ISSN: 1118-5872

FOS

MULTI-DISCIPLINARY JOURNAL

(Alvan Ikoku Federal University of Education)

CONTENT VALIDITY AND RELIABILITY OF MATHEMATICS TESTS FOR BASIC EDUCATION CERTIFICATE EXAMINATION IN ABIA STATE, NIGERIA

[1]Ike, Iheanacho C, [2]Alwell, Uzoma., [3]Onyegbuchelem, Chialuka A., & [4]Uwakwe, Joy I.
[1,2,3,4] Department of Mathematics, AIFUE, Owerri Imo State. ikeachor5@gmail.com

Abstract

This research assessed the Content Validity and Reliability of Mathematics Tests for the Basic Education Certificate Examination (BECE) in Abia state, Nigeria. The study was guided by three research questions and two null hypotheses, employing an evaluation research design. The study population consisted of 240 items (also served as the sample size), utilizing BECE mathematics tests from 2019 to 2022 as research instruments. Percentage was used to answer research questions one and two and Kuder-Richardson 20 was used to answer research question three, chi-square goodness of fit test was employed to test the null hypotheses at a 0.05 significance level. Results indicated that content percentages in the tests aligned significantly with those in the Mathematics Curriculum, but the objectives' percentages did not. The instruments exhibited reliability indices ranging from 0.9338 to 0.9790. Consequently, it was concluded that while the constructed mathematics tests were reliable, they lacked sufficient content validity. The researchers recommended employing a well-developed test blueprint to enhance content validity during test construction.

Keywords*; Mathematics, Tests, Content Validity, Reliability, BECE, and Test Blue Print.*

Introduction

The education system has gone through tremendous changes and innovations over the years. These innovations are targeted towards achieving educational goals and solving educational problems. Improvement of any country may be a degree of advancement within the zone of science and innovation, within the world nowadays, science and innovation have ended up a prevailing control advancement marker. Mathematical science may be a centre science subject, at the essential and auxiliary schools that have found expression within the social, political, logical, and innovative achievement (Unodiaku, 2012). The common approach in mathematical science is to build a mathematical model of a phenomenon, solve the model, and develop recommendations for performance improvement (Elaine & Gordon, 2021). Mathematics can be overwhelming for some, but with artificial intelligence's (AI's) personalization powers, it can become an accessible and fun topic to learn for all. According to Nirmal (2023), AI enhances mathematics education by; explaining a concept or a problem, personalizing problems based on interest, generating project ideas on a topic, generating structured lesson plans for teachers, generating worked examples for students, converting abstract problems into word problems and providing feedback. In this study, the AI tool was explored in assessing the cognitive objectives of the items in the mathematics tests.

Mathematics provides an effective way of building mental discipline and encourages logical reasoning and mental rigor (Park, Brombacher, Brocardo, & Steen, 2017). In addition, mathematical knowledge plays a pivotal role in grasping the contents of other school subjects such as science, social studies, music, and art. Mathematics has provided the mental discipline required for other disciplines, mathematical literacy (basic computational skills, quantitative reasoning, spatial ability, and others.) plays a vital role in empowering individuals to lead more productive lives as engaged and thoughtful members of society.

Education in Mathematics at the secondary school level according to Buba and Kojigili (2020), is the level during which the innate potentialities of the students are to be explored, identified, and properly channelled. A student whose mathematical potential is not recognized, nurtured, and cultivated during this stage may miss out on it indefinitely. Once solid foundations are established during this stage, the expectation is for a lasting structure to be built upon them. The importance of secondary education cannot be over-emphasized. In addition to bridging primary and tertiary education, secondary school offers children the chance to gain supplementary knowledge, skills, and desirable attitudes after the primary education level. It provides children with attitudes and abilities that can enable them to acquire tertiary education that is aimed at developing the child for better education beyond primary education by acquiring literacy, numeracy, and communication skills. Many occupations like accountancy, banking, tailoring, carpentry, taxation, insurance, and others, which fulfils the needs of man, can be carried out by the use of mathematics. These agencies depend on mathematics for their successful functioning. Teaching mathematics in schools is the most effective and systematic way to impart the knowledge and skills needed for counting, subtraction, multiplication, and division (Pedagogy of Mathematics, 2019). It is regrettable that despite the importance of mathematics, students' academic achievement, attitude, and interest in mathematics have not improved to a great extent.

