



Assessment of Random and Systematic Errors in 2024 Neco Ssce Mathematics Multiple Choice Test Items in Ethiope East Local Government Area, Delta State

Praise Kehinde ADEOSUN (PhD)

Department of Educational Evaluation and Counseling Psychology
University of Benin, Edo State, Nigeria.
<https://orcid.org/0009-0003-1870-978X>
E-mail: praise.adeosun@uniben.edu
Tel: 08060481557

and

Emamuyovwi Esume

Department of Educational Evaluation and Counseling Psychology
University of Benin, Edo State, Nigeria.
<https://orcid.org/0009-0002-9737-2793>
Email: emamuyovwiesume@gmail.com
Tel: 08032307561

Accepted 17th February 2026

Published 27th February 2026

How to Cite:

Adeosun, P. K., & Esume, E. Assessment of Random and Systematic Errors in 2024 Neco Ssce Mathematics Multiple Choice Test Items in Ethiope East Local Government Area, Delta State. *Msingi Journal*, 9(1). <https://doi.org/10.33886/mj.v9i1.760>

ABSTRACT

This study investigated random and systematic errors in 2024 NECO SSCE Mathematics Multiple Choice Test Items among senior secondary school three students in Ethiope East Local Government Area, Delta State, Nigeria. The study specifically assessed extent of random and systematic errors and also the difference between observed random and systematic errors in 2024 NECO SSCE mathematics multiple choice test items administered to examinees in 2024. Three research questions guided the study. The study adopted a survey research design on a population of 1,270 SSS III students in Ethiope East Local Government Area of Delta State. The sample size was 320 selected through random sampling techniques in two stages. The instrument used for data collection test was adopted from NECO. The instrument was the

official NECO standardized achievement mathematics multiple choice 60 test items with four options A-D administered and used for the in-school Senior School Certificate Examination (SSCE) in 2024. The instrument was not validated because it was standardized. For the reliability of the instrument, it was administered on twenty students twice with three weeks' time gap and subjected to Pearson Product Moment Correlation statistics to establish its dependability. The reliability coefficient of .84 was obtained. The data used for the study were collected independently by the researcher with the assistance of classroom teachers. Data were analysed using Cronbach Alpha, Paired t-test and variance statistics respectively in R software. Findings revealed that the 2024 NECO SSCE Mathematics multiple-choice items exhibited 15.6% random error, 6.09% systematic error, suggesting a small but significant level of consistent bias in measurement with random error of 90.48% substantially higher than systematic error of 9.52%. It was concluded that the 2024 NECO SSCE Mathematics multiple-choice test items were more affected by unpredictable inconsistencies than by systematic bias. It was therefore recommended that item development procedures need to be enhanced by NECO in order to improve in the aspect of item-writing, moderation, and pilot-testing processes to ensure all multiple-choice items are clear, discriminating, and free from ambiguity that could increase random error. Also, that teachers; should adequately expose students to standardized testing conditions and develop classroom assessments aligned with NECO standards to minimize performance fluctuations in summative assessment.

Key Words: assessment, Random Errors, Systematic Errors

INTRODUCTION

Test is one of the measurement instruments in the educational system used to measure students' traits or behaviour, especially achievement test. Tests can be described as structured assessments designed to evaluate knowledge, skills, and competencies across various subjects and grade levels. Test items construction often employed systematic method to write final items to measure student's understanding, abilities, or performance in a specific areas or subject. Test, as an educational measurement tool, plays a vital role in assessing students' learning, guiding instructional practices, and informing educational policies. It seems to typically involve a set of questions or tasks that students must complete under controlled conditions. The responses supplied by students on test items presented to them are scored and often quantified numerically to ascertain the total score of test takers, students are further judge based on their total score; a process that can be described as measurement and evaluation.

