ISSN 2962-5548
https://doi.org/10.56773/ejer.v3i2.56
Eureka: Journal of Educational Research, 3(2), 144-152

Development of standardized mathematics achievement test instruments for measuring senior secondary school students' learning outcomes

Bashir Mohammed Ibrahim[1]*, Mohammad Ibrahim Sani[2], Ibrahim Maryam Abdallah[3]
[1]Faculty of Education, Ahmadu Bello University, Nigeria
[2]Faculty of Management Science, Bayero University—Kano, Nigeria
[3]Department of Epidemiology and Disease Control, School of Health Technology Kano, Nigeria
*bashiribrahimmohd@gmail.com

Received: August 20, 2024
Revised: December 26, 2024
Accepted: January 23, 2025

Abstract

This study developed a Standardized Mathematics Achievement Test (SMAT) instrument for measuring senior secondary school students' learning outcomes in Dala education directorate, Kano State, Nigeria. The study adopted an instrumentation design. Three research questions were raised to guide the study. The population of the study comprised all 6,462 SS 3 students of the Dala education directorate. The sample size of the study was 132 SS 3 students drawn from the target population in 7 schools with a total number of 248 students using purposive sampling. Fifty multiple-choice items of the SMAT was developed by the researchers. The content validity of the SMAT was determined using a table of specifications. The reliability of the instrument was calculated using Kuder-Richardson formula 20 (KR-20). The result indicated that the content validity as presented in the table of specification of the SMAT items covered the six main sections of the Mathematics scheme. The result indicated that the internal consistency index of the SMAT was found to be a highly reliable instrument with a reliability coefficient of 0.79. The result also indicated that the difficulty index (DifI) of the SMAT items is moderately difficult, and the discriminatory power indicated that the test items distinguish between intelligent and weaker students. The study recommended that mathematics teachers and researchers should always determine the difficulty and discrimination indices of the test items before using such instruments as a tool for measuring students' learning outcomes.

Keywords

Item analysis, reliability, standardized mathematics achievement test, students learning outcomes, validity.

INTRODUCTION

Mathematics education and knowledge are acquired through effective teaching of mathematical concepts in schools, and they aim at producing individuals who can find solutions to identified mathematical problems. Mathematics plays an important role in our daily life, it also plays an important role in academic and business careers. Most students choose Students choose Mathematics as a subject in senior secondary school because of their interest, ability, and its relevance to their future careers [1]. In the assessment of mathematics achievement learning, a developed test that could measure an individual's knowledge or skills after a period of training or mathematical instruction is called an achievement test. Rani and Anisha [2] expressed that the achievement tests give reliable information regarding the decisions taken in the

context of Mathematics education. The purpose of an achievement test in a particular subject area is to determine student's knowledge as stated by Devi and Sharma [3]. Ibrahim and Sani [1] posited that questions of Mathematics Achievement Test (MAT) should be valid and reliable when the questions make a fair coverage of the topic and performance objectives emphasized in the Mathematics curriculum.

The standardization of an achievement test is the performance of an achievement test under standard conditions [4]. It is constructed by the experts or specialists by following the standard procedure of test construction and administration of test, scoring, analysis and interpretation of test results [5]. Basically, standardized achievement tests are designed and prepared by or with the assistance of measurement and evaluation experts for a large number of students. According to Thorndike and Thorndike-Christ [6], standardized achievement tests are used for placement decisions, remedial and diagnostic decisions, selection decisions, guidance and counseling decisions, curricular decisions, decisions on public policy and alternative programs, and decisions on how well a school is doing. Standardized achievement tests are the most popular type of achievement test used to assess the knowledge of students in a specific subject area.

A SMAT is the one which is developed to measure knowledge and skills of students for assessing their level of understanding. For the SMAT, the most important variables to be considered are validity and reliability. The validities to establish are face and content validities. Inko-Tariah and Okon [7] stated that face validity is concerned with the level of English language used, whether the multiple-choice items are ambiguous, whether the keys follow a pattern and are properly keyed, and whether there are overlapping items. As for content validity, the test items are supposed to reflect the entire topics that are expected to be covered in a subject by the students [8].

