

Development and Validation of Mathematics Achievement Test based on Gender of Middle Basic Pupils in Akwa Ibom State

Ibanga, Magdalene Antyhony

Ovute, Alphonsus Odo

Eluwa, Idowu Obukohwo

Department of Science Education,

Michael Okpara University of Agriculture, Umudike

Email: ibangamagdalene@gmail.com, Email: ovutealphonsus@gmail.com,

Email: eluwa.idowu@mouau.edu.ng

ABSTRACT

The development and validation of a mathematics achievement test in the middle basic schools in Akwa Ibom State aims to estimate the item parameters (content validity (CVI), difficulty indices, discrimination indices and the reliability index) of the Mathematics Achievement Test (MAT) for basic 6 pupils and ascertain the conceptual understanding of male and female basic 6 pupils in mathematics using MAT in Akwa Ibom State, Nigeria. The study sample comprised 200 basic 6 pupils (102 females and 98 males), selected using a multistage sampling technique. Two research hypotheses were formulated and tested at a 0.05 level of significance. The instrument for data collection was a researcher-developed, 50-item, 4-option, multiple-choice, structured diagnostic test titled: Mathematics Achievement Test (MAT), with a reliability coefficient of 0.79, determined using the Kuder Richardson 20 formula. Data were analysed using mean, standard deviation and t-test to test the hypotheses. The findings showed that with a content validity index (CVI) of 0.80, difficulty indices of the items ranging from 38.0% to 67.0%, discrimination indices of the items ranging between 0.35 and 0.64 and distractor indices ranging between 0.00 and 0.08, the MAT is a valid and reliable instrument for measuring achievements in mathematics. Hence, it was recommended, among other things, that teachers should always measure the educational achievements of pupils and ensure that the items are valid and relevant to the course objectives.

Keywords: *Mathematics achievement test, difficulty index, discrimination index, distractor index, reliability*

INTRODUCTION

Education in Nigeria has undergone a long historical evolution, shaped by indigenous traditions, colonial influences, missionary interventions and post-independence reforms. Understanding this trajectory is crucial for appreciating the current challenges and opportunities in primary school education, especially as they relate to the development and validation of diagnostic tests in mathematics. Akwa Ibom State embraced UBE policies, with the State government declaring free and compulsory education at the primary and secondary levels in 2008. However, challenges of quality teacher training and infrastructure persist (Udofia & Udo, 2017). UNESCO (2021) estimates that Nigeria has one of the highest out-of-school populations globally, with over 10 million children out of school. The federal government affirms mathematics as a core subject at the basic level; emphasis on assessment for learning, teacher quality, language of instruction (mother tongue in early primary), inclusive education, and quality assurance (FRN, 2013). It also produced national school evaluation frameworks and standards for basic education, covering curriculum delivery, assessment, teacher competence and school environment. The Nigerian Educational Research and Development Council (NERDC, 2017) listed domains to be included in the primary school mathematics: number and numeration, basic operations (addition, subtraction, multiplication, and division), fractions and decimals, measurement and geometry, data handling and simple statistics and problem-solving skills in real-life contexts.

Primary education empowers both boys and girls by giving them access to knowledge and self-confidence. In Akwa Ibom State, challenges of rural poverty, cultural attitudes and child labour affect enrolment and retention, particularly among fishing and farming communities (Etim & Udoh, 2016). For Akwa Ibom State, success in primary education will determine the country's overall educational success and its ability to compete globally in the knowledge economy. In primary education, mathematics achievement testing has become critical in identifying learning gaps, guiding instruction and improving achievement. The primary concern of testing in the teaching and learning process is that the measurements derived from tests will help make decisions. Testing is a fundamental part of the teaching and learning process used not only as a basis for ranking students at the end of the teaching - learning process but to guide teaching and aid in the development of curriculum, as well as in the assessment of needs, learning difficulties, level of mastery and differences among students (Onah and Jiwueze 2020). Persistent difficulties in key mathematical concepts such as place value, regrouping, fractions, magnitude, area/volume, reasoning and interpretation of graphs have continued to characterise the learning outcomes. These challenges are largely rooted in weak foundational numeracy and ineffective instructional approaches at the basic education level. Empirical evidence from Akwa Ibom State shows that Primary Six pupils experience significant difficulties in basic numeration, which includes place value

and number relationships. Okri, Esuong and Standhope (2022) found that pupils in Uyo metropolis demonstrated low mastery of numeration concepts before instructional intervention, indicating persistent gaps in understanding foundational mathematics at the upper primary level. Pupils' performance in mathematics is strongly influenced by instructional methods, with traditional teaching approaches contributing to poor conceptual grasp and retention (Akpan, Utibe, Babayemi and Nelson, 2023). UNESCO (2015) also stated that classroom assessment relies heavily on end-of-term tests and few teachers use validated diagnostics instrument or item analysis routinely.

