percentage-difference of the coefficient of correlation between the test forms and syllabi contents with a formula  $r_{ts} = 1 - \sum |d| / 100$ , (where  $r_{ts}$  is the correlation between test and syllabus contents),  $d$  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, the instrument were given to Economics experts and as well test and measurement experts (experienced teachers) to give keys to the items. While reliability coefficient for test form A was 0.75, form B was 0.89 and form C had 0.88 using split half method of internal consistency. The instrument for data collection were the adapted 2023 WAEC, NECO, and NABTEB Economics multiple-choice papers which comprises of unique and anchor items that were positioned as a block at numbers 11 - 30 in each test form for this study. This is referred to as appended internal anchor (Ryan, 2011). The data collected for this study were subjected to descriptive statistics of mean, standard deviation and standard score deviate of T-Score to answer the four generated research questions with the aid of SPSS Version 23 to equate the scores obtained from the test forms.

## Presentation of Results

**Research Question One:** What is the form of student's performance with regards to WAEC, NECO and NABTEB test form on Economics multiple choice papers in North West, Nigeria?

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

**Table 1: Form of student's performance with regards to WAEC, NECO and NABTEB test form**

Test forms	No of Students	Minimum	Maximum	Mean	Std. Deviation
WAEC (A)	511	2	28	12.28	4.022
NECO (B)	510	2	36	11.92	4.463
NABTEB(C)	511	1	36	13.49	5.692

Table 1 above showed the result of mean form of student's performance with regards to the various test form under consideration. The mean form of students' performance in test form A was 12.28, while the mean form of student's performance in test form B and C are 11.92 and 13.49 respectively. Respondents' highest scores in the three test forms were 28 and 36 respectively and their lowest score was 1 and 2. Also, test form A had standard deviation of 4.022, while test form B and C had 4.463 and 5.692 as their standard deviation respectively.



*Analysis of Linear Equating System in ... Olutola, A. T., Et. al. (2025)*

To further understand the differences between the test forms, t-test were perform to compare the means of each pair of forms. The results of the t-test are as follows in table 2:

**Table 2: Paired Samples t- Test of Form A-B, A-C and B-C**

Test Form Pair	Mean	Std. Deviation	Std. Error Mean	T	Df	Sig	Decision
A-B	0.355	5.708	0.253	1.404	509	0.161	NS
A-C	-1.213	6.132	0.271	-4.473	510	.000	Sig
B-C	-1.553	0.289	0.289	-5.375	509	.000	Sig

$\alpha = 0.05$ , NS= non-significant

T-tests were conducted to compare the means of each form pair. The t-values represent the difference between equated test forms and a higher absolute value of the t-value suggests a greater difference. Results in Table 3 above showed that there is significant differences in difficulty between the pair of form A-C ( $t=4.473=.000$ ,  $p < 0.05$ ) and B-C ( $t=5.375=.000$ ,  $p < 0.05$ ), while a pair of A-B form showed no significant difference ( $t=1.404=.161$ ,  $p > 0.05$ ), this meaning that there is no statistical difference in students' ability or proficiency in Economics multiple-choice test items of form A and B (WAEC & NECO) in year 2023. While the pairs of other test forms A-C and B-C indicated significant differences, thus equating process can take place using Linear equating methods.

**Research Question Two:** What is the pattern of student's performance on the common items of WAEC, NECO and NABTEB Senior School Certificate Multiple Choice papers in Economics in North West, Nigeria?

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

**Table 3. Mean of Respondents' Pattern of Performance in Common Items**

Test forms	No of Students	Minimum	Maximum	Mean	Std. Deviation
WAEC (A)	511	0	13	4.94	2.121
NECO (B)	510	0	17	4.55	2.114
NABTEB (C)	511	0	18	5.10	2.504

Table 3 above showed the result of mean performance of respondents in common items test. Test form A (WAEC) mean performance was 4.94, while the mean performance in test form B (NECO) and C (NABTEB) are 4.55 and 5.10, (indicating equal mean of 5 if round off to 1 sig. fig.) respectively. Respondents' highest scores in the three test forms were 13, 17 and 18 respectively and Zero (0) was their lowest score. Also, test form A had standard deviation of 2.121, while test form B and C had 2.114 and 2.504 as their standard deviation respectively.

