

## VALIDITY AND RELIABILITY OF 2021 KATSINA STATE MATHEMATICS BASIC EDUCATION CERTIFICATE EXAMINATION (BECE) MULTIPLE CHOICE ITEMS

USMAN MU'AZU Ph.D<sup>1</sup>, ABUBAKAR ABDULLAHI Ph.D,<sup>2</sup> & LABIBA HABIB<sup>3</sup>

<sup>1</sup>Nigeria Customs Service Kano/Jigawa Area Command Kano State, Nigeria

<sup>2</sup>Department of Education, Bayero University, Kano Nigeria

<sup>3</sup>Department of Educational Psychology and Counseling,

Federal University Dutsin-Ma, Katsina State, Nigeria

umshagari5@gmail.com; habiblabiba7@gmail.com; a.abdullahi07g@gmail.com

### Abstract

This paper assessed the validity and reliability of 2021 Katsina State BECE Mathematics multiple-choice items. Three objectives and three research questions were guided the study. Descriptive survey design was adopted. The population consisted of all Junior Secondary School three (JSS III) students who sat for the 2021 Mathematics BECE in public junior secondary schools, totaling 68,219 students. A multi-stage sampling technique was used to select a sample of 1,428 students' responses. The collected data was analyzed using MS Excel 2010 to answer the research questions using Lawshe's formula to establish CVRs and CVI, while KR-20 was used to determine internal consistency. The findings revealed that 49(82%) of the items had very high CVRs, and the CVI value was 0.79. It was also found that the 2021 Katsina State Mathematics BECE possessed a high reliability coefficient of 0.83. It was concluded that the 2021 BECE Mathematics multiple-choice items had content validity and were homogeneous. Therefore, it is recommended that test experts should be involved in test construction and validation, and item analysis should be maintained in test development. Also, Katsina State BECE multiple-choice items should be pilot-tested to establish their item statistics.

**Keywords:** Achievement, Psychometrics, Content Validity Ratio, Content Validity Index, Reliability

### Introduction

Achievement tests are designed to measure students' knowledge in specific subjects or skills. They assess the level of mastery or accomplishment that students achieve through their efforts. These tests are commonly used in educational settings such as schools and universities to evaluate the effectiveness of a particular programme of instruction or training. The primary purpose of an achievement test is to determine a student's understanding of a specific subject. They serve as tools to assess students' progress, identify areas for improvement, inform instruction and curriculum development and evaluate the overall effectiveness of educational programs. It is important to note that achievement tests reflect knowledge based on what has been learnt. Achievement tests can be categorized into standardized and non-standardized tests.

According to Sidhu (2012), standardized tests are developed by experts and are administered, scored and interpreted under uniform conditions. In contrast, non-standardized tests are developed by teachers for local use. Typically, achievement tests are conducted after a teaching-learning programme to evaluate the extent to which educational objectives have been met. These tests are critical assessment tools and must be designed to be as objective as possible, ensuring that their results, whether indicating success or failure, are accepted by candidates. The quality of tests is a top priority for test constructors, test-takers, and anyone who relies on test scores for decision-making. Therefore, it is essential that

these tests effectively and efficiently measure learners' abilities.

Basic Education Certificate Examination (BECE) is a crucial assessment for junior secondary school students, evaluating their knowledge in core subjects like English, Mathematics, and Integrated Science, as well as elective subjects. The exam format combines multiple-choice and written questions, with results significantly impacting students' progression to senior secondary education. BECE is an annual examination conducted by the Katsina State Education Resource Centre (ERC) for JSS III students. It aims to develop patriotic citizens with diverse knowledge, skills, and values that promote national consciousness, entrepreneurship, and moral uprightness. Its objectives, as outlined in the National Education Policy (NPE, 2014), include fostering citizens who contribute positively to society, appreciate the dignity of labor, and coexist harmoniously despite differences in abilities, religion, ethnicity, or socio-economic background.