As such, education is a fundamental tool when answering or solving the different problems confronting any nation (Akobi, Ezugwu, Madu, Foluke and Asongo, 2021). Akobi *et al.* (2021) opined that there is a need to ensure quality education through valid assessment tools. Therefore, the quality of education in any given nation or country can predict the progress or attainment of that country. Nigeria, in its National Policy on Education, adopted education as an instrument 'par excellence' for affecting national development and harnessing the potentials of the citizens (Federal Republic of Nigeria, FRN 2014).

Education is widely recognised by Nigerian scholars as the central instrument for national transformation and sustainable development. It serves as a strategic framework through which societies transmit values, culture, skills and innovations across generations while also driving socio-economic and political advancement. In addition, Nwachukwu and Azuka (2024) highlight that education goes beyond cognitive development to include character formation and moral values, which are essential for societal stability and continuity. They assert that education ensures the intergenerational transmission of cultural heritage and ethical standards necessary for national cohesion and development.

Therefore, if Nigeria is to maintain a high standard of education through high-quality assessment, all the dimensions of learning - cognitive (knowledge), affective (attitude), and psychomotor (practical/motor skills) domains must be assessed adequately in any examination (teacher-made or standardised examination). The three educational objective domains are interrelated. Among primary school subjects, mathematics has been at the centre of reform and its curriculum is regarded as a very essential aspect in mathematics education (Awofala, 2017). The benefit of introducing mathematics at all levels of the school system, including the College of Education, is to promote in students a deeper understanding of the physical, space/geometric figures (Gada, 2024), strengthen the arithmetic programme to promote more fundamental development of the nature of measurement and measuring process, and encourage imaginative thinking in science and technology. The development and validation of a mathematics achievement test for primary school education is an essential step towards improving the quality of mathematics education.

The primary reasons of Nigerian students having a poor score or lack of interest in mathematics are low quality teacher made tests, the inability of instructors to design quality items for evaluation and learners' lack of foundation skills (Wanlor, Wakjissa and Mustapha 2025a). These contribute to students' poor performance (Dadughun, 2015; WAEC 2019-2023). Wanlor, Dalong and Olakunle (2023) advised researchers, teachers and assessment specialists to develop geometry diagnostic tests to assist learners in remedying their difficulties and misconceptions. These evaluations help teachers tailor their instruction, develop individualised lesson plans and provide remedial education by identifying learners' strengths and shortcomings (Nuraini, Chulifah and Laksono 2019). The development of diagnostic test entails producing top-notch items to evaluate learners' proficiency (Mawak, 2019). Examiners and other test users are usually interested in the results yielded by the test administration. They are not conversant with the characteristics or technical features of the test and test items (Ohiri, 2023). He further stated that many persons using test results do not realise that the usefulness and appropriateness of test score interpretation is a direct result of the test's internal characteristics. The psychometric properties or characteristics are the internal attributes of a test and its items.

A diagnostic test is considered one-dimensional if it assesses only one trait and multidimensional if it evaluates multiple traits. One-dimensional items provide dichotomously scored data that are calibrated using the 1, 2, 3 and 4 parametric logistic model (PLM) while multidimensional items produce polytomous (multiple response) scored data that are calibrated using the partial credit model (PCM) or Graded Response Model (GRM) (Columbia Public Health, 2020).

The item response theory (IRT) provides a test information function as an alternative to classical test theory (CTT) reliability and provides detailed information on the accuracy level for various ability levels. By using properly selected items, psychometricians can precisely design the degree of reliability information for various ranges of skill (Ayanwale, 2021). Eleje, Nkedi, Esomonu, Koye, Obasia and Onah (2016) found good item difficulties with decimating indices ranging from 0.22 to 0.65 and suitable difficulty ranging from 0.24 to 0.79. Dadughun (2015) found that the primary school Mathematics Diagnostic Achievement Test (PRISMADAT) was reliable and one-dimensional with items difficulty and discrimination parameters ranging from -0.97 to 3.21 and -0.29 to 4.95, respectively.