One-way ANOVA was conducted if there is any significant difference in the test-takers' performance at 0.05 level of significant in the common items. This is to ensure that condition for equating using Non-equivalent method is not violated, Check the table 4 below:

**Table 4: ANOVA Summary Table of Respondents' Performance in Common Items**

	Sum of Squares	Df	Mean Square	F	Sig.	Decision
Between Groups	10.331	3	10.331	2.314	1.909	NS
Within Groups	8012.812	1526	15.249			
Total	8023.143	1529				

$\alpha = 0.05$ , NS= non-significant

Table 4 above reveals an F-value of 2.314, which is not significant at the 0.05 alpha level, since the p-value of (1.909) is greater than 0.05 ( $1.909 > 0.05$ ). This therefore, implies that test-takers are not statistically different in proficiency/ability in common items of Economics multiple-choice test items conducted by the WAEC, NECO, and NABTEB in 2023 Senior School Certificate Examination. This is a further confirmation of the usefulness of this test equating process.



*Analysis of Linear Equating System in ... Olutola, A. T., Et. al. (2025)*

**Research Question Three:** What is the pattern of student's performance on the unique items of WAEC, NECO and NABTEB Senior School Certificate Multiple Choice papers in Economics in North West, Nigeria?

The pattern of students' scores in unique items of test form A, B, and C are independently summed and their mean and standard deviation are computed to determine the pattern of respondents.

**Table 5. Descriptive Statistics of the Pattern of Students' Performance on Unique Items**

Test forms	No of Students	Minimum	Maximum	Mean	Std. Deviation	Skewness	Kurtosis
WAEC (A)	511	0	19	7.30	2.977	.310	.197
NECO (B)	510	0	24	7.40	3.524	.616	1.182
NABTEB (C)	511	0	22	8.34	4.178	.645	.157

Table 5 showed the statistical pattern of student's performance on the unique items. Test form C with a mean performance of 8.34, standard deviation of 4.178, positive Skewness of .645 and kurtosis of .157 indicated best performance. Form C has the highest Skewness .645, suggesting a greater concentration of high scores and have Kurtosis values of .157 that is closer to zero. Test form B with a mean performance of 7.40, standard deviation of 3.524, positive Skewness of .616 and kurtosis of 1.182 indicated a better performance. Form B Kurtosis has a more peaked distribution with heavier tails compared to forms A and C. While testees who did test form A had the least performance with mean of 7.30, standard deviation of 2.977, positive Skewness of .310 and kurtosis of .197. Forms A and C have kurtosis values closer to zero, suggesting a distribution closer to a normal distribution.

**Research Question Four:** What are the results of linear equating of WAEC, NECO and NABTEB Senior School Certificate Multiple Choice Test Items in Economics papers in North West Nigeria, using standard score deviates?

Here, students' scores in test form A, B, and C (WAEC, NECO and NABTEB) are independently standardized using T-score. The T-scores is used as a scale on which examinees' raw score in each of the test form were compared.

**Table 6: Linear Equating of WAEC, NECO, and NABTEB Economics SSCE Multiple Choice Papers**

T-SCORE	Students' raw scores in WAEC (A)	Students' raw scores in NECO (B)	Students' raw scores in NABTEB (C)
25	2	-	-
26	-	-	-
27	-	-	-
28	3	-	1
29	-	-	-
30	4	2	2
31	-	-	-
32	-	3	3
33	5	-	4
34	-	4	-
35	6	-	5
36	-	5	-
37	-	-	6
38	7	6	7
39	-	-	-
40	8	7	8
41	-	-	-
42	-	8	9
43	9	-	10
44	-	9	-
45	10	-	11
46	-	10	-
47	-	-	12
48	11	11	13
49	-	-	-
50	12	12	14
51	-	-	-
52	-	13	15