Psychometrics is a discipline that focuses on the theory and practice of psychological measurement. This includes assessing knowledge, abilities, attitudes, personality traits, and educational outcomes (Jones & Thissen, in Gawthrop, 2014). Psychometrics employs standardized tests that measure psychological traits using a mathematical approach. The field involves the development and validation of instruments to ensure they provide reliable and valid measurements. Psychometrics is primarily concerned with the theory of psychological measures, encompassing the assessment of knowledge, abilities, attitudes, personal capacities, traits, and educational measurements. Psychometricians operate on the assumption that evaluation tests are objective, motivating them to create and refine standardized tests. This contrasts with clinometric, which favors clinical expertise over formal statistical models or biometrics (Alkharusi, 2011). Psychometrics often focuses on measuring mental capacities. Over time, it has evolved as a sub-discipline of psychology, becoming a science centered on quantifying individual non-observable characteristics. This includes the design and analysis of measurement instruments and the development of theoretical methodologies for measurement (Aborisade & Fajobi, 2019). There are several psychometric properties among which is validity and reliability which are crucial statistical measures that determine the quality of a test.

Validity refers to how well a test measures what it is designed to measure while reliability indicates the consistency, stability, and dependability of the test in measuring its intended constructs (Anikweze, 2010; Anastasi & Urbina, 2014; Creswell, 2012). Validity of a test is exclusively concerned with what the test measures and how it does so. Therefore, validity evidence of a test could be established through content, criterion, face and construct validation. Content validity simply refers to the extent to which an instrument measures the features for which it is made. When the content of a test is valid, the test items represent the range of possible items that the test should cover. It is the degree to which a test or instrument adequately covers the area of behavior it intends to measure (Sidhu, 2012). Content validity assesses whether the dimensions and components of a concept are appropriately defined (Sekaran & Bougie, 2011). It involves validating each item of the instrument to ensure it accurately represents the construct being measured (Miller & Lovler, 2013). In psychometrics, the validity of a scale reflects its efficacy in assessing a particular construct.

Content validity ratio is a process used to assess an instrument by assessing whether it includes all necessary particulars and removes any that are not applicable to the specific construct being measured. It's an extensively habituated system for determining content validity developed by Lawshe. This approach helps judges or raters agree on the significance of each point in relation to the construct. Lawshe (1975) proposed that each member of the panel should answer a specific question for every point "Is the skill or knowledge covered by this point essential or useful, but not essential, or not necessary for the performance of the construct?" The content validity rate is a numerical score that

reflects the validity of the instrument grounded on the conditions handed by the experts. According to Lawshe, if more than half of the panelists consider a point essential, then it is considered to have a certain position of content validity.

A greater level of content validity exists as a larger number of panelists agree that a particular item is essential. It is a method of measuring mutual agreements between evaluators and judges concerning the importance of a particular item. Likened to other alternative methods in quantifying content validity, such as Cohen's kappa; Tinsley Weiss, T index; and James, Demaree, and Wolf's index; the CVR method is more straightforward, user-friendly, and simple. It also provides a table for finding the critical cutoff value (Wilson et al., 2012). For these reasons, CVR has been widely utilized by local and foreign researchers as the first step in the instrument development process (Doust et al., 2017; Kim, 2011; Matore & Khairani, 2015; Norashady et al., 2016). Lawshe (1975), suggested that each rater should respond to each item with (1) Essential, (2) Useful but not essential, and (3) Not necessary. Based on these assumptions, Lawshe developed a formula known as the content validity ratio formula:

$$CVR = \frac{\left(n_e - \frac{N}{2}\right)}{\left(\frac{N}{2}\right)}$$

$n_e$  = number of panelists indicating essential,

N = total number of panelists.

Content Validity Index is commonly used in quantitative evaluations. The content validity index is a statistic designed to indicate the degree to which a test measures what it is intended to measure. The higher the validity indexes of the test, the greater the validity of the test. A scale with excellent content validity should be composed of CVI of 0.78 and above. Content validity index refers to the mean of the CVR values of items. Content validity index is widely used to estimate the validity of the content by researchers, teachers and test experts. This index does not account for the possibility of inflated values due to the chance agreement. The content validity index is mostly used by researchers, test experts and teachers because it is simple to calculate, easy to understand and provides information about tests (Sangoseni et al., 2013).