Consequently, Attah, Ngachu, Amuche and Obiji (2025) emphasized that many teacher-made and standardized test items in Nigeria fail to meet acceptable psychometric standards, particularly in terms of discrimination power and appropriate difficulty levels. This further supports the argument that instability in test items is a systemic issue affecting the credibility of qualifying examinations. Overall, these findings indicate that qualifying examination items in Nigeria often exhibit poor stability due to fluctuating difficulty levels and weak discrimination indices, which undermine their effectiveness in

accurately assessing learners' abilities. Furthermore, there is a strong need for rigorous item analysis, test calibration and continuous validation to ensure the development of stable and reliable assessment instruments. The internal consistency of the instrument was determined using the Kuder–Richardson (KR-20) formula, and a reliability coefficient of 0.70 and above was considered acceptable, indicating that the instrument was reliable (Oloyede and Adeyemi, 2023; Adebule and Ayoola, 2020). Reliability refers to the consistency of test scores across different administrations of the test. The reliability of internal consistency of the test estimated using the Kuder-Richardson formula 20 (KR-20) because items were dichotomously scored, though items had multiple options, there is only one key that is the correct answer and no point for incorrect answer.

Furthermore, the Rasch model of test standardisation, developed by George Rasch, is a psychometric framework used to analyse categorical data such as test scores and questionnaire responses, focusing on the interaction between a person's ability or trait and item difficulty (Istiqomah, Hasanati and Nida, 2022). This model is particularly effective in educational research for evaluating multiple-choice or binary-scored items, where each response is classified as correct or incorrect (Abe, Gbore, Owabumoye and Omotoyinbo, 2025). The model operates on the premise that a person's response is determined by both a person – specific and item-specific parameter, which remain consistent across different interactions (Owabumoye, 2024). The model is described as a simple trait model, asserting that the probability of a correct response depends solely on the interaction between the individual's ability and the item's difficulty (Alonge & Gbore, 2016). As such it offers a robust framework for test standardisation with the potential to mitigate test anxiety and improve students' engagement and attitudes towards subjects like Mathematics. The model posits that the probability of a correct response increases with a person's ability and decreases with item difficulty (Abe et al 2025). They asserted that an important advancement in test theory would be the identification of item parameters that remain stable across varying analysis groups.

Consequently, educational assessment means a systematic process of gathering data from a variety of sources to understand, describe and improve learning (Onah and Jiwueze, 2020). The various sources of gathering educational data include, but are not limited to, classroom tests, assignments, projects, home works, examinations, etc. According to the author, this concept of educational assessment implies that it encompasses both measurement and evaluation; decisions or judgments are made based not only on one single measurement or on a source of data, but multiple measurements or sources of data. Ugodulunwa (2020), has it that evaluation in education is the process of making value judgement over a level of achievement or performance or level of implementation of a programme. In education, assessment is an important part or component of teaching and learning. With these, it shows that assessment is the systematic way or method of gathering and synthesising information about pupils' progress. It is also a method used for analysing and evaluating the achievement of pupils

or the success of a programme. Assessment can be either formative or summative. From Nworgu (2015), formative assessment is undertaken while a programme, lesson or course is still in progress or on going to collect relevant data and use the feedback to improve the learning/ course of the programme. On the other hand, summative assessment or assessment of learning is carried out to determine what pupils have been able to learn at the end of a given lesson, unit programme or period of schooling. The most popularly used technique for obtaining information in the school or educational system is the test. The test enables teachers to systematically obtain data for making comparisons across individuals, classes, schools, districts or countries (Akobi *et al.*, 2021). Nworgu (2015) stated that a test refers to a structural situation comprising a set of items (i.e. questions or statements) with preferred responses given to individuals, learners or testees to determine the amount of the relevant trait or attribute they possess.

In addition, Udofia and Udoh (2017) assert that tests should involve both quantitative and qualitative description of a pupil or testee's behaviour, the passing of value judgment concerning the desirability of that behaviour. So, examination agencies were set up to promote education, coordinate educational programmes and monitor the quality of education in educational institutions. Since examination is the basis for comparison of the achievement of the National Policy of Education, the psychometric properties of the test items must be of an acceptable standard (Ogunbamowo, Adediwura and Diyan, 2019).

This study aims at addressing the gap by developing and validating a basic middle school mathematics achievement test in Akwa Ibom State.

Hypotheses

The following hypotheses were formulated to guide this study and tested at 0.05 level of significance

H₀₁: There is no significant difference between the mean mathematics achievement scores of male and female Basic 6 pupils assessed using researcher-developed Mathematics Achievement Test.