*Analysis of Linear Equating System in ... Olutola, A. T., Et. al. (2025)*

53	13	-	16
54	-	14	-
55	14	-	17
56	-	15	-
57	-	-	18
58	15	16	19
59	-	-	-
60	16	17	-
61	-	-	-
62	-	18	21
63	17	-	22
64	-	19	-
65	18	-	23
66	-	20	-
67	-	-	24
68	19	21	25
69	-	-	-
70	-	22	26
71	-	-	-
72	-	23	27
73	21	-	28
74	-	24	-
75	22	-	29
76	-	26	-
77	-	-	-
78	23	-	-
79	-	-	-
80	-	27	-
81	-	-	-
82	-	-	-
83	25	-	-
84	-	-	-
85	-	-	-
86	-	-	-
87	-	-	36
88	-	-	-
89	-	-	-
90	28	-	-
91	-	-	-
92	-	-	-
93	-	-	-
94	-	34	-
95	-	-	-
96	-	-	-
97	-	-	-
98	-	36	-

However, in order to establish the outcome of Linear Equating under the NEAT design, the summary of equated scores of WAEC, NECO and NABTEB are presented in Table 7 below:

**Table 7: Linear Equating Comparison Summary of Economics SSCE Multiple Choice Papers.**

T-SCORE	Students' raw scores in WAEC (A)	Students' raw scores in NECO (B)	Students' raw scores in NABTEB (C)
30	4	2	2
38	7	6	7
40	8	7	8
48	11	11	13

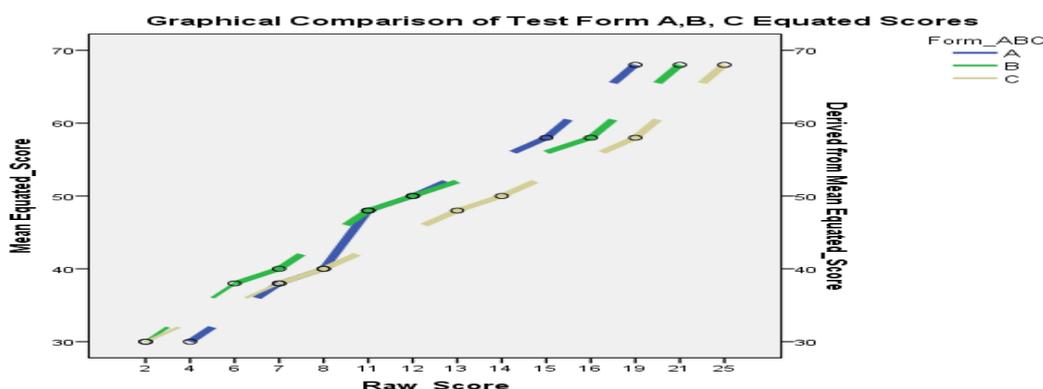


*Analysis of Linear Equating System in ... Olutola, A. T., Et. al. (2025)*

<b>50</b>	12	12	14
<b>58</b>	15	16	19
<b>68</b>	19	21	25

The result of table 7 above revealed linear relationship between test Form A (WAEC), B (NECO) and C (NABTEB) and the resulting standard score deviate across the test forms under the NEAT design, meaning that equal increases in form A, B, and C lead to equal increases in T-scores. However, when the Students' raw score in test form A (4), B and C were (2) respectively, the corresponding standard score deviate is (30), when A (7), B (6) and C (7), the corresponding standard score deviate is (38), when A (8), B (7) and C (8), the corresponding standard score deviate was (40), also, when A and B were (11) respectively and C having (13), the corresponding standard score deviate is (48). Linear equating also revealed that standard score deviate of (50) corresponding to form A and B (12) respectively with C having (14), while standard score deviate of (58) corresponding to form A (15), B (16) and C (19) respectively, while standard score deviate of (68) was equivalent to form A (19), B (21) and C (25) respectively. Thus, T-score increases as the value of test form A, B and C increases suggesting a positive relationship using Linear equating under NEAT design. This result further revealed that linear equating of WAEC and NECO Economics multiple choice items were more equivalent when compared to NABTEB multiple choice items as display in figure 2 line graph below.