Reliability is the degree to which a test is consistent, stable, and dependable in measuring what it is measuring. Reliability of a test is its standard of consistency, stability, dependability, precision, and accuracy. However, a test may give very consistent day-to-day results without measuring what it is intended to measure. This means that a test can be reliable when it is not valid. It is estimated that the scores of a particular test should generate the same or nearly the same when questions are administered multiple times or at different times (Anikweze, 2010; Kubiszyn & Borich, 2013; Sidhu 2012). Internal consistency measures a test's ability to produce reliable results across different items that assess similar concepts. It evaluates whether multiple items intended to measure the same general construct yield comparable scores. The Kuder-Richardson (K-R20) is a method for estimating internal consistency; it calculates the reliability of test scores based on a single administration of one test form using a formula developed by Kuder-Richardson. K-R20 and K-R21 methods provide procedures for assessing whether all items within a test measure the same abilities without being divided into two halves.

Classical Test Theory (CTT) is a foundational framework in educational and psychological measurement that posits an individual's observed score which consists of a true score and measurement error. It emphasizes evaluating and improving test reliability focusing on individual performance and item difficulty. According to CTT, each individual has a true score representing their ability, achievable in a measurement devoid of error. This theory is essential for guiding test development and validation in various assessments (METRIC, 2006).

In Katsina State, Nigeria, the state government, through Education Resource Centre (ERC), conducts Basic Education Certificate Examination (BECE) for all Junior Secondary School three (JSS III) students. Despite the importance of this examination and its widespread popularity particularly in Nigeria, there is no clear evidence to prove that the test questions pass through the assessment of essential qualities such as validity and reliability among others. This clearly indicates that the statistical qualities of the items used for BECE are not known. As such, the items may probably be lacking in one or more psychometric properties which may affect the performance of the students.

It is possible that there are some technical weaknesses in the construction of the test items and probably in their administration. Performance in Mathematics BECE in particular is critical as obtaining a credit serves as a benchmark for decision making with regards to performance of a student. However, passing of Mathematics examination in BECE has been a major challenge to students. Any student who fails to pass Mathematics has to re-sit or repeat JSS 3 class. As a result, the process of learning is to some extent delayed. This study is focused on measuring the validity and reliability of the 2021 Katsina State Mathematics Basic Education Certificate Examination (BECE).

### **Objectives of the Study**

The main objective of this study is to measure the validity and reliability of 2021 Katsina State Mathematics Basic Education Certificate Examination (BECE) Multiple Choice Items. The specific objectives of the study are to find out the:

1. content validity ratios (CVRs) of 2021 Mathematics BECE multiple-choice items
2. content validity index (CVI) of 2021 Mathematics BECE multiple-choice items
3. reliability coefficient of 2021 Mathematics BECE multiple-choice items

### **Research Questions**

1. What are the content validity ratios (CVRs) of 2021 Mathematics BECE multiple-choice items?
2. What is the content validity index (CVI) of the 2021 Mathematics BECE choice items?
3. What is the reliability coefficient of 2021 Mathematics BECE

### **Methodology**

The study adopted a descriptive survey design because it systematically collects data from a large sample, allowing for the identification of patterns, trends, and correlations. The population of the study consisted of all the students who sat for the 2021 Mathematics BECE in public Junior Secondary Schools in Katsina State. In this study, there are two types of populations: students and Mathematics teachers. The population comprised 68,219 JSS III students who sat for the 2021 Mathematics BECE. The population of the experts is ten (10) Mathematics teachers. A total of 1,428 students' responses were drawn from twelve (12) public junior secondary schools using a multi-stage sampling technique. Ten (10) Mathematics teachers were randomly selected from different schools as the experts' ratings for content validation. The 2021 BECE question paper served as the instrument for data collection. Lawshe's formula was used to establish Content Validity Ratio (CVR) and Content Validity Index (CVI) to answer research questions 1 and 2 while Kuder-Richardson (KR-20) was used to determine the internal consistency of the test for research question 3.

### **Results**

**RQ1:** What are the content validity ratios (CVR) of 2021 Mathematics BECE multiple-choice items? This research question was answered by using formula proposed by Lawshe (1975). Each expert panel responded to each item with essential=1 or useful but not essential and not necessary = 0. Content validity ratios (CVRs).

$$CVR = \frac{\left(n_e - \frac{N}{2}\right)}{\left(\frac{N}{2}\right)}$$

$n_e$  = number of panelists indicating essential,

N = total number of panelists.