H₀₂: There is no significant difference between the mean mathematics achievement scores of male and female Basic 6 pupils assessed using the researcher-developed Mathematics Achievement Test.

METHOD

This research adopted the survey design. The area of study is Akwa Ibom State, Nigeria. Akwa Ibom State is in the South-South geopolitical zone of Nigeria. The target population for the study was primary school pupils in the middle basic 6 in public primary schools in Akwa Ibom State. The sample for this study comprised 200 Basic 6 pupils (102 females and 98 males) selected from 2 public primary schools and 2 private primary

schools. The breakdown includes 66 male pupils, 32 male pupils, 53 female pupils and 49 female pupils. The sampling technique used was a multistage sampling technique. The instrument for data collection was a researcher-developed, 50-item, 4-option, multiple-choice, structured diagnostic test titled: Mathematics Achievement Test (MAT), drawn from the government-approved scheme for Mathematics in Basic 6. Each correct answer on the Mathematics Achievement Test (MAT) was scored 1 mark, and an incorrect answer was scored zero. This gave a maximum score of 50 and a minimum score of zero. This involved both face and content validity and item analysis. Validity of a research instrument is the ability of an instrument to measure what it is designed to measure.

To ensure face and content validity, the instruments, MAT, were submitted to three independent experts, two content experts in Mathematics education and one measurement and evaluation expert, all of Michael Okpara University of Agriculture, Umudike. To further strengthen the validity of the research instruments, content validity was conducted. A test or an instrument should adequately cover the domain of behaviour it intends to measure to the extent possible. Hence, to ensure content validity the items were drawn from basic concepts in Basic 6 Mathematics curriculum using the Table of Specification for Mathematics Achievement Test (MAT). The content validity index (CVI) of the instrument was computed based on the joint ratings of relevance of MAT items in a 4-point rating scale (4 = very relevant, 3 = quite relevant, 2 = somewhat relevant, and 1 = not relevant) by two content specialists using the relationship.

The reliability of the MAT was established using the Kuder Richardson formula - 20 (KR-20 was 0.79). The data collected, after the administration of the validated instrument, from the respondents were answered using mean and standard deviation to provide answers to the research question, while an independent t-test was used to test the hypotheses at a 0.05 level of significance. The decision to accept or reject the null hypotheses stated was based on the following rules. When the calculated p-value was greater than the 0.05 level of significance, the hypothesis of no significant difference was upheld. When the calculated p-value was less than the 0.05 significance level, the hypothesis of no significant difference was rejected.

RESULTS AND DISCUSSION

Table 1: Summary of Mean and Standard Deviation scores of pupils on mathematics achievement classified by gender (N=200)

Gender	N	\bar{X}	Std. Deviation
Male	98	23.20	7.24
Female	102	30.66	8.42

Table 1 shows the mean mathematics achievement and standard deviation scores of male and female basic 6 pupils on the concept of areas and perimeter, number and numeration, basic operation and fractions assessed using a researcher-developed mathematics achievement Test. The mean and standard deviation scores of the male and female basic 6 pupils displayed are 23.20 and 7.24, and 30.66 and 8.42, respectively. These results show that the female basic 6 pupils outperformed their male counterparts.

Table 2: Summary of Independent t-test analysis of Basic 6 pupils' scores classified by gender (N= 200)

Gender	N	\bar{X}	Std. Deviation	df	t	Sig.
Male	98	23.20	7.24		198	-6.70
Female	102	30.66	8.42			0.00

The independent t-test analysis of Basic 6 pupils' scores, classified by gender, in Table 2 shows a significant level of 0.00, which is less than 0.05 alpha, and a calculated t-value of -6.70, indicating that there was a significant difference between the mean mathematics achievement scores of male and female Basic 6 pupils in favour of the females. Hence, the null hypothesis was rejected.

Table 3: Summary of Mean and Standard Deviation scores of pupils on mathematics achievement classified by gender (N=200)

Gender	N	\bar{X}	Std. Deviation
Male	119	26.99	9.43
Female	81	27.02	7.52

Table 3 shows the mean mathematics achievement and standard deviation scores of male and female Basic 6 pupils on the concept of area and perimeter, number and numeration, basic operation, and fraction assessed using the researcher-developed Mathematics Achievement Test. The mean and standard deviation scores of the male and female Basic 6 pupils displayed are 26.99 and 9.43, and 27.02 and 7.52, respectively. These mean scores show that the female Basic 6 pupils outperformed their male counterparts. The standard deviation scores show that the raw scores of the female pupils were closer to the mean than those of their male counterparts.