Mathematics is a study of concepts such as quantity (numbers), structure, space, and change. In the Basic Education Mathematics Curriculum, there are five themes that are broken into content/ topics. Percentages (Weights) are assigned to each theme based on the number of periods used in teaching the content. According to Thorndike and Hagen as cited in Ogugua, Agah, Ene, Acholonu, Azubuike, Okereke, and Agbo, (2020) the weights of contents should depend on personal judgment as guided by the amount of time devoted to each content during instructions. Hence, it is expected that the Basic Education Certificate Examination (BECE) mathematics tests be constructed in the same manner. The researchers deemed it fit to ascertain the content validity of the test instruments of the Education Development Centre (EDC) as one of the bodies saddled with the responsibility of conducting BECE for the states Ministry of Education in various subjects which Mathematics is one of them.

Test (Achievement tests) is a circumstance in which a learner is required to reproduce what he or she has learned either verbally or in writing or in concrete practical terms to a teacher from which the knowledge, adjudged by the teacher or competent authority as either right or wrong (Nwana, 2007). Achievement test measures an individual's knowledge or skill in a given area or subject (Fraenkel & Wallen, 2013). Achievement tests are often used in educational and training settings. In schools, for example, achievement tests are frequently used to determine the level of education for which students might take such a test to determine if they are ready to pass a particular subject or grade level and move to the next level (Kendra, 2021). The author, also stated that Standardized achievement tests are also used extensively in educational settings to determine if students have met specific learning goals. Each grade level has certain educational expectations and testing is used to determine if schools, teachers, and students meet those standards.

The Federal Republic of Nigeria (FRN) (2013), in her National Policy on Education, introduced a 6-3-3-4 system of education, a six-year duration for secondary education given

in two stages of three years each, that is, the junior and senior secondary respectively. Junior secondary school level three (JSS3) students and senior secondary school level three (SS3) students have separate examination bodies responsible for conducting certification examinations. The Junior Secondary Certificate Examination (JSCE), introduced in 1992 and administered by each state of the federation was renamed the Basic Education Certificate Examination (BECE) in 2009 in line with the 9-3-4 educational system (Buba & Kojigili, 2020). Essentially, the 9-3-4 educational system implies that the first nine (9) years of the education programme is basic and compulsory for all Nigerian children of school age. The nine years start from primary up to the JSS3 level. All these changes are to improve the standard of education in Nigeria.

In this regard, the BECE is a testimonial to ones; ability and aptitude. Because of this, this examination has to be valid, reliable, and useable as an evaluation instrument. According to Nwana (2007), the validity of a test and the reliability of the same test are both interrelated as well as inter-twined and while one is determining the one, one is making assumptions as well as utilizing strategies that are implied in the other. Content validity according to Ukozor, as cited in Ogomaka, Ekwonye, Ukozor, and Onah (2016), is the extent to which the test measures both the subject matter content and the instructional objectives designed for a given course. According to Iwuji, as cited in Ogomaka, et al, (2016) validity refers to the accuracy of a measure, it relates to whether the results do represent what they are supposed to measure, it focuses on the outcome, while reliability refers to the consistency of a measure, it relates to whether results can be reproduced under the same conditions, it focuses on maintaining the consistent result. Omorogiuwa cited in Oribhabor (2016) defined Reliability as the consistency with which a test measures what it sets out to measure. This means that reliability is the dependability or trustworthiness of a test. All valid examination instruments are relatively reliable and useable but not all reliable instruments may be valid, therefore validity takes precedence over the other qualities possessed by evaluation instruments. An examination is not valid and so is the result if the following principles are not considered during construction; questions should be drawn from all parts of the syllabus; the number of questions set from the syllabus sections must reflect the contents (relative importance) according to the syllabus; the questions should test all the intellectual skills (cognitive objectives) expected to result from the course (Nwana, 2007). A test is said to have content validity if it is a representative sample of the contents and objectives in the syllabus. If any section of the curriculum (syllabus) is not represented by the questions, then the test ceases to be adequately relevant; it ceases to be a valid measure of the subject. Suppose that all the sections of the curriculum are covered by questions but that in addition some questions are set on contents outside the curriculum. These additional questions, not prescribed by the syllabus, will reduce the relevance of the whole test. The first principle of relevance in an achievement test, therefore, holds that all the sections of the syllabus, no more, are to be examined.