Table below summarized the CVR or indices of the 2021 Katsina state Mathematics Basic Education Certificate Examination multiple choice items.

**Table 1: Content validity ratio (CVR) of 2021 Katsina state Mathematics Basic Education Certificate Examination multiple choice items**

Items	CVR	Items	CVR	Items	CVR
1	0.8	21	0.8	41	0.8
2	0.8	22	0.8	42	0.8
3	1	23	0.4	43	0.8
4	1	24	0.8	44	0.8
5	0.8	25	1	45	0.8
6	0.8	26	0.6	46	0.8
7	0.8	27	1	47	0.8
8	1	28	0.8	48	0.8
9	1	29	0.8	49	0.8
10	0.8	30	0.6	50	0.8
11	0.6	31	0.8	51	0.2
12	1	32	0.8	52	1
13	1	33	0.8	53	0.4
14	0.8	34	1	54	0.4
15	0.4	35	1	55	1
16	0.6	36	0.4	56	0.8
17	0.8	37	0.8	57	0.8
18	0.6	38	0.8	58	0.8
19	0.8	39	1	59	1
20	0.8	40	1	60	1

Table 1 indicates the analysis of the Content Validity Ratio (CVR). Sixty (60) items were content validated based on the 10 experts judgment using CVR based on the formula proposed by Lawshe (1975). However, all the experts have more than 5 years of teaching experience in the area (i.e. Mathematics) and the experts were chosen from different schools. The results showed that 49 items have a CVR value greater than 0.78 (> 0.78) while 11 items have a CVR value < 0.78 which is in line with (Lawshe, 1975).

**RQ2:** What is the content validity index (CVI) of the 2021 Mathematics BECE?

**Table 2: Content validity Index (CVI) of 2021 Katsina state Mathematics Basic Education Certificate Examination**

Indicator	Value
Items	60
CVI	0.79

The content validity index (CVI) of the 2021 Mathematics BECE multiple-choice items was investigated based on the view of 10 experts in the area of Mathematics. Thus, the CVI value is 0.79 and the results showed that the instrument has a good content validity index based on recommendations given by Songoseni et al., (2013) that a scale with an excellent content validity index ranges from 0.78 and above.

**RQ3:** What is the reliability coefficient KR-20 (internal consistency) of 2021 Mathematics BECE?

**Table 3: Internal Consistency Reliability of 2021 Katsina state Mathematics BECE multiple-choice items Using Kuder-Richardson (KR-20)**

Indicator	Value
No. of Items	60
KR-20 or Reliability coefficient	0.83

Table 3 indicates the reliability coefficient of a dichotomous variable in the 2021 Mathematics BECE multiple-choice items. The coefficient of the internal consistency was reported at 0.83 which is high in line with Brown, (2005) who had recommended an acceptable value of 0.70 to 0.79.

**Discussion**

The aim of the study was to measure the validity and reliability of the 2021 Katsina State Mathematics Basic Education Certificate Examination. The validity and reliability investigated include: Content Validity Ratios (CVRs), Content Validity Index (CVI), and Reliability (Internal consistency). The findings of this study were discussed under each research question.

The finding of the study revealed that the items on the 2021 Katsina state Mathematics Basic Education Certificate Examination (BECE) possessed an acceptable Content Validity Ratio (CVR) within the range of 0.78 and Above in which 49 (82%) of the items in the test are considered to have acceptable CVR value based on the recommendation given by Lawshe (1975). This finding was supported by Gharamani and Khalkali (2015); Ahari et al., (2018); Ramli et al., (2018); Feleye and Ayeloja (2018); Chong et al., (2019); Othman and Mature (2020); Cardenas et al., (2021) where they indicated in their studies that the test items (in terms of CVR) are good because they possessed the acceptable CVR value of greater than 0.78 (> 0.78). The findings of this study differ from those of Adechi et al., (2017), Ghazali et al., (2017), in which their findings showed that items are poor and fail to possess the acceptable range, i.e., CVR value < 0.78.