The development of effective SMATs follows certain procedures. These procedures are used to identify the content areas, format, and table of specifications for the test items. Most standardized achievement tests are developed by professional test publishing organizations, however, the laid down procedures could be adopted by anyone constructing such a test [9]. This is because the standardized test is designed to measure the educational development of all students. This can also be achieved when the Mathematics teachers have a good knowledge of developing an effective standardized Mathematics test for assessing students' knowledge.

Several researchers conducted different studies on the development of achievement tests. For instance, Chime's [10] study developed and validated an Economics Achievement Test. 125 SS II students were sampled for the study. The researcher postulated four research questions and two hypotheses to guide the study. Based on the research questions and hypotheses, the design and methodology, a table of specification was constructed and used. Fifty test items were developed based on the level of cognitive ability measured by the test, and the draft EAT was validated by the experts. The data obtained were analyzed using the mean, standard deviation, and KR-20 reliability. The results of the analysis in his study revealed that the developed Economics achievement test items for senior secondary schools have high psychometric properties in terms of facility and Discrimination Index (DI); the instrument has high-reliability index.

Gourav [11] constructed and validated achievement tests in Economics. 90 test items initially on the selected topics of Economics of class XI on the basis of the blueprint prepared for the achievement test in the light of specific objectives. After the items were written, the researcher consulted the language and subject experts to check the items framed with respect to faulty language or inadvertent defects in wording, and also to verify whether the items measure what they were designed to measure. In a similar study, Opara and Magnus-Arewa [12] developed and validated MAT for primary six pupils. An instrumentation research design was adopted for the study. The study used a simple random sampling technique via balloting, 10 primary schools were drawn from 54 primary schools in Obio/Akpor L.G.A. Through the stratified random sampling technique, 858 primary six pupils were drawn from a population of 2,928 primary six pupils from Obio/Akpor L.G.A.

Kunwar [13] developed and standardized the achievement test to measure students' cognitive level of grade X mathematics. The construction and development of the test consist of the following steps: preparation of the test blueprint, preliminary draft and answer key, pilot testing, item analysis, preparation of the final test,

reliability and validity of the test, and norms. The discriminative index and the power of difficulty level were determined, and the test reliability coefficient was also calculated, which was found to be 0.86. Additionally, the internal consistency using Cronbach's α was found to be 0.85. In his study, validity was also established through expert judgment, and the age norm was also established.

Nwuchegbuo and West [14] developed and standardized a MATs for Junior Secondary Schools (JSS 3) students in Rivers State. Instrumentation design was adopted. Six research questions guided the study. The population of the study comprised all the 34,495 JSS 3 students in Rivers State. The sample size of 600 JSS 3 students was obtained through multi-stage sampling techniques. The validity of the MAT was determined using a test blueprint. The instrument was also validated by two experts in Measurement and Evaluation and Mathematics specialists. The reliability of the instrument was estimated using the KR-20 reliability method, which yielded a reliability coefficient of 0.96 for internal consistency. The data analysis in their study showed that the items of the MAT had good difficulty and discrimination indices.

Statement of the problem

Due to the inability of mathematics teachers to develop effective SMAT instruments for measuring students' learning outcomes with quality, validity, and reliability, some teachers face problems while developing test instruments of poor quality, which fail to measure what they are supposed to measure exactly. At times, the items of the test are not developed based on the cognitive domain of learning. Some teachers set test questions to measure students' achievement in mathematics without knowing the effectiveness of the test items by determining the psychometric properties of the test items such as difficulty and discrimination indices. To support these statements, a questionnaire was administered in a study conducted by Ibrahim and Sani [1], which found that some of the teachers did not construct tests based on the above-mentioned areas. All these difficulties of mathematics teachers contributed to the poor performance of students in mathematics. Therefore, there is a need for researchers in mathematics education to develop SMAT instruments for mathematics teachers. Against this background, the present study aims to develop SMAT instruments for measuring senior secondary school students' learning outcomes in the Dala Education Directorate of Kano State, Nigeria.