**Figure 2: Line Graph of linear equating showing equated scores from test forms A (WAEC), B (NECO) and C (NABTEB) SSCE Economics multiple choice papers**



Graph presented in Figure 2 above showed the scores obtained from the three test forms. Test form C had higher values of equated scores 25, while test forms A and B had very close values, 19 and 21 respectively. It can be said that the two test forms (A & B) produced almost equivalent scores.

### Findings of the Study

The key results of this study are:

1. There was significant differences in difficulty between pair of forms A-C and form B-C, but no significant difficulty differences in form A-B at 0.05 level because the test forms appear to be equivalent.
2. There was no significant differences in examinee proficiency on common items across the three test forms.
3. Performance of testees on unique items varied significantly across the three test forms. It revealed that Form C (NECO) had the highest performance, while form A (WAEC) had the lowest.
4. Linear equating result indicated that T-scores increased as WAEC, NECO, and NABTEB raw scores increases, which suggests positive linear relationship between the forms, meaning each form contributes to the overall T-score.

### Discussion of Findings

The results of the study provided an overview form of students' performance on three different test forms A (WAEC), B (NECO) and C (NABTEB) SSCE of the year 2023 in Economics multiple-choice test items. The results of test forms A, B, and C revealed some differences in difficulty and variability, suggesting that examinees differed in their proficiency level. Form C appears to be the easiest, with the highest mean score of (13.49), Form B seems to be the most difficult, with the lowest mean score (11.92), while Form A falls in between, with a mean score of (12.28), indicating moderate performance of test takers. Also, the standard deviation of test Form C shows the highest variability in scores (5.692), indicating a wider range of performance, Form A has the lowest variability of (4.022), suggesting a narrower range of



*Analysis of Linear Equating System in ... Olutola, A. T., Et. al. (2025)*

scores, while Form B falls between A and C in terms of variability (4.463). However, to further understand if there are differences between the test forms (A,B &C), t-tests were conducted to compare the means of each form pair and the resulted t-test revealed that Form C is significantly different from both Form A and Form B, with pair of form A-C ( $t=4.473=0.000$ ,  $p < 0.05$ ) and pair of B-C ( $t=4.375=0.000$ ,  $p < 0.05$ ). These p-values are less than the significance level of 0.05, indicating that the differences between a pair of forms A-C and as well forms B-C are statistically significant. This finding is supported by the study of Alfred (2013) that there is a significant difference in the difficulty level of Economics multiple choice items conducted by WAEC, NECO and NABTEB. Therefore, the researcher considered linear equating methods to adjust for these differences due to the significant differences between Form C and the other two forms, as direct score comparisons may not be appropriate. While a pair of form A-B showed no significant difference ( $t=1.404=.161$ ,  $p > 0.05$ ), which means that there is no statistical difference in students' ability or proficiency in Economics multiple-choice test items of WAEC and NECO in year 2023. Hence, the two examinations could be used interchangeably. The finding corroborated those of Bandele and Adewale, (2013) where they submitted that there is no significant difference in the difficulty level, reliability and validity coefficients of mathematics items constructed by WAEC, NECO and NABTEB.

The result of the common items revealed that respondents are similar in proficiency/ability. Thus, test forms A, B and C (WAEC, NECO & NABTEB) had their mean 2.121, 2.114 and 2.504 respectively, which are not significant at 0.05 alpha level. This suggested that the average difficulty levels of the year 2023 Economics multiple-choice test items conducted by WAEC, NECO, and NABTEB (SSCE) are the same under the common items and could be used interchangeably. The inclusion of common items is to ensure that the conditions for carrying out scores equating for WAEC, NECO and NABTEB multiple choice items in Economics was not violated. Therefore, this finding is in line with the submission of Doran, Moses and Eignor (2010) 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. But the finding is contrary to the study of Arokoyo (2021) that examinees differed in their proficiency level because the mean performance on their common items scores on the three test forms A, B and C were 5.46, 6.04 and 7.36 respectively, with submission that, the result might be attributed to differences in the proficiency level of testees who sat for the three test forms. Meanwhile, differences in difficulty are what test score equating take care of but, differences in examinees ability or proficiency are a confounding factor that needs to be eliminated before the equating process can take place Daramola et al. (2022).