The findings of the study also revealed that the items on the 2021 Katsina state Mathematics Basic Education Certificate Examination (BECE) had an acceptable content validity index (CVI) within the range of 0.78 and above (i.e., 0.79). This finding was in line with that of Sangoseni et al., (2013); Larsson et al., (2015); Gharamani and Khalkhali (2015); Masuwai et al., (2016); Adachi et al., (2017) and Eleje et al., (2018) in which in terms of CVI their findings possessed an acceptable value which is more than 0.78. However, the finding of this study differs with that of Hadie et al., (2017) in their study which shows that a low level of CVI was exhibited (<0.78).

Meanwhile, the internal consistency of 2021 Katsina state Mathematics Basic Education Certificate Examination was estimated using Kuder and Richardson (KR-20) method. It exhibits a high level of reliability of .83. This indicates that the test is reliable because it is in line with Brown's (2005) recommendation stated earlier. This finding was supported by Osadebe (2014); Dorothy et al., (2015); Larsson et al., (2015), Sa'ad and Ado (2020), and Ahmad, et al., (2020) in their studies that possessed an excellent and acceptable internal consistency reliability coefficient within the acceptable range of

>0.70. The findings of the study also revealed that the 2021 Katsina state Mathematics Basic Education Certificate Examination possessed an acceptable reliability coefficient within the range of 0.80 and above, i.e., 0.83, which is high. This indicated that the tests (in terms of reliability) are good. The finding is in line with that of Agu et al., (2013), and Aggarwal (2018) and Matthew (2019) in their studies in which it was revealed that a high reliability index was exhibited. The study contrasted with that of Toyin (2015), and Margaret and Anthonia (2017), who indicated that their findings have low or poor reliability coefficients within the acceptable range of <0.70.

### **Conclusion**

From the findings of the study, the researchers conclude that the test had content validity and the items were homogeneous.

### **Recommendations**

From the findings of this study, the researchers were able to come up with the following recommendations;

1. Test construction experts should be involved in the construction and validation of tests.
2. Item analysis should be kept up in test development and evaluation, because of its importance in the probe of reliability and in minimizing measurement errors.
3. Katsina state BECE multiple-choice items need to be pilot tested to establish their item statistics.
4. Regular sensitization workshops and seminars should be organized for ERC to familiarize them with the necessary techniques needed for test construction.

### **References**

- Aborisade O. J. & Fajobi O. O. (2019). Comparative analysis of psychometric properties of mathematics items constructed by WAEC and NECO using item response theory approach, *Educational Research, and Reviews*, 15(1), pp. 1-7. <https://doi.org/10.5897/ERR2019.3850>.
- Adachi, D. J., Rodrigues, B. I., Beattie, A. K., & MacDermid, C. J. (2017). Development and Validation of a New tool to measure the Facilitators, Barriers, and Preferences to Exercise in people with Osteoporosis. *BMC Musculoskeletal Disorders*.
- Aggarwal, V. (2018). Construction and Validation of an Achievement Test for Concept Assessment in Biology. *International Journal of Education* Vol. 9, S.No.63022, <http://ijoe.vidyapublications.com>, 32-57.
- Agu, N. N., Onyekubu, C., & Anyichie, A. C. (2013). Measuring Teachers' Competencies in Constructing Classroom-based tests in Secondary Schools: Need for a test construction skills inventory. *Educational Research and Reviews* Vol. 8(8), <https://doi.org/10.5897/ERR12.219>., 431-439.
- Ahari, M. N., Azad, A., Zarei, A. M., Ebadi, A., Parand, A., & Mohammadi, P. (2018). Development and Validity of the School Interim Competency of Performance Skill Battery Scale (SICPSBS). *International Journal Pediatr*, 6(11), <http://ijip.mums.ac.ir>., 8451-8473.
- Alkharusi, H. (2011). Psychometric Properties of the Teacher Assessment literacy Questionnaire for Preservice Teachers in Oman. (procedia-paper presented at International Conference on Education and Educational Psychology (ICEEPSY, 2011). *Procedia-Social and Behavioural Science*. 29. Pp 1614-1624.
- Ahmad, S., Sultana, N., & Jamil, S. (2020). Considerations for Construction and Validation Biology Achievement Test at Secondary Level. *Education Review*, B 10 (1), <http://doi: 10.17265/2161-6248/2020.01.002>., 13-25.
- Anastasi, A & Urbina, S. (2014). *Psychological Testing* (7th ed.). Perfect Impression, ND: PHI-Learning Private Limited.
- Anikweze, C. M. (2010). *Measurement and Evaluation for Teacher Education* (2nd ed.). SNAAP Press Nig Ltd.