Purpose of the study

The purpose of this study was to develop a SMAT instrument for measuring senior secondary school students' learning outcomes in the Dala Education Directorate, Kano State, Nigeria. The specific objectives of this study are to determine the: (1) content validity index of the SMAT instrument for measuring senior secondary school students' learning outcomes; (2) internal consistency index of the SMAT instrument for measuring senior secondary school students' learning outcomes; (3) difficulty and discrimination indices of the SMAT instrument for measuring senior secondary school students' learning outcomes.

Research questions

The following research questions were raised to guide the study: (RQ1) What is the content validity index of the SMAT instrument for measuring senior secondary school students' learning outcomes?; (RQ2) What is the internal consistency index of the SMAT instrument for measuring senior secondary school students' learning outcomes?; (RQ3) What are the difficulty and discrimination indices of the SMAT instrument for measuring senior secondary school students' learning outcomes?

RESEARCH METHOD

This study adopted an instrumentation research design. The design is appropriate for this research because the study intends to develop a valid and reliable SMAT for measuring senior secondary school students' learning outcomes. According to Nworgu [15], an instrumentation research design "is a type of design which aims at the development and certification of the efficacy of an instrument for the measurement of a given behavior or construct."

Population, sample, and sampling technique

This population is spread across fifty-nine (59) public secondary schools. For the purpose of this study, seven (7) schools were selected as the target population, with a total number of 248 students, out of which one hundred and thirty-two (132) students were selected from the seven (7)

schools as the sample size for the study. Each school was characterized, and the data collected from those schools were compared to determine how the difficulty levels and discrimination indices vary across those characteristics. The sample was drawn using the Research Advisor [16] table for determining the sample size with 0.05 degree of accuracy at the 95% level of confidence. Simple random sampling technique was also used to select the students. The sample of students was selected proportionally according to the population of students in each school.

Instrumentation

The instrument for this study was a multiple-choice test items titled; SMAT developed by the researchers. The test comprised 50 items, which were derived from the senior secondary school Mathematics syllabus. This syllabus was obtained from the Kano Education Resources Department (KERD) in the Ministry of Education, Kano State. The SMAT test items consisted of 4 options (A, B, C, and D), with one correct response and three incorrect options (distractors). The SMAT was effectively developed to measure senior secondary school students learning outcomes under the following steps.

Step 1: Planning the test. This step consists of a series of activities that the researchers planned for the development of a SMAT. These activities included: (i) preparing the content areas to be covered; (ii) preparing the behavioral objectives to be measured; (iii) deciding on the test format; (iv) and preparing a table of specifications for the SMAT.

Step 2: Preparation of the MAT. The first draft of the SMAT consists of the selection of the items and types of test items to be selected. After constructing the preliminary draft, it was given to Mathematics experts and two experts in the field of measurement and evaluation for criticism and possible suggestions. A detailed blueprint of SMAT was prepared. After these items were evaluated by the experts, the second draft of the test was also prepared by the researchers, who also prepared the criteria for scoring. Scoring of the SMAT items was done as 1 for a correct response and 0 for an incorrect response. The total mark for the test was 50 marks. The details of the test after the second trial are given in the form of a blueprint.

Step 3: Item writing. The very next step is to develop an item bank by keeping in view the specific learning objectives (knowledge, comprehension, application, analysis, synthesis, and evaluation). The mathematical content included in the test was Numbers and Numeration with 18 items, Algebraic Processes with 10 items, Geometry with 6 items, Introductory Calculus with 6 items, and Statistics with 10 items, making the total distribution across the content areas and process objectives. The writing of test items by the test maker was followed by the development of the table of specifications for the SMAT. These items were written in line with the specifications.

Step 4: Administration of the test: The developed SMAT was administered to SS 3 students who had already been taught the mathematics concepts and covered the content. Sixty (60) minutes were fixed as the duration for the tryout of the test. After the first tryout, the answer sheets were scored using the scoring guide, which had already been prepared by the researchers. One mark was assigned to each correct answer and zero to each incorrect answer.