Content validity is particularly well-suited for objective-type achievement tests. This assessment is facilitated through the use of a test blueprint and the collective judgment of test experts and subject matter experts. Nunnally, as referenced in Abonyi (2011), suggests that a key method to establish content validity is to clearly define the objectives of an instructional course and design examinations that align with those objectives. This can be accomplished by

the creation of a test blueprint (a table of specifications), which serves as a detailed blueprint to guide test developers and ensure that the content is adequately represented in the assessment items. A test blueprint, as described by Nworgu as cited in Ogomaka et al. (2016) and Abonyi (2011), is a comprehensive chart that outlines the relationship between the objectives and the subject matter. In this blueprint, the content areas are listed vertically, while the objectives are listed horizontally.

A comprehensive table of specifications should encompass all six major levels in the cognitive domain as identified by Bloom (1956). Watson (2020) stated that Bloom's taxonomy was designed with six levels to encourage higher-order thinking, which was knowledge, comprehension, application, analysis, synthesis, and evaluation. These levels were later revised by Anderson and Krathwohl in 2001 as remembering, understanding, applying, analysing, evaluating, and creating. For higher-level classes, there should be a greater number of higher-order cognitive question categories. In contrast, for beginners (junior classes), the higher-order categories may be excluded or have fewer questions at this level since they are not expected to acquire such skills at that stage of their academic development. The lower-order thinking skills include remembering and understanding, while the higher-order thinking skills involve applying, analysing, evaluating, and creating (analysis, synthesis, and evaluation). Lower levels entail memorization, while higher levels require understanding and applying that knowledge. In the Nigerian educational system, assessment at the junior secondary school level is based on 60% lower-level cognitive skills (LLCS) and 40% higher-level cognitive skills (HLCS). At the senior secondary level, assessment is based on 40% LLCS and 60% HLCS, while in tertiary education, there is an advocacy for 30% LLCS and 70% HLCS (Ihekwaba & Osuala, 2017).

Statement of the Problem

There has been a general outcry on the fluctuating performance of students in BECE as conducted by the Education Development Centre (EDC). According to Faleye, as cited in Oribhabor & Emafor, (2016) many of the items utilized in the classroom assessments across secondary schools in Nigeria are not validated before their use. In summary, researchers, test users, mathematics educators and educational evaluators, cited in the background of the study, were doubtful or worried about the validity (relevance and reliability) tests constructed by teachers and some examination bodies. In this study in particular, the concern, worries and doubts focused on the validity and reliability of Mathematics tests and the test items constructed and used by Basic Education Certificate Examination Centre of the Ministry of Education of Abia State. Do the tests, over some years, have content (and cognitive) validity? Could the tests have acceptable levels of reliability (in terms of internal consistency)? The answers to these questions prompted this study.

Purpose of the Study

The researcher in this study:

I. ascertained how closely the content of tests constructed by the Abia State's EDC align with the expected content of the Basic Education Mathematics Curriculum.
II. determined the degree to which the cognitive objectives assessed by the Abia State's EDC test items reflect those outlined in the Basic Education Mathematics Curriculum.

III. determined the reliability index of each the Mathematics achievement tests constructed by the Abia state's EDC, from 2019 to 2022.

Research Questions

1. In what content areas do the percentage distributions of test items in the Basic Education Certificate Examination (BECE) differ from the expected distribution in the Basic Education Mathematics Curriculum (BEMC)?
2. In what cognitive level do the percentage distributions of test items in the Basic Education Certificate Examination (BECE) differ from the expected distribution in the Basic Education Mathematics Curriculum (BEMC)?
3. What were the reliability coefficient indices of each BECE mathematics test constructed by Abia state's EDC; from 2019 to 2022?