However, measurement often contains errors which may be due to the imprecision of educational measurement instrument, inappropriate test administration, or other factors related to test takers. Measurement errors due to chance factors is often termed as random errors while errors due to technical factors can be correctly described as systematic errors. Random errors may occur from unpredictable factors such as guessing, fatigue or inconsistent test environment, while systematic errors may result from technical factors such as item parameters, item difficulty and item discrimination indices and ambiguous wording. Also, in subjects like Mathematics, symbols could be confusing to students (such as + and * as well as \ and $\sqrt{}$). Students may overlook a symbol completely such as a minus sign, skip a question and not to come back to it due to lack of item clarity. Hence, test takers total scores could be affected by errors due to chance factors and is often termed as random errors while errors due to technical factors can be correctly described as

systematic errors. Identifying and reducing this error is very important in assessment of students. Test as an instrument for measurement serves as form of data for assessment.

Assessment can be described as a systematic process used to evaluate and measure students' knowledge, skills, abilities, and performance because it is a critical tool for measuring students' academic performance, providing valuable insights into their understanding and skills acquisition. Assessment data often identify strengths and weaknesses, guiding instructional strategies and interventions. Assessment feedback may be culminated to formative or summative assessments depending on its use. Formative assessment is a continual process that tracks students' development and offers feedback to improve teaching methods. Summative assessment occurs at the end of an instructional period to evaluate overall learning outcomes and effectiveness. Summative Assessment seems to be high-stakes assessment that contributes to final grades and provides a comprehensive overview of what students have learned. Assessment at the end of a programme may lead to evaluation where valued-judgements are made using some form of test items.

Evaluation may be used for placement, promotion, and certification among others. Evaluation and assessment are often related because the data obtained from measurement using tests can provide vital information for decision making. Hence, summative assessment done at the end of a term or programme is important and should be free from bias due to random and systematic errors. Various examining bodies conduct examinations in Nigeria using test items that could provide data for measurement, assessment and evaluation programmes such as the West African Examination Council (WAEC) and National Examination Council (NECO). But NECO is the researcher's primary emphasis.

NECO has grown to become one of the most well-known testing organisations, administering tests in a variety of academic topics. All students are required to take Mathematics, which is also one of the disciplines on which NECO administers tests each year. The Multiple-Choice (MC) and essay sections of the Mathematics test are intended to evaluate students' understanding and application of mathematical topics. Mathematics as a school subject plays a crucial role in the educational system of Nigeria, influencing various fields such as science, technology, engineering, and economics among others. The teaching and learning of Mathematics is essential for developing critical thinking and problem-solving skills among students. Mathematics is the abstract science of numbers, quantity, structure, space, and change. It encompasses various fields, including arithmetic, algebra, geometry, calculus, and statistics. Mathematics is not only a body of knowledge but also a language used to describe patterns, relationships, and phenomena in the natural world. Its teaching and performance is therefore crucial to the Nigerian education system. Critical observation showed that most candidates sitting for the subject fail during the Senior School Certification Examination across the nation. Could random and systematic errors be the cause of this failure in the summative assessment of NECO Mathematics assessment?

The precision of NECO test items related to random and systematic errors is vital for policy and practice in ensuring quality assurance in assessment data. However, there seems to be lack of comprehensive research focusing on random and systematic errors combined in NECO Mathematics Multiple-Choice Items (NMCTI) used for high/large scale assessments from the literature and

personal experience. It appears that no comprehensive research specifically focused on random and systematic errors in NECO Mathematics Multiple-Choice Items; rather studies often focus on item psychometric properties like difficulty and discrimination. Hence there is need for investigation.

Concept of Summative Assessment

Summative assessment in measurement refers to a form of evaluation conducted at the conclusion of an instructional period, such as the end of a term, to measure academic programmes. Its main goal is to assess the extent to which students have achieved the intended learning outcomes. These assessments are typically formal, graded, and carry substantial weight in determining a student's final academic standing. Summative assessment could be in form of standardized exams like WAEC, NECO, and JAMB among others. Effective summative assessments allow students to demonstrate mastery of key learning objectives, receive comprehensive feedback, and showcase their ability to integrate course concepts into a cohesive understanding.