In unique items, this study revealed that there were differences in the performance of examinees across the three test forms. The unique items of test forms A, B and C had mean performance of 7.30, 7.40 and 8.34 respectively, with form C having the highest mean (8.34) and form A having the lowest (7.30). This suggests that Form C might be slightly easier than the other two forms. Also form A, B and C revealed standard deviation of 2.977, 3.524 and 4.178 respectively, with scores on Form C are mostly spread out (4.178) and scores on Form A are the least spread out (2.977). This implies that there is more variation in the performance on Form C. This finding is consistent with Kolen and Brennan's (2004) submission that it is virtually impossible to construct multiple forms of tests that are strictly parallel. Hence, equating is necessary because it adjusts for differences in difficulty among or across test forms constructed to be similar in difficulty and content just like the senior school certificates examinations investigated. The finding also corroborated those of Daramola et al, (2022) and Arokoyo, (2021) where they submitted in their separate studies in Economics and Chemistry that there were differences in the performance of examinees across the three test forms in unique items.

Results from linear equating between test Form A (WAEC), B (NECO) and C (NABTEB) Senior School Certificate Economics multiple choice papers with the use of standard T-score deviates showed that T-scores increases as the values of WAEC, NECO and NABTEB increases after the transformation, suggesting a positive linear relationship between the forms, meaning each form contributes to the overall T-score. The T-score increment are generally consistent, reinforcing the assumption of linearity in equating process, though, differences observed in T-scores at different score levels suggest the relative difficulty and contribution of each test form. Findings on linear equating of WAEC, NECO, and NABTEB reveal a low score range of 4 in WAEC, 2 in both NECO and NABTEB respectively corresponding to T-score of 30, while the mid score range is 12 in both WAEC and NECO respectively and NABTEB having 14 corresponding to average T-score of 50, suggesting a balance level of performance. And the highest raw score of 19, 21, and 25 in WAEC, NECO and NABTEB respectively corresponding to a T-score of 68, indicating above average performance.



*Analysis of Linear Equating System in ... Olutola, A. T., Et. al. (2025)*

Therefore, the relatively even increases in T-score (except for minor variations) reinforce that each test form contributed equally or proportionally to the equated score which are comparable in difficulty when adjusted using the linear model. This finding is in line with the studies of Daramola et al. (2022) and Arokoyo (2021) that there are closely equivalent scores in WAEC and NECO when compared to NABTEB Economics (SSCE) papers which can be attributed to a relatively long period of operation of these two Senior School Certificate Examination bodies. This finding also corroborates the assertion of Obinne, Nworgu, and Umobong (2013) that there is no significant difference in the DIF of Biology tests conducted by WAEC and NECO which is also supported by the submission of Obinne (2011) that Biology test conducted by WAEC and NECO were equally reliable. But the finding is contrary to that of Olutola (2011) who found out that WAEC have more difficult items than NECO multiple choice Biology items.

### Conclusions

The WAEC, NECO, and NABTEB SSCE on Economics multiple-choice tests form of year 2023 were analyzed using linear equating method in North-West Nigeria. However, based on the findings from this research work, the researchers therefore concluded that form C (NABTEB) was distinctly easier, indicating significant differences from test form A and B (WAEC & NECO) that shown no statistical difference in students' ability or proficiency, as both test items are comparable and of similar standard on the Economics multiple-choice items. Conversely, no significant differences in examinee proficiency on common items across the three test forms, while respondents on unique items performed differently across test form A, B and C respectively. Also, the researcher concluded that the linear equating method was found to be consistent and précised while showing a positive linear relationship between the forms, with T-scores increasing as WAEC, NECO, and NABTEB raw scores increased. This indicates that each form contributes to the overall T-score, supporting the idea that the tests measure similar underlying constructs and that the linear equating method is suitable for these forms.