Hypotheses

Ho_1. There is no significant difference between the percentage distributions of the contents in Abia state's EDC-constructed mathematics tests and the percentage distributions of contents in the basic mathematics curriculum.

Ho_2. There is no significant difference between the percentage distributions of the objectives of Abia state's EDC-constructed mathematics tests and the percentage distributions of the cognitive objectives of the basic mathematics curriculum.

Methodology

This research utilized an evaluation research design to scrutinize the content of the BECE mathematics past question papers from Abia state, covering the period from 2019 to 2022. Stake's evaluation model was chosen for its focus on the relationship between observed outcomes and intended objectives. This model also entails a comparison of these elements to establish criteria to ensure that the procedures adhered to are in line with the prescribed standards. The study was carried out in Abia state, Nigeria. The study population comprised two hundred and forty (240) items from all mathematics examination past question papers for BECE developed by the state's EDC from 2019 to 2022. The sample size for the study included the 240 items in the BECE mathematics question papers spanning four years (2019, 2020, 2021, and 2022), with the census population technique used to determine the sample size. The data collection instrument consisted of mathematics past questions created by Abia state's EDC for BECE mathematics. The validity of the instruments was ensured before their utilization. To establish the reliability indices of the state's EDC-constructed basic mathematics test, the instruments were administered to one hundred and twenty (120) senior secondary level one (SS1) students in private secondary schools in Owerri, Imo state. SS1 students were selected because they have covered junior secondary mathematics content (JSS1 to 3). The Kuder-Richardson 20 formula was employed to calculate the reliability indices for each instrument. Research questions one and two were addressed using descriptive statistics, including frequency counts and percentages. Conversely, research question three was tackled using the Kuder-Richardson-20. Null hypotheses were tested using the Chi-square test at a significance level of 0.05.

Results

Research Question 1

In what content areas do the percentage distributions of test items in the Basic Education Certificate Examination differ from the expected distribution in the Basic Education Mathematics Curriculum?

Ho_1. There is no significant difference between the percentage distributions of the contents in Abia state's EDC-constructed mathematics tests and the percentage distributions of contents in the basic mathematics curriculum.

BECE Curriculum content		2019	2020	2021	2022	TOTAL
Number and Numerations	(12%)	**06**(10%)	**11**(18%)	**09**(15%)	**9**(15%)	**35**(15%)
Basic Operations	(27%)	**14**(23%)	**10**(17%)	**14**(23%)	**13**(22%)	**51(**21%)
Algebraic Operations	(26%)	**14**(23%)	**07**(12%)	**11**(19%)	**13**(22%)	**45(**19%)
Mensuration / Geometry	(24%)	**21**(35%)	**26**(43%)	**17**(28%)	**19**(31%)	**83(**34%)
Statistics	(11%)	**05**(9%)	**06**(10%)	**09**(15%)	**06**(10%)	**26(**11%)
		60	60	60	60	240
$\chi^2_{cal} - value$		6.68	29.37	5.35	4.42	2.25
df		4				
$\chi^2_{crit} - value$		9.49				

Table 1. Presents the percentages of basic mathematics curriculum contents and percentages of Abia state's EDC-constructed mathematics contents from 2019 to 2022. The table shows that in 2019, 2020, 2021, and 2022 number and numeration, has a percentage of 10%, 18%, 15%, 15%, and overall of 15% against 12% assigned to it in the curriculum, the basic operation has 23%, 17%, 23%, 22%, and overall of 21% against 27%, algebraic operations has 23%, 12%, 19%, 22% and overall of 19% against 26%, mensuration/ geometry has 35%, 43%, 28%, 31% and overall of 34% against 24% while statistics has 9%, 10%, 15%, 10% and overall of 11% against 11% respectively.