Summative assessment evaluates student learning at the end of a course, focusing on the final product and determining grades and progression. According to Lane (2018), summative assessments are typically designed with standardized criteria, to ensure they can be administered consistently across large groups of students, multiple cohorts, at different time frame. The data gathered from these summative assessments whether at the individual student, cohort, school, or system level provides educators and administrators with valuable insights to evaluate student knowledge in relation to intended learning objectives. Additionally, this data enables comparisons between current and previous cohorts, as well as across different schools.

Blankman (2024) emphasized the importance of Mathematics summative assessment as a force for good. Also, summative assessments can measure the overall performance of a district, state, or even country when assigned rigorously and judiciously. Similarly, Ngunjiri (2022) explained that assessment in Mathematics classrooms reviews the role of assessment in bringing changes in students' Mathematics performance. Also, that item quality assessment is a key factor in improving the learning of Mathematics as the relationship between learning and summative assessment is strong and robust.

Mathematics as a school subject is a fundamental subject that deals with numbers, quantities, and shapes. It involves logical reasoning, problem-solving, and critical thinking to understand and analyze various mathematical concepts. Applications of knowledge expand beyond the science-based subject to other fields of study. The benefits of studying Mathematics equips individuals to develop problem-solving skills such as critical thinking, analytical skills, and problem-solving strategies. It also improves logical reasoning which potentially enhances logical reasoning and analytical skills. Mathematics is fundamental to other fields of study putting it as a career opportunity enhancer ranging from sciences, social sciences, humanities among others. Mathematics as a subject is examined for certification for senior secondary graduating class generally described as summative assessment by WAEC, JAMB and NECO and other examination bodies in a large scale assessment. NECO as an examination body introduced amidst the existence of WACE have gained popularity over the years.

In summative assessments, tests are used to evaluate cumulative learning outcomes. Random errors pose a significant threat to reliability, because assessment inform high-stakes decisions, such as graduation or certification, it implies that, score fluctuations due to random error can misrepresent a learner's true competence.

Concept of Measurement Errors

Measurement errors in educational assessment refers to the discrepancy between a student's observed score and their true score on a given construct or trait being measured. It represents the difference between the actual (potentially inaccurate or biased) data yielded by the assessment and the ideal, error-free information that accurately reflects the learner's true ability, knowledge, or characteristic.

In simplified terms, measurement error can be expressed as:

$$x = xo - xt$$

Where:

- x is the measurement error,
- xo is the observed (possibly distorted) score obtained from the assessment, and
- xt is the true score, which represents the student's actual level of performance or ability, free from error.

This framework reflects the principle that assessments in education do not perfectly measure the target trait (e.g., intelligence, reading ability, or test anxiety), and that every score contains some level of error, whether from the test-taker, the instrument, or the testing environment. Measurement error may arise in two main forms: systematic error, which is associated with issues of construct validity, and random error, which pertains to the reliability of the assessment. While efforts can be made to reduce it, random error can never be fully eliminated. Yashim et al., (2021) characterized measurement errors in educational assessment sources of both systematic and unsystematic errors that are prevalent in educational assessment related to students test misconduct of any kind and error associated with teachers such as items development, test administration and scoring process among others. Yashim et al., further stated that, errors can be controlled to some extent by standardizing the measurement procedures and increasing the number of observations that are sampled for each person hence; the score will be formed by averaging over a sample of observations.

Random Error in Assessment

Random errors in educational assessment refer to unpredictable fluctuations that occur during the assessment process, impacting the reliability of test scores. Random errors are caused by temporary factors like student fatigue, distractions, misinterpretation of questions, or random guessing, particularly in multiple-choice test items (Adeosun & Ukana, 2024; Blankman, 2024). In summative assessments, where tests are used to evaluate cumulative learning outcomes, random errors pose a significant threat to reliability; because assessments inform high-stakes decisions, such as graduation or certification. It implies that, score fluctuations due to random error can misrepresent a learner's true competence.