Table 4: Summary of Independent t-test analysis of Basic 6 pupils' scores classified by gender (N= 200)

Gender	N	\bar{X}	Std. Deviation	df	t	Sig.
Male	119	26.99	9.43	198	-0.03	0.98
Female	81	27.02	7.52			

The independent t-test analysis of Basic 6 pupils' scores classified by gender in Table 4 shows a non-significant level of 0.98, which is greater than the 0.05 alpha, and a calculated t-value of -0.03, indicating that there was no significant difference between the mean mathematics achievement scores of male and female Basic 6 pupils. Hence, the null hypothesis was upheld.

CONCLUSION AND RECOMMENDATIONS

Based on the findings of the study, it is hereby concluded that the researcher-developed Mathematics Achievement Test (MAT) is a valid and reliable instrument for measuring achievements in mathematics. It is useful as a diagnostic tool for determining learners' areas of weakness in middle basic mathematics. Based on the findings, the following recommendations were made:

- Teachers should measure the educational achievements of pupils to ensure that the items are valid and relevant to the course objective.
- Teachers, researchers and relevant educational agencies should ensure that the instrument they develop for measuring achievements is reliable, as instruments for measuring achievement are used in the decision-making of achievement of pupils.
- Teachers, researchers and relevant agencies should always carry out item analysis to determine the difficulty and distractor indices of each of the items, especially when carrying out achievement tests. This is recommended because items that are too difficult or too easy are not effective for measuring achievements; rather, items of average difficulty are more appropriate.
- Seminars should be organised for mathematics teachers on how to develop, validate and use diagnostic mathematics tests in the classrooms.

REFERENCES

- Abe, T. O., Gbore, L. O., Owabumoye, O. S. & Omatoyinbro, D. W. (2025). Verification of Rasch Model in evaluation test Anxiety and Attitude of Students towards learning of Mathematics. *International Journal of Mathematics and Statistics studies* 13(2), 21-33. Publication of the European centre for Research training and development – UK.
- Adebule, S. O., & Ayoola, O. O. (2020). Test construction and validation practices in Nigerian educational research. *Journal of Educational Foundations*, 10(1), 112–125.
- Akobi, T. O., Ezugwu, I. J., Madu, B. C., Foluke, B. E. & Asongo, S. T. (2021). Analysis of Psychometric Qualities of National Examinations Council (NECO) Mathematics

Essay Test Using Generalized *International Journal of Humanities and Social Studies* Partial Credit model. The www.theijhess.com.

- Akpan, I. I., Utibe, U. J., Babayemi, J. O., & Nelson, M. F. (2023). Effect of geo-board and charts on students' achievement in mathematics in Essien Udim, Akwa Ibom State, Nigeria. *Journal of Research in Education and Society*.
- Alonge, M. F. & G, L. O. (2016). The concepts of probability and probabilistic models of item analysis. In *Concurrent Issues in Education*. Lagos: Premier Publishers.
- Awofala, A. O. (2017). Improving mathematics teaching in Nigeria through diagnostic assessment. *International Journal of Education, Learning and Development*, 5(7), 1–12.
- Attah, F., Ngachu, B., Amuche, C. I., & Obiji, D. (2025). Assessment of psychometric properties of economics teacher-made tests using item response theory in Jalingo Education Zone, Taraba State, Nigeria. *International Journal of Innovative Education Research*, 13(4), 296–310.
- Ayanwale, M. A. (2021). Calibration of Polytomous Responses Mathematics Achievement Test using generalized partial credit model. *British Journal of Education*, 5(2). 21-41.
- Columber Public Health (2020). Item response theory <https://www.publichealth.columbia.edu/research/health-methods>. Columbia University. (2025). Research Instrument Examples. Teachers College.
- Dadughun, S. I. (2015). Development and calibration of a primary school Mathematics diagnostic test based on item response theory (Doctoral thesis, University of Nigeria, Nsukka, Nigeria).
- Eleje, L. I., Nkedi, P., Esomonu, M. & Koye, R. O. (2016). Development and validation of diagnostic economic test for secondary schools. Retrieved from <https://www.researchgate.net/publication/10.5430/wje.v6n3p9>.
- Etim, J., & Udoh, E. (2016). Determinants of pupils' dropout in primary schools in Akwa Ibom State. *Journal of Educational Studies*, 22(1), 45–58.
- Federal Republic of Nigeria. (2013). *National Policy Education (6th ed.)*. Lagos National Educational Research council NERC Press.
- Federal Republic of Nigeria. (2014). *National Policy on Education*. Abuja: National Educational Research Council NERC Press.
- Federal Republic of Nigeria. (2014). *National policy on education (6th ed.)*. NERDC Press.
- Gada, B. A. (2024). Development and Validation of the geometry achievement test for colleges of education students. *International journal of computer science and Mathematical Theory (IJCSMT)*.
- Istiqomal, I., Hasant, I. & Nida, N. (2022). Development of students' academic performance determinants using ranch model analysis. *Psymphatic Junal Ilmiah Psikologi*, 9(1), 17-30.
-