The result of the Chi-square goodness of fit test, shows that in 2019, 2021, and 2022 the χ^2_{cal}= 6.68, 5.35, and 4.42 respectively and the degree of freedom from the contingency table is (2-1) (5-1) = 4 while the $\chi^2_{crit.}$= 9.49 at 0.05 level of significance. Since$\chi^2_{cal}. < \chi^2_{crit.}$, we accept the null hypothesis that the percentage distributions of the observed fit the percentage distribution of the expected. In 2020, the χ^2_{cal} = 29.37 while the $\chi^2_{crit.}$ = 9.49. $\chi^2_{cal}. > \chi^2_{crit.}$, we reject the null hypothesis. But for the overall percentage of the Abia state EDC-constructed mathematics test

content, χ^2_{cal}. (2.25).< $\chi^2_{crit.}$ (9.49), we accept the null hypothesis that the percentage distributions of the state's constructed contents fit the percentage distribution of the basic education curriculum contents.

Research question 2.

In what cognitive level do the percentage distributions of test items in the Basic Education Certificate Examination differ from the expected distribution in the Basic Education Mathematics Curriculum?

Ho_2. There is no significant difference between the percentage distributions of the objectives of Abia state's EDC-constructed mathematics tests and the percentage distributions of the cognitive objectives of the basic mathematics curriculum

BECE Cognitive objectives		2019	2020	2021	2022	Total
Remembering	25%	11(18%)	10 (17%)	07 (12%)	15 (25%)	43(18%)
Understanding	35%	16 (27%)	18 (30%)	15 (25%)	18 (30%)	67(28%)
Applying	25%	25 (42%)	24 (40%)	25 (42%)	18 (30%)	92(38%)
Analysing	10%	08 (13%)	08 (13%)	12 (20%)	08 (13%)	36(15%)
Evaluating	04%	00 (00%)	00 (00%)	01 (01%)	01 (02%)	02(01%)
Creating	01%	00 (00%)	00 (00%)	00 (00%)	00 (00%)	00(00%)
		60	60	60	60	240
$\chi^2_{cal} - value$		21.25	18.17	34.38	4.21	15.82
df		5				
$\chi^2_{crit} - value$		11.07				

Table 2. Presents the percentages of basic mathematics (expected) cognitive objectives and percentages of Abia state's EDC-constructed (observed) mathematics cognitive objectives from 2019 to 2022. The table shows that in 2019, 2020, 2021, and 2022: remembering, has a percentages of 10%, 18%, 15%, 15%, and overall of 18% against the expected 25%, understanding has 27%, 30%, 25%, 30%, and overall 28% against expected 35%, applying has 42%, 40%, 42%, 30%, and overall of 38% against expected 25%, analysing has 13%, 13%, 20%, 13%, and overall of 15% against 10%, evaluating has 0%, 0%, 01%, 2% and overall of 1% against expected 5% while creating has 0%, 0%, 0%, 0% against 1% respectively. The result of the Chi-square goodness of fit test, it shows that in 2019, 2020, and 2021 the χ^2_{cal}= 21.25, 18.17, and 34.21 respectively while the $\chi^2_{crit.}$ = 11.07. Since$\chi^2_{cal}. > \chi^2_{crit.}$, we fail to accept the null hypothesis and accept the alternative hypothesis. In 2022, the χ^2_{cal}= 4.21 and the $\chi^2_{crit.}$value = 11.07. $\chi^2_{cal}. < \chi^2_{crit.}$, we accept the null hypothesis. But for the overall percentage distribution of the Abia state EDC-constructed mathematics test cognitive objectives, χ^2_{cal} (15.82) > $\chi^2_{crit.}$

(11.07), we fail to accept the null hypothesis and accept the alternative hypothesis that the percentage distributions of the Abia state's EDC-constructed cognitive objectives do not fit significantly with the expected percentage distribution of the cognitive objectives for the basic education curriculum.

Research Question 3

What are the reliability coefficient indices of each BECE mathematics test constructed by Abia state's EDC; from 2019 to 2022?

YEARS/STATES	2019	2020	2021	2022
r	0.9485	0.9790	0.9338	0.9555

Table 5. Shows the summary of each years' reliability coefficients of internal consistency indices of the constructed mathematics tests calculated using Kuder- Richardson 20 ranged from 0.9485 – 0.9555. This indicated that the constructed tests were reliable. Having determined these reliability coefficients of the basic mathematics constructed tests, the last objective of the study has been achieved. Also the research question, 'What are the reliability coefficient indices of each BECE mathematics tests constructed by the state's EDC; from 2019 to 2022? 'Answered.