Multiple-choice items, commonly used in Mathematics assessments, are prone to random errors due to guessing or misinterpretation. Environmental inconsistencies, such as noise or poor lighting,

can also introduce errors. In Mathematics multiple-choice items, random errors can arise from misreading symbols, overlooking units, or encountering test environmental distraction such as noise. In such a case, random errors interfered with total score whereby, actual mathematical ability will not be reflected due to the conditions under which students respond to test items. Understanding random errors is crucial, because chance errors can mask students' true proficiency, leading to misclassification and poor instructional decisions (Thorndike & Thorndike-Christ, 2010). This highlights the importance of minimizing errors to ensure accurate summative assessments data. Ajayi and Fakorede (2020) stated that the impact of test administration practices on students' performance in standardised assessments especially inconsistencies in administration related to timing, supervision, and noise introduced random fluctuations in students' performance in WAEC and NECO examinations which in turn impact the reliability of the examination outcomes. Flaws like guessing contribute to random errors (Alruwais. 2020; Daramola & Akinkuotu, 2022). Imasuen and Adeosun (2023) investigated the application of Generalizability Theory on measurement errors in the 2019 WAEC Mathematics objective examination for senior secondary students. The results showed that the student-teacher interaction accounted for the largest proportion of measurement error.

Systematic Error in Assessment

Systematic errors in educational assessment are consistent, repeatable inaccuracies that occur due to flaws in test design, administration procedures, or content bias, which affect test scores in a predictable manner. Unlike random errors, systematic errors have a directional influence consistently increasing or lowering examinees scores. In Mathematics multiple-choice assessments, systematic errors can arise from encountering of poorly worded items or item with inappropriate difficulty index. Systematic errors either have an advantage or disadvantage for certain group of students not due to difference in ability but due to error from the test instrument or test items. Also, in subjects like Mathematics, symbols could be confusing to students (such as + and * as well as \ and $\sqrt{}$). Students may overlook a symbol completely such as a minus sign, skipped a question and not to come back to it due to lack of item clarity. As Thorndike and Thorndike-Christ (2010) explained; systematic errors compromised the validity of summative assessments by ensuring test scores to reflect something other than the intended construct. Batubo, and Goodness (2022), conducted a study on analyses of reliability and the content validity of basic education certificate examination mathematics objective test items from 2018 – 2020. The findings revealed that objective test items content area were not adequately covered.

Both types of errors threaten the validity and reliability of test scores, leading to inaccurate conclusions about student ability. However, it should be noted that it is difficult to completely eliminate random and systematic errors. Hence, this study is investigating random and systematic errors on summative assessment of NECO 2024 Mathematics multiple choice assessment in Delta State from the view point of external factors (random errors) and test quality (systematic errors). This constitutes a gap which this study is intended to fill.

Statement of Problem

Mathematics plays a crucial role in the educational development of nations, serving as a foundational discipline for scientific, technological, and economic progress. As a compulsory subject for admission

into most tertiary institutions, Mathematics performance potentially influences students' academic progression and future career prospects especially in summative assessment. Summative assessment plays critical role in decision taking about students potential using total test score and shaping educational outcomes. NECO as an accredited examination body in Nigeria conducts summative assessment that is used to take high-stake decision on examinees. The precision of NECO test items related to random and systematic errors is vital for policy and practice in ensuring quality assurance in assessment data. However, there seems to be lack of comprehensive research focusing on random and systematic errors combined in NECO Mathematics Multiple-Choice Items (NMCTI) used for high/large scale assessments from the literature. Nevertheless, as revealed by NECO statistics, students' performances are dwindling by the year. Could errors be an issue in NECO summative assessment related to random and systematic errors? This study aims to address assessment measurement errors in 2024 NECO SSCE mathematics multiple choice test items with a focus on random and systematic errors detection among students in Ethiopie East Local Government Area, Delt

RESEARCH QUESTIONS

Three research questions were raised to guide this study.