- Brown, J. D. (2005). *Testing in Language Programs: A Comprehensive Guide to English Language assessment*.
- Cardenas-Gutierrez, R. A., Bernal-Guerrero, B.A., & Montoro-Fernandez, E (2021). Construction and Validation of the Basic Scale of Entrepreneurial Competencies for the Secondary Education level. *Journal Plus One*. 16(4) <https://doi.org/10.1371/journal.pone.0249903>, 1-21.
- Chong, J., Mokshein, E. S., & Mustapha, R. (2021). A Content Validity Study for Vocational Teachers' Assessment Literacy Instrument (VOTAL). *International Journal of Academic Research in Business & Social Sciences*, 11(4), <http://hrmars.com/index.php/pages/detail/publocation-ethics>., 868-883.
- Creswell, J. W. (2012). *Educational Research: Planning, Conducting and Evaluating Quantitative and Qualitative Research* (4 ed.). Boston: Pearson Education.
- Dorothy, C., Nebife, E., & Ezeudu, S. A. (2015). Development and Validation of an Instrument for Assessing Junior Secondary School Mathematics Classroom Environment. *Global Journal of Educational Research*: 109-118. [http://dx.doi.org/10.4314/gjedr.14\(2\)5](http://dx.doi.org/10.4314/gjedr.14(2)5).
- Doust M. A., Omidvar, N., Keshavarz M. N., Abdollahi, M., Amini, M., & EiniZinab, H. (2017). Developing and validating a scale to measure Food and Nutrition Literacy (FNLIT) in elementary school children. *PLoS ONE*, 12(6). <https://doi.org/10.1371/journal.pone.0179196>, 118.
- Eleje, L. I., Abanobi, C. C., & Obasi, E. (2017). Development and Validation of Economics Achievement Test for Secondary Schools. *Asian Journal of Education and Training* 3(1), 6-17.
- ERC, (2021). Katsina State Ministry of Education Centre. Junior Secondary School Three (JSSIII) Population. Katsina State Ministry of education.
- Feleye, A. B., & Ayelaja, O. R. (2014). Taxonomical Analysis of Selected Teacher-made Multiple Choice Tests. *Journal of Educational and Social Research* 4(3), [doi:10.5901/jesr.2014.v4n3p315](https://doi.org/10.5901/jesr.2014.v4n3p315)., 315-324.
- Gawthrop, J. (2014) *Measuring Students Achievement: A Study of Standardized Testing and its Effects on Students Learning*. Escamilla: P.
- Ghahramani, A., & Khalkhali, R. H. (2015). Development and Validity of a Safety Climate Scale for Manufacturing Industry. *Journal homepage:www.e-shaw.org*.
- Ghazali, N., Nordin, M. S., Hashim, S., & Hussein, S. (2017). Measuring Content Validity: Students' Self-efficacy and Meaningful Learning in Massive Open Online Course (MOOC) Scale. *Advances in Social Science, Education and Humanities Research*, Volume 115 3rd International Conferences on Education in Muslim Society (ICEMS).
- Hadie, S.N.H., Hassan, A. Ismail, Z.I.M., Asari, M.A., Khan, A.A., Kasim, F., (2017). Anatomy Education environment measurement inventory: a valid tool to measure the anatomy learning environment. *Anatomical Sciences Education*, 10(5):42332. <https://doi.org/10.1002/ase.1683>.
- Kim, J. (2011). Developing an instrument to measure social presence in distance higher education. *British Journal of Educational Technology*, 42(5), <https://doi.org/10.1111/j.1467-8535.2010.01107.763777>.
- Kubiszyn, T., & Borich, G. (2013). *Educational Testing and Measurement: Classroom Application and Practice*. (10 ed.). Donnelley Jefferson City: John Wiley & Sons, Inc
- Larsson, H., Tegern, M., Monnier, A., & Skoglund, J. (2015). Content Validity Index and Intra- and Inter-Rater Reliability of a New Muscle Strength/Endurance Test Battery for Swedish Soldiers. *PLOS ONE/DOI:10.1371/Journal.pone.0132185*.
- Lawshe C.H. (1975) A quantitative approach to content validity. *Personnel Psychology* 28(4), 563-575.
- Margaret, O. I., & Anthonia, E. M.-A. (2017). Development and Validation of the Mathematics Achievement Test for Primary School Pupils. *British Journal of Education* 5(7), Pp. 47-57.
- Masuwai, A., Tajudin, M. N., & Saad, S. N. (2016). Evaluating the Face and Content Validity of a Teaching and Learning Guiding Principles Instrument (TLGPI): A prospective study of Teacher Educators. *Malaysian Journal of Society and Space* 12(3): 11-21.
- Matore, E. M & Khairani, A. Z (2015). "Assessing the Content Validity of IKBAR using Content