### Recommendations

1. The researcher recommends a thorough analysis of form C's content, question types, and cognitive demands. Then Compare C to forms A and B to identify the source of the difficulty difference.
2. Despite the consistency of the common items in this study, the researcher still recommends a periodic review of the alignment of future test form with the curriculum to ensure that the same intended constructs are measure continuously.
3. The researcher recommended a detailed analysis of the unique items to further understand why testees' performance on unique items varied significantly across the three test forms. This is vital for future test development.
4. The researcher recommends the establishment of regulatory body to standardize and monitor examinations conducted by these examining bodies in Nigeria.
5. After considering linear equating appropriateness in this context, the researcher recommended linear equating method for equating scores across these WAEC, NECO, and NABTEB Economics multiple-choice test,

### References

- 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>
- Agah, J.J. (2013). *Relative efficiency of test scores equating methods in the comparison of students' continuous assessment measures*: PhD thesis. Nsukka: University of Nigeria.
- Albano, A. D. (2011). *Equate: Statistical methods for test equating [Computer software manual]*. Available from <http://CRAN.R-project.org/package=equate> (R package).
- Alfred, O. (2013). *Assessment of the equivalence of 2009 senior school certificate multiple -choice Economics test items in Ilorin, Nigeria*. M.Ed thesis, University of Ilorin. Ilorin.
- Arokoyo (2021), *Levine Linear and Chained Equipercentile Equating Of Senior School Certificate Examination Chemistry Multiple-Choice Papers in South-West, Nigeria*. A Ph.D. Thesis at the University of Ilorin, Nigeria.
- Bande, 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.
- Dorans, N. J., Moses, T. P., & Eignor, D. R. (2010). *Principles and practices of test score equating* (ETS Research Report No. RR-10-29): New Jersey: ETS Princeton.
- Daramola, D.S.; Olutola, A. T. and Jimoh, M. I. (2022). Analysis of Linear Equating of Senior School Certificate Multiple Choice Examination Papers in Economics. *SPEKTA: Jurnal Pengabdian Kepada Masyarakat: Teknologi dan Aplikasi*, 3 (2), 147-158.



*Analysis of Linear Equating System in ... Olutola, A. T., Et. al. (2025)*

- Daramola, D.S.; Owolabi, H.O. and Olutola, A.T. (2018). Equipercentile Equating of WAEC, NECO and NABTEB Senior School Certificate Multiple Choice Test Items in Economics. *Journal of Education in Developing Areas (JEDA)*, 26 (2), 658-668.
- 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.
- Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices* (3rd ed.). Springer Science & Business Media.
- Obinne, A. D. E. (2011). A psychometric analysis of two major examinations in Nigeria: Standard error of measurement. *International Journal of Educational Sciences*, 3(2), 137-144.
- Obinne, A.D.E., Nworgu, B.G., & Umobong, M.E. (2013). An investigation into differential item functioning of tests conducted by the two major examination bodies in Nigeria. *Advances in Educational Research* 2 (1), 1-8.
- 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.
- Olatunji, D. S. (2015). *Analysis of linear and equipercentile equating of senior school certificate examination economics multiple-choice papers In Kwara State, Nigeria*. Unpublished Ph.D. Thesis, University of Ilorin.
- Olutola, A.T. (2011). Analysis of item parameters of senior school certificate multiple choice in biology in Ekiti State, Nigeria. Unpublished Ph.D Thesis, University of Ilorin, Ilorin.
- Olutola, A.T. (2015). Item Difficulty and Discrimination Indices of Multiple Choice Biology Tests. *Liceo Journal of Higher Education Research*, 11(1), 16-30.
- Ryan, J. (2011). *A Practitioner's Introduction to Equating with Primers on Classical Test Theory and Item Response Theory*. Arizona State University Frank Brockmann, Center Point Assessment Solutions.
- Uysal, I. & Kilmen, S. (2016). Comparison of item response theory test equating methods for mixed format tests. *Int. Online Journal of Educational Sciences*. 6(2) 1-11.
- Wiberg, M. & von Davier, A. A. (2017). Examining the impact of covariates on anchor tests to ascertain quality over time in a college admissions test. *International Journal of Testing*. 1-22. 105-126.