1. To what extent does the 2024 NECO SSCE Mathematics multiple choice test items exhibit random errors among Senior Secondary School students in Ethiopie East Local Government Area?
2. To what extent does the 2024 NECO SSCE Mathematics multiple choice test items exhibit systematic errors among senior secondary school students in Ethiopie East Local Government Area?
3. What is the difference in the random and systematic errors of the summative assessment of 2024 NECO SSCE Mathematics multiple choice test items administered towards secondary school students examination performance in Ethiopie East Local Government Area?

METHODOLOGY

The study was carried out using the survey research design. The population for this study was 1,270 SSS III students. The sample size was 320 students selected through random sampling techniques in two stages. The instrument that was used for this study was an adopted official NECO 2024 Mathematics multiple choice questions for the Senior School Certificate Examination (SSCE). The 2024 Senior School Certificate Examination contains 60 items with four options A-D. The multiple choice items constructed were validated and standardized by the National Examination Council before it was administered to students. Hence, the instrument was not subjected to further validation by the researcher. The reliability of the instrument was established by administering the test items to 20 senior secondary school student using test-retest method. It yielded a reliability coefficient of .84. The responses of examinees were scored 1 for correct option and 0 for incorrect option using a scoring key. Thereafter, research question one, two and three were answered using descriptive statistics, paired t-test statistics and variance respectively in R software.

PRESENTATION OF RESULTS AND DISCUSSION OF FINDINGS

Research Question One: To what extent does the 2024 NECO SSCE Mathematics multiple choice test items exhibit random errors among senior secondary school students in Ethiopia East Local Government Area?

Table 1: Descriptive Statistics of Random Error of 2024 NECO SSCE Mathematics Multiple Choice Items

N	r	Random Error (%)
320	.844	0.156 (15.60%)

Table 1, show a p-value of .844 and random error of 0.156 which is 15.60%. This implies that 2024 NECO SSCE Mathematics Multiple Choice Items exhibited random error. The random error accounted for 15.60%.

Research Question Two: To what extent does the 2024 NECO SSCE Mathematics multiple choice test items exhibit systematic errors among senior secondary school students in Ethiopia East Local Government Area?

Table 2: Paired t-test Statistics of Systematics error of 2024 SSCE NECO Mathematics Multiple Choice Items

Test	N	Mean	Std Deviation	Mean Difference	t	Df	Sig.
1	320	19.91	8.49	-1.28	-3.276	319	.000
2		21.19	7.27				

Table 2: shows a calculated t of -3.296, a p- value of .000 and mean difference of -1.28. It is significant p .05. This implies that 2024 NECO SSCE Mathematics Multiple Choice Item exhibited systematics error. The systematics error accounted for 6.09%. The relatively small systematic error (6.09%) indicates that the test maintained reasonable fairness, while the random error (15.6%) reflects some limitations in precision and reliability. Consequently, the NECO SSCE Mathematics test can be described as a fair but slightly imprecise assessment tool.

Research Question Three: What is the difference in the random and systematic errors of the summative assessment of 2024 NECO SSCE Mathematics multiple choice test items administered towards secondary school students' examination performance in Ethiopia East Local Government Area?

Table 3: Descriptive of Difference between Random and Systematics Errors of 2024 NECO SSCE Mathematics Multiple Choice Items

Error Type	Variance	% of Total Error
Random Error	15.60	90.48
Systematic Error	1.64	9.52

Table 3 shows variance of 15.60 and 1.64 for random and systematics errors respectively. It also shows the value of 90.48% and 9.52% for random and systematic errors respectively. Since random error has higher percentage than the systematic error, this implies that the 2024 NECO SSCE Mathematics Multiple Choice Item is mostly affected by random inconsistencies rather than the systematic error.