Step 5: Determine the reliability of the test. The reliability coefficient of the SMAT was determined using the KR-20. The researchers carried out a confirmatory reliability test using the KR-20 reliability method to determine the reliability of the instrument. This was done by administering the instrument to one hundred and thirty-two (132) SS 3 students selected from the education directorate. The students were given the SMAT multiple-choice test items simultaneously. The internal consistency of the SMAT multiple choice test was determined using KR-20.

Kuder and Richardson [17] developed various reliability formulas, each with its assumptions. These formulas are presented as follows:

$$\mathrm{KR}-20=\frac{k}{k-1}\left(1-\frac{\sum pq}{\sigma^2}\right)$$

Where k is the number of items, p is the proportion of testees who responded the item correctly, q is the proportion of testees who responded the item wrongly, and σ^2 is the variance of the total test score.

Step 6: Item analysis: After the SMAT had been administered and scored, the effectiveness of the test items was evaluated. This was done through item analysis. The result from the item analysis provides information for improving the

quality of the test items. Item analysis includes difficulty level and discriminating power. According to Evroro [18], DifI and DI are calculated using the following formulas:

$$\text{DifI} = \frac{\text{RU} + \text{RL}}{\text{NU} + \text{NL}} \text{ and } \text{DI} = \frac{\text{RU} - \text{RL}}{\frac{1}{2}\text{T}}$$

Where, DifI is the difficulty index of the test item, DI is the discrimination power of the test-item, RU is the number of testees in the upper group 27% (36 students) who responded the items correctly, RL is the number of testees in the lower group 27% (36 students) who responded the item correctly, NU or NL is the number of testees in either upper or lower group, and T is the total number of testees in the two groups.

Step 7: Composition of the test items. After item analysis 43 items were selected for the final form of the SMAT. The selected SMAT items that met the standard were composed to form a test. The test was further standardized by experimental validation of the test which included establishing reliability and validity.

Procedure for data collection

The researchers visited the sampled schools, presenting the introduction letter to the principals and requesting permission to administer the SMAT instrument. The SMAT was administered by the researchers to the selected students with the help of research assistants.

Procedure for data analysis

The data collected were analyzed using MS-Excel and SPSS software (version 20.0). Specifically, RQ1 was answered using percentages through the table of specification for content validity, and RQ2 was answered using the KR-20. In order to estimate the reliability of SMAT multiple choice test items when scored using number right scoring, the p, q, sum of all pq, and σ^2 were determined and finally the reliability were estimated, RQ3 was answered using the proportion of the testees in terms of DifI, where items greater than or equal to 0.90 were considered very easy, items range from 0.80 to 0.89 were considered easy items, items range from 0.60 to 0.79 were considered moderately easy items, items range from 0.40 to 0.59 were considered moderately difficult items, items range from 0.20 to 0.39 were considered difficult items, and items less than 0.20 were considered very difficult, while for the DI, items greater than or equal to 0.40 were considered excellent, items range from 0.30 to 0.39 were considered very good, items range from 0.20 to 0.29 were considered good, items less than 0.20 were considered poor items, and items with negative sign indicated a defective items.

RESULT AND DISCUSSION

The analysis from the research study addressing each research question is presented based on the following: (1) Table of specification for SMAT's content validity, (2) SMAT's internal consistency using KR-20, and (3) DifI and DI of the SMAT.

The SMAT's content validity results are presented in Table 1.

Table 1. Computation of the SMAT table of specifications for content validity

Content Area	Behavioural Objectives						Total (100%)
	K (30%)	C (30%)	A (20%)	An (12%)	S (5%)	E (3%)	
Numbers and Numeration (36%)	5	5	4	2	1	1	18
Algebraic Processes (20%)	3	3	2	1	1	0	10
Geometry (12%)	2	2	1	1	0	0	6
Introductory Calculus (12%)	2	2	1	1	0	0	6
Statistics (20%)	3	3	2	1	1	0	10
Total (100%)	15	15	10	6	3	1	50

Note: K=Knowledge, C=Comprehension, A=Application, An=Analysis, S=Synthesis, E=Evaluation.