Discussion of the Findings

The percentage of the content in the Abia state's EDC-constructed BECE mathematics tests.

The outcome in 2020 highlighted discrepancies between the distribution of content percentages observed in the tests and those outlined in the BECE curriculum. The results reveal that, on average, Abia state's Educational Development Centre (EDC) closely adhered to the expected percentage allocations of content in mathematics tests over four years. This indicates a significant correlation between the expected and observed content percentages, emphasizing importance of maintaining consistency in content allocation. These findings contrast with previous studies by Ogugua et al. (2020), Iweka (2008), and Buba and Kojigili (2020), which reported inconsistencies between observed and expected content percentages. In terms of the emphasis placed on content categories, there was a notable correspondence between the expected and observed constructions. However, certain contents were excessively highlighted, while others were neglected or omitted. This suggests that teachers and students may focus their efforts on the emphasized content and pay little to no attention to the neglected or omitted material. This could potentially lead to complacency among teachers and students, as they concentrate on certain areas while disregarding others. As a result, when the omitted content is eventually presented, students may struggle with it. This situation may arise due to a lack of attention from teachers, curriculum developers, and education experts in formulating the necessary content, objectives, and exercises based on the content of textbooks and syllabi.

The percentage of the objectives assessed by items in the tests constructed by Abia state's EDC.

A detailed analysis of the data exposes a notable disparity in the allocation of percentages to objectives. However, the percentages assigned to objectives in the mathematics tests developed by the states' Education Development Centre (EDC) did not significantly align with the expected percentages outlined in the Basic Education Mathematics curriculum. In summary, the observed values did not align significantly with the expected values, a notion supported by the works of Ogugua et al. (2020), Buba et al. (2020), Oribhabor et al. (2016), and Gittinet et al. (2021). The weighting of cognitive objectives did not adhere to a predetermined plan or specification that considers the student's academic level. Bloom (1956), Ihekwaba et al. (2017), and Abonyi (2011) emphasized that cognitive objectives for junior secondary education should consist of 60% lower-level cognitive skills (LLCS) and 40% higher-level cognitive skills (HLCS), yet objectives percentages observed did not correspond to these expected percentages. This implies that neglecting cognitive level validity of tests is leading students in the wrong direction of the syllabus goals, resulting in lower scores in their test results and less development in solid mathematics knowledge, skills, and attitudes. The results of the study show that the items distribution on the (expected) model has little relationship with items distribution on the (observed) constructed test. This affects the motivation of teachers and students.

The reliability coefficient indices of each BECE mathematics test constructed by Abia state's EDC; from 2019 to 2022?

Moreover, the reliability coefficients (internal consistency) across the state during the years under review indicate that the constructed test items are reliable. This aligns with Oribhabor and Emafo (2016) and Buba and Kojigili (2020), which suggest that tests used by mathematics teachers have moderate internal consistency reliability but low content validity. However, ensuring reliability does not guarantee validity. Additionally, it was noted that some items in the EDCs' constructed mathematics tests lack correct answers in the listed options, rendering these items inaccurate. In conclusion, the EDCs' constructed mathematics test instruments were reliable but lacked validity. This implies that the tendency of the fluctuating performances of students will continue if nothing is done.

Conclusion

The research work has been an effort to assess the validity of mathematics tests for BECE in Abia state, Nigeria. In the study, past question papers on mathematics constructed by the state's EDC from 2019 to 2022 were used. From the data analysed and discussed, it was observed that the fluctuating performances of students in mathematics within the period under review might not solely be based on students' factors but could also be attributed to errors in test constructions and development. The percentages (weights) assigned to the contents and cognitive objectives in the constructed tests under review, were not in proportion with the percentages

allotted to the contents and cognitive objectives in the researchers' model. Furthermore, the findings indicate that the constructed tests lacked content validity in terms of cognitive objectives, meaning that the examinations did not assess the required learning objectives and did not consider the students' level.