DISCUSSION OF FINDINGS

The findings from research question one revealed that the 2024 NECO SSCE Mathematics Multiple Choice Items exhibited random error. The random error accounted for 15.60%. This could be as a result of unpredictable or chance factors. These factors could be due to students' fatigue, illness, students' misread question, student guessing in multiple choice item etc. Students may have misread some items, responded to other items aberrantly even when they are highly intelligent and guessing of options in perceived difficult items since errors usually fluctuate across students and items. This result suggests that while the 2024 NECO Mathematics test demonstrates acceptable reliability related to true-score variance it still contains a moderate level of random fluctuations. The findings collaborated the conclusions of Esomonu and Okeaba (2021) who found that NECO Mathematics examinations are prone to hidden sources of random error, affecting score dependability. Also, Imasuen and Adeosun (2023) reported that random error arising from student-item interactions constituted the largest share of error variance in WAEC Mathematics tests. The implication is that the random error found in this study is within the normal range observed in large-scale assessments. It suggests that there is need for students to avoid guessing in test-taking and control their emotions to prevent test anxiety often resulting in random error. The need for clearer instructions on test items used in examinations by item developers reduces random fluctuations.

Findings from research question two showed that the 2024 NECO SSCE Mathematics Multiple Choice Item exhibited systematics error. The systematics error accounted for 6.09%. This could be as result of bias across test items, some items may contain words that are not clear to examinees. It may also be due to poorly calibrated difficulty ranges, leading to uniform bias inherent in the test items. It can be said that that test items favour students who have more in-depth exposure to the content of items used for the study. This findings agree with Bassey et al. (2021), who reported that systematic error remains a challenge in Nigerian assessments due to inadequate test-construction skills and lack of standardization. Similarly, Oyetunde (2022) found residual systematic variance in NECO Mathematics due to rater effects and item-bias patterns. Internationally, Brookhart (2015) and the AERA, APA & NCME (2014) standards emphasis that systematic errors undermine test validity by introducing consistent distortions in measurement. The implication arising from this

finding to measurement and evaluation experts, is that it is necessary to conduct item analysis calibration and differential item detection periodically to provide trend evidence needed for proper items development associated with quality assurance.

The findings from research question three shows the value of 90.48% and 9.52% for random and systematic errors respectively. This implies that the 2024 NECO SSCE Mathematics Multiple Choice Items are affected more by random inconsistencies rather than the systematic error. In other words, the 2024 NECO SSCE Mathematics Multiple Choice Items are basically, affected by random inconsistencies rather than systematic bias. Psychometrically, this reflects an assessment that is fairly unbiased but not optimally precise because of an observed slight technical error across all examinees who took the test in this study. These findings potentially reflect that examinees variations in test-taking strategies, anxiety and stress level introduces variability that are basically random. It is also possible that some items are easier for others while some examinees find similar items to be more difficult potentially accounting for more random errors than systematic errors due to predominate random variations.

The predominance of random error mirrors findings by Imasuen and Adeosun (2023), who observed that random fluctuations were the dominant source of score variability in WAEC Mathematics. Similarly, Ohiozua and Omorogiuwa (2025) findings affirms that random errors commonly occur moderately in 2018 and 2019 May/June Biology objective questions of National Business and Technical Certificate Examination in Nigeria. Oyetunde (2022) noted that random components (student $\times$ item interaction) contributed more to total error than systematic facets in NECO Mathematics. This finding aligns with international research suggesting that large-scale multiple-choice tests often have 80–95% of total error attributable to random variance (Shavelson et al., 2019). Hence, while the 2024 NECO Mathematics test can be regarded as relatively fair, its precision could be enhanced by increasing item numbers, improving discrimination indices, and ensuring consistent test administration conditions.

CONCLUSION

Based on the findings the researcher concluded that the 2024 NECO SSCE Mathematics multiple-choice test items were more influenced by random inconsistencies than by systematic biases. In other words, the test scores were fairly free from directional bias but moderately affected by unpredictable variations in students' performance and testing conditions.

RECOMMENDATIONS

Based on the findings and conclusions drawn from this study, the following recommendations are made:

- Item Development Procedures need to be enhanced by NECO in order to improve in the aspect of item-writing, moderation, and pilot-testing processes to ensure all multiple-choice items are clear, discriminating, and free from ambiguity that could increase random error.
- Adoption of innovations by examination bodies whereby reliability coefficients and standard errors of measurement along with official results are reported to the public to promote transparency and public confidence in national assessments.