Determination of SMAT's content validity

This is addressed to RQ1. Table of specifications was used to determine the SMAT's content validity. It has six (6) behavioural objectives: knowledge; comprehension; application; analysis; synthesis; and evaluation. Therefore, Percentages was used to examine the SMAT's content validity.

The analysis is presented in Table 1. Table 1 displays the table of specification, which serves as a guide for the development of the SMAT instrument for measuring senior secondary school students' learning outcomes based on the Ministry of Education curriculum. From the table, the items of the draft copy of the SMAT were distributed across the six (6) behavioral objectives and covered the five (5) content areas of the SS 3 students' syllabus. Based on the curriculum content, Table 1 indicates that a total of 50 items were developed from the five curriculum content areas, with 18 (36%) items drawn from Numbers and Numeration, 10 (20%) items from Algebraic Processes, 6 (12%) items from Geometry, 6 (12%) items from Introductory Calculus, and 10 (20%) items from Statistics. Under the behavioral objectives of the cognitive domain, the table shows that 15 (30%) items were on Knowledge, 15 (30%) items on Comprehension, 10 (20%) items on Application, 6 (12%) items on Analysis, 3 (5%) items on Synthesis, and 1 (3%) item on Evaluation. This result indicated that all the topics under content areas and process objectives were covered by the study. This implies that the developed SMAT instrument for measuring students' learning outcomes is valid and effective in terms of the content areas and process objectives in the cognitive domain based on the senior secondary school (SS 3) curriculum.

Table 2. Computation of the SMAT for internal consistency, DifI and DI

KR-20 Internal Consistency Index							
P_{KR-20}	Item 1	Item 2	Item 3	…	Item 48	Item 49	Item 50
p	0.41	0.42	0.44	…	0.25	0.21	0.39
q	0.59	0.58	0.56	…	0.75	0.79	0.61
pq	0.24	0.24	0.25	…	0.19	0.17	0.24
Difficulty Index (DifI)							
Ranges		Number of Items			Percentage		Remark
< 0.20		-			0%		Very difficult
0.20 – 0.39		29			58%		Difficult
0.40 – 0.59		21			42%		Moderately difficult
0.60 – 0.79		-			0%		Moderately easy
0.80 – 0.89		-			0%		Easy
> 0.90		-			0%		Very Easy
Total		50			100%		
Discrimination Index (DI)							
Ranges		Number of Items			Percentage		Remark
Negative		-			0%		Defective items
< 0.20		7			14%		Poor
0.20 – 0.29		7			14%		Good
0.30 – 0.39		15			30%		Very Good
≥ 0.40		21			42%		Excellent
Total		50			100%		

Note: *p*=Proportion of students who answered correctly, *q*=Proportion of students who answered wrongly, DifI=Difficulty index, DI=Discrimination index.

Determination of SMAT's internal consistency
To address RQ2, KR-20 was used to assess the SMAT's internal consistency.

The analysis is presented in Table 2. From the data in Table 2, the KR-20 was used to compute the internal consistency index of the SMAT instrument for measuring senior secondary school students' learning outcomes. This approach became necessary because the SMAT is a multiple-choice objective test with expected students' responses of either pass (1) or fail (0). Thus, a reliability estimate of 0.79 was obtained. This implies that the developed SMAT instrument is highly reliable and useful for measuring senior secondary school students' learning outcomes.

Determination of the DifI and DI of the SMAT
This is addressed to RQ3. Table 2 also indicates the following. The DifI and DI of the SMAT.

First, the analysis from Table 2 indicated the DifI for the SMAT instrument for measuring

senior secondary school students' learning outcomes. Table 2 shows that the test instrument had a DifI between 0.21 to 0.54. Out of the total 50 items, twenty-nine (29) items, or 58% of the items, were considered difficult in terms of their difficulty level. Therefore, there is a need to modify the items before they can be used for measuring senior secondary school students' learning outcomes. Meanwhile, twenty-one items (42%) were moderately difficult. This implies that the developed test instrument has a moderate difficulty level.