Recommendation

i. In developing test items, attention should be given to validity and reliability. To develop a mathematics model examination that attains content validity, first, the concerned office in charge should have to prepare a well-developed plan of test that represents the contents and cognitive outcomes of the syllabus appropriately.
ii. Experts in the field of mathematics, and measurement and evaluation should be consulted.

References

Abonyi, O. S. (2011). *Instrumentation in behavioural research:* A practical approach. Enugu: Fulladu Publishing Company.

Anderson, L. W. & Krathwohl, E. (2001). *A taxonomy for learning, teaching, and assessing. A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Ayebale, L., Habaasa, G. & Tweheyo, S. (2020). Factors affecting students' achievement in mathematics in secondary school in developing countries. *Statistical journal of IAOS*. 36(1), 73-76.

Bloom, B. S. (1956). *Taxonomy of educational objectives.* Handbook: the cognitive domain. New York: Mc Kay.

Buba, M. A., & Kojigili, S. T. (2020). Content validation of basic education certificate examination questions in basic science in Adamawa State, Nigeria. *Journal of education, society, and behavioural science*, 33(2), 42-54. https://doi.org/10.9734/JESBS/2020/v33i23020

Education Development Centre, (2022). EDC mission and visions. https://www.edc.org/site-credits. Retrieved November 6, 2022.

Elaine, J.N. & Gordon, J. (2021).*What is mathematics?* https://www.livescience.com Retrieved March 19, 2022.

Federal Republic of Nigeria (2013): National Policy on Education. Abuja NERDC.

Fraenkel, T. R. & Wallen, N. E. (2013). *How to design and evaluate in education research*. (5th ed). Boston: McGraw Hill.

Ihekwaba, C. N. & Osuala, R. C. (2017). Principles of measurement and evaluation in education. Owerri: Zionbless publishers Nig.

Kendra, C. (2021). *How achievement tests measure what people have learned*. www.verywellmind.com. Retrieved April 6, 2022.

National Research Council (1993). *Measuring what counts: a conceptual guide for mathematics assessment.* Washington DC. The National Academic Press. https://doi.org/10.17226/2235

Nirmal, P. (2023). AI tools for mathematics education; use cases and tools to make mathematics fun and engaging. https://www.smartpaperapp.com. Retrieved May 5, 2024

Nwana, O. C. (2007). *Textbook on educational measurement and evaluation*. Owerri: Bomaway publishers.

Ogomaka, P.M.C. (2022). *Toward uniformity in education and research designs and reports*: An Introduction. Owerri: Odesaa Educational Books.

Ogomaka, P.M.C., Ekwonye, E.C., Ukozor, F. I., Ebubechukwu, A. O., Onah, F.E., Dimson, J.N., Ajuonuma, J.O., Uzoma, P.N., Offor, L., & Acholonu, V.N. (2016). *Measurement and evaluation: a comprehensive textbook for students and teachers*. Owerri: Flashpoints publishers.

Ogugua, B. C., Agah, J.J., Acholonu, V.V, Azubuike, R.N., Okeke, M.A., Agbo, P.L. & Ene, C. U. (2020). Content validity of WAEC financial accounting questions. *International journal of learning, teaching, and educational research*. 19(6), 161-178.

Oribhabor, C.B. & Emafo, O. D. (2016). Determining the reliability and content validity of the mathematics tests constructed by senior secondary school mathematics teachers in Edo state, Nigeria. *African journal of education, science and technology.* 3(2): 83-90

Park, K., Brombacher, A., Brocardo, J., & Steen, L. (2017). Role of mathematics in the overall curriculum. *International commission on mathematics instruction.* https//www.mathunion.org/icmi/role-mathematics-overall-curriculum

Pedagogy of Mathematics (2019). *Aims and objectives of teaching mathematics.* Retrieved March 20, 2022

Unodiaku, S.S. (2012) Development and validity of mathematics readiness test for senior secondary school students. *African journal of science, technology and mathematics education*. 2(1), 57-71

Watson, S. (2020). *Higher-order thinking skills in education.* Thoughts.com/higher-order-thinking-skills-hots-education. 3111297