- The 2024 NECO SSCE Mathematics test, though generally reliable and fair, can still be improved by addressing random inconsistencies and residual biases. Through the regular psychometric analyses to estimate the reliability and error structure of test items before and after administration; the margin of error can be reduced. The implementation of regular psychometric audits in assessment should be enhanced as a policy.
- Teachers should provide students with adequate exposure to standardized testing conditions and develop classroom assessments aligned with NECO standards to minimize performance fluctuations in summative assessment.

REFERENCES

- Adeosun, P. K., & Ukanah, I. A. (2024). Estimating Standard Error of Measurement in BECE 2022 Basic Science Multiple Choice Test Items “in Etsako West Local Government Area of Edo State” Using Bootstrap Methods. *NIU Journal of Social Sciences*, 10(2), 175-180.
- Alruwais, N., & Zakariah, M. (2023). Evaluating student knowledge assessment using machine learning techniques. *Sustainability*, 15(7), 6229.
- Ajayi, A. O., & Fakorede, A. O. (2020). Impact of test administration practices on students’ performance in standardized assessments. *Journal of Educational Practice and Policy*, 5(2), 42–50.
- Bassey, A. E., Ovat, A. O., & Ofem, E. E. (2021). Systematic error in measurement: Ethical implications in learners’ assessment in the Nigerian educational system. *Pan-African Journal of Education*, 2(1), 45–56. <https://openaccessglobal.com>
- Batubo, L. O., & Goodness, W. O. (2022). Psychometric analysis of 2018–2020 Basic Education Certificate Examination mathematics objective test items. *East African Scholars Journal of Psychology and Behavioural Sciences*, 4(1), 33–39.
- Blankman, R. (2024). Math Summative Assessment: Explanation and Examples.
- Daramola, C. A., & Akinkuotu, J. A. (2022). Enhancing item validity through item analysis: A study of West African mathematics assessments. *Journal of Educational Development*, 11(3), 85–98.
- Esomonu, N. P.-M., & Okeaba, J. U. (2021). Analysis of hidden sources of error in NECO Mathematics examination. *European Journal of Education Studies*, 8(10), 145–157. <https://www.ej-edu.org>
- Imasuen, K., & Adeosun, P. K. (2023). Estimating sources of measurement error in WAEC Mathematics objective test using generalizability theory. *International Journal of Psychological and Behavioral Sciences*, 8(2), 1–12. <https://www.sciencepg.com>
- Lane, R. (2018). Explainer: what’s the difference between formative and summative assessment in schools?
- Macy, M., Thorndike-Christ, T., & Lin, Y. C. (2010). Parental reports of perceived assessment utility: A comparison of authentic and conventional approaches. *Infants & Young Children*, 23(4), 286-302.

- National Examination Council (2024). *2024 SSCE External Results*. NECO website
- National Examinations Council. (2024). 2024 SSCE External Results.
- Ngunjiri, M. (2022). The role of assessment in mathematics classrooms: A review. *International Journal of Advanced Research*, 5(1), 156-160. <https://doi.org/10.37284/ijar.5.1.887>
- Ohiozua & Omorogiuwa (2025). DETERMINATION OF ABERRANT Examinees' responses Of 2018 and 2019 May/June Biology Objective Questions Of National Business And Technical Certificate Examination. *Journal of Evaluation*, 10(1), 185-199.
- Oyetunde, K. C. (2022). *Application of generalizability theory in estimating measurement error in NECO Mathematics examination in Osun State*. Obafemi Awolowo University Repository. <https://ir.oauife.edu.ng>
- Shavelson, R. J., Baxter, G. P., & Pine, J. (2019). Performance assessments: Reform, pitfalls,
- Yashim, A. U., & Mhab, L. C. (2021). SKEPTICISM IN USING STUDENTS' TEST SCORES TO APPRAISE TEACHERS' EFFICIENCY. *Journal of Educational Theory and Practice (JETP) Online Journal: Vol, 1(1) a State*.