Second, this Table also shows the discriminatory index of the SMAT instrument for measuring senior secondary school students' learning outcomes. The DI as shown in Table 2 ranges between 0.03 to 0.84. Out of the total 50 items, seven (7) items, or 14% of the items were considered poor. Therefore, there is a need to modify the poor items before they can be used for measuring senior secondary school students' learning outcomes. While, seven (7) items or 14% of the items were considered good, fifteen (15) items or 30% of the items were considered very good, and twenty-one (21) items or 42% of the items were considered excellent in terms of their discrimination power. This implies that the developed test discriminates between knowledgeable and weaker students.

Discussion of findings

The findings of RQ1 indicated that the content validity as presented in the table of specification of the SMAT perfectly aligned with the National Curriculum for Mathematics. To ensure the content validity of the SMAT, a table of specification was developed, and the alignment of test items with various levels of Bloom's Taxonomy revealed that the majority of test items fall within the knowledge and comprehension domains. The second highest number falls within the application domain, followed by the analysis domain, and the lowest percentage of items falls within the synthesis and evaluation domains. Therefore, the test items covered the six main sections of the Mathematics scheme; hence, the SMAT has content validity. The finding of this study is in agreement with the findings of Osadebe and Jessa [19] who indicated that the social Studies Achievement Test instrument were valid as shown in the table of specifications. This finding is consistent with the findings of Osadebe [20], who constructed the Economics Achievement Test for the assessment of students and revealed that the Economics achievement test has high face and content validity. Kunwar [13] reported that the achievement test instruments used in their study were reasonably valid.

The findings of RQ2 indicated that the internal consistency index of the SMAT as presented in Table 2 based on KR-20 reliability estimate was found to be a highly reliable instrument. The computed analysis showed a high correlation coefficient of 0.79 which is significant for a standard test. This shows that the SMAT meets the requirement of the internal consistency reliability of the test. This finding is consistent with the finding of Nwuchegbuo and West [14] who established the reliability of the instrument using KR-20 reliability method and found the internal consistency reliability coefficient of 0.96. This result is in agreement with the result of Kunwar [13] which showed a reliability coefficient of 0.85 internal consistency of SMAT.

The findings of RQ3 indicated the difficulty and discrimination indices of the SMAT as also presented in Table 2 which shows that the majority of the SMAT items fall within the acceptable range. The moderate DifI falls within the confidence interval of 0.40-0.59, while the accepted discriminative index falls within the confidence interval of 0.20-0.84. This shows that the DifI of the SMAT items are moderately difficult, and the discriminatory power indicates that the test items distinguish between intelligent and weaker students. This result supported the findings of Chime [10] revealed that the test items of Economics achievement for senior secondary schools have high psychometric properties in terms of facility and DI. The finding of this study agrees with the findings of Osadebe and Jessa [19]; and Nwuchegbuo and West [14] whose indicated that the achievement test items had a good difficulty and discriminatory indices. This finding also agrees with the findings of Gourav [11], who found that 70 items had Difficulty Values (DV) ranging from 0.20 to 0.75, as well as retained items with Discriminating Power (DP) ranging from 0.20 to 0.90.

CONCLUSION

In conclusion, the findings showed that the SMAT instrument indicated that the content validity as presented in the table of specifications of the SMAT perfectly aligned with the National

Curriculum for Mathematics. This implies that the SMAT items is reasonably valid in terms of content areas and process objectives based on the mathematics syllabus. Therefore, it can be used for measuring senior secondary school students' learning outcomes. The finding indicated that the internal consistency index of the SMAT was found to be a highly reliable instrument with a reliability coefficient of 0.79. This implies that the SMAT items are reliable instruments for measuring students' learning outcomes. The result also indicated that the DifI of the SMAT items was moderately difficult, and the discriminatory power indicated that the test items distinguished between intelligent and weaker students. This implies that the majority of the SMAT items fall within the acceptable range.