Validity Ratio. *Australian Journal of Basic and Applied Sciences* 9(7): 255257.

- Matthew, A. (2019). Construction and Validation of Chemistry Achievement Test For Senior Secondary Schools. *International Journal of Innovative Social and Science Education Research* 7(2):106-110.
- METRIC, (2006). Classical Test Theory and Item Response Theory. An Overview. <http://www.measurementexperts.org/learn/theories>.
- Miller, L.A & Lovler, R. L (2013). *Foundations of psychological testing: a practical approach*. SAGE Publications Inc.
- Norashady, M., Muhammad Aziz Shah, M. ., Eftah, M. ., & Haniza, Y. (2016). Measuring the content validity of MEPI using Content Validity Ratio. *Journal of ICT in Education*, 3, pp. 81-88.
- Osadebe, P. U. (2014). Construction of Economics Achievement Test for Assessment of Students of Senior Secondary School two (SSII). *World Journal of Education*, 4 (2) <http://dx.doi.org/10.5430/wje.v4n2p58>.
- Othman, N., & Matore, E. (2020). Validation of Motor Observation Questionnaire for Teachers (MOQ-T) Measurement Items Using Content Validity Ratio (CVR). *Journal of Mechanics of Continua and Mathematical Sciences*, 15(6), <http://doi.org/10.26782/jmcms.2020.06.00018>, 224-244.
- Ramli, N. B., Talib, O. b., Abdul Manaf, U. B., & Hassan, S. B. (2018). Content Validity of STEM TIP Using CVR Method. *International Journal of Academic Research in Business & Social Sciences*, 8(7), <http://hrmars.com/index.php/pages/detail/publication-ethics.>, 1118-1125.
- Sa'ad, N., & Ado, I. M. (2020). Assessing Psychometric Properties of 2018 Mathematics Basic Education Certificate Examination (BECE) Multiple Choice Items. *FUDMA Educational Foundations (FUJEF)* 3(1).
- Sangoseni, O., Hellman, M., & Hill, C. (2013). Development and Validation of a Questionnaire to Assess the Effect of Online Learning on Behaviors, Attitude, and Clinical Practices of Physical Therapists Regarding Evidenced-based Clinical Practiced. *Internet Journal of Applied Health Sciences and Practice* 11(2): 1-12.
- Sekaran, U. & Bougie, R. (2011). *Research methods for business: A skill-building approach*. John Wiley & Sons.
- Sidhu, K. S. (2012). *New Approaches to Measurement and Evaluation*. SterlingPublishers Ltd.
- Toyin, M. D. (2015). Analyzing the Psychometric Properties of Mathematics in Public Examinations in Nigeria. *Research on Humanities and Social Sciences* 5(24).
- Wilson, F. R., Wei Pan, & Schumsky, D. A. (2012). Recalculation of the Critical Values for Lawshes Content Validity Ratio. *Measurement and Evaluation in Counseling and*

*Development*, 45(3). <https://doi.org/10.1177/0748175612440286>.