- Mawak J. J. (2019). Development and Calibration of economics achievement test for secondary school students in Plateau State Nigeria, University of Jos Nigeria.
- NERDC. (2017). National assessment of learning achievements in basic education. Abuja: NERDC Press.
- Nuraini, P. S., Chulifah, N. L. S. & Laksono, W. C. (2019). *Mathematics errors in elementary schools: A meta-synthesis study advances in social science, education and humanities research (ASSEHR)*, 1st International Conference on early childhood and primary education, 24(4), 244-259 Retrieved from <https://www.researchgate.net/pm>.
- Nwachukwu, E. L., & Azuka, C. V. (2024). Education and character reformation in Nigeria contemporary society: Challenges and prospects. *Journal of Digital Learning and Distance Education*, 2(12).
- Nworgu, B. G. (2015) Educational Research: Basic issues and methodology (3rd Ed. Nsukat: University Trust Publisher.
- Ogunbamowo, A. O., Adediwura, A. A. & Diyan, R. O. (2019). Psychometric Properties of 2017 West African Examinations Council and National Examinations Council's Economic Senior School Certificate Examination Items. *Journal of Advances in Education and Philosophy*. <http://saudijournals.com/jaep>.
- Ohiri, S. C. (2023). Psychometric Analysis at Item Level of WAEC May/June Mathematics Multiple Choice Questions Using the Item Response Theory. *Journal: Unizik Journal of Educational Research and Policy Studies*, 16(5). DOI: 10.13140/RG.2.2.15731.94247.
- Okri, J. A., Esuong, U. U., & Standhope, E. (2022). Mathematics scaffolds learning strategy and primary school pupils' knowledge of basic numeration in Uyo metropolis of Akwa Ibom State. *International Journal of Studies in Education*.
- Oloyede, O. I., & Adeyemi, B. A. (2023). Assessment of internal consistency of objective test items using Kuder–Richardson approach in Nigerian secondary schools. *Nigerian Journal of Educational Evaluation*, 15(2), 45–58.
- Onah, F. E. & Jiweze, C. C. (2020). Psychometric properties of West African Examination Council (WAEC) O'Level multiple choice past question papers on economics for the year 2013. *Journal of Science Computer Education (JOSCED)*.
- Owabumoye, S.O. (2024). Application of Probability Models on Evaluating Test Anxiety Attitude toward Mathematics Learning in Secondary School in Ondo State (Unpublished PhD thesis). Adekunle Ajasin University, Akungba-Akoko, Ondo State.
- Primi, T. D. N. & Primi, R. (2014). Rasch-Master's partial credit model in the assessment of children's creativity in drawings. *The Spanish Journal of Psychology*. 17(35),1-16.
- Udofia, I., & Udo, E. (2017). Educational reforms in Akwa Ibom State: Challenges and prospects. *Journal of Educational Research and Development*, 12(3), 45–56.
-

- Ugodulunwa, C. A. (2020). Fundamentals of Educational Measurement and Evaluation 2nd Ed. Fab Anieh Nig. Ltd. Jos, Plateau State.
- UNESCO (2015). Education 2030: Incheon Declaration and Framework for Action. UNESCO.
- UNESCO (2021). Global education monitoring report 2021/2: Non-state actors in education. UNESCO Publishing.
- Wanlor, S. B., Dalong, O. M. & Olakunle, F. J. (2023). Perception of geometric concepts learnt by upper basic school students to enhance quality education in plateau State, *Nigeria Journal of Educational Research and Evaluation*.
- Wanlor, S. B., Wakjissa, S. G., & Mustapha, Y. A. (2025). Development and Calibration of Geometry Diagnostic Test for Upper Basic Education Students in Plateau State, Nigeria. *International Journal of Research and Innovation in Social Science*. 22(22) 238-148.
- West Africa Examination Council (WAEC 2019-2023). Chief examiners annual reports. <https://www.waecgh.org/examiner>.