Based on these results, the following recommendations were raised: (1) mathematics teachers and researchers should always prepare table of specification to determine the content validity index of any achievement test they developed for measuring students' learning outcomes in mathematics in order to ensure that the items are valid and relevant; (2) mathematics teachers and researchers should always ensure that the instrument developed for measuring students' learning outcomes is reliable and useful by estimating the reliability of the test items; (3) mathematics teachers and researchers should always determine the difficulty and discrimination indices of the test items before using such instruments as a tools for measuring students' learning outcomes.

ACKNOWLEDGEMENT

We are very thankful to Almighty God, who kept and preserved us for the successful completion of this research study. The authors would also like to express their deep appreciation to the director of KERD in the Ministry of Education, the principals of the sampled schools, and the sampled students who contributed to the successful completion of this study.

REFERENCES

[1] B. M. Ibrahim and M. I. Sani, "Test construction and validation procedures of senior secondary schools' mathematics teachers of Kano Central Senatorial District, Nigeria," *Eureka J. Educ. Res.*, vol. 3, no. 1, pp. 77–85, Jul. 2024.

[2] R. Rani and A. Anisha, "Construction and Standardization of Mathematics Achievement Test for IXth Grade Students," *Educ. Quest- An Int. J. Educ. Appl. Soc. Sci.*, vol. 8, no. 3, pp. 629–633, 2017.

[3] S. D. Sharmila Devi and H. L. Sharma, "Construction of an Achievement Test for the students of VIII class in the Subject of Mathematics," *Int. J. Sci. Res.*, vol. 2, no. 7, pp. 41–43, Jun. 2012.

[4] B. Gogoi and S. Bhuyan, "Construction and Standardization of an Achievement Test in Mathematics and English grammar for Class IX Students," *Int. J. Creat. Res. Thoughts*, vol. 11, no. 7, pp. 98–115, 2023.

[5] Chanda Rani and Puja Ahuja, "Construction and standardization of Achievement test in Accountancy," *Int. J. Eng. Technol. Manag. Sci.*, vol. 7, no. 1, pp. 61–69, Feb. 2023.

[6] R. Thorndike and T. Thorndike-Christ, *Measurement and Evaluation in Psychology and education*, 8th ed. Pearson, 2009.

[7] D. C. Inko-Tariah and E. J. Okon, "Knowledge of Test Construction Procedures Among Lecturers in Ignatius Ajuru University of Education, Port Harcourt, Nigeria," *Acad. Res. Int.*, vol. 10, no. 1, pp. 130–138, 2019.

[8] A. Matthew, "Construction and Validation of Chemistry Achievement Test for Senior Secondary Schools in Bayelsa State," *Int. J. Innov. Soc. Sci. Educ. Res.*, vol. 7, no. 2, pp. 106–110, 2019.

[9] E. N. Chinyere, M. I. Clementina, and I. A. Chika, "Construction and standardization of mathematics achievement test for senior secondary," *Int. Acad. J. Educ. Lit.*, vol. 2, no. 1, pp. 37–45, 2021.

[10] U. M. Chime, "Development and validation of economics achievement test for senior secondary school students," University of Nigeria, 2012.

[11] G. Mahajan, "Construction and validation of achievement test in Economics," *Int. J. Humanit. Soc. Sci. Stud.*, vol. 1, no. 6, pp. 54–60, 2015.

[12] I. M. Opara and E. A. Magnus-Arewa, "Development and validation of mathematics achievement

test for primary school pupils," *Br. J. Educ.*, vol. 5, no. 7, pp. 47–57, 2017.
[13] R. Kunwar, "Development and standardization process of mathematics achievement test for the students of grade X," *Int. J. Curr. Res.*, vol. 10, no. 11, pp. 75451–75455, 2018.
[14] B. I. Nwuchegbuo and J. West, "Development of Mathematics Achievement Test for Junior Secondary Schools Three Students in Rivers State," *Int. J. Adv. Res. Learn.*, vol. 2, no. 4, pp. 100–112, 2023.
[15] B. G. Nworgu, *Educational Research: Basic Issues and Methodology*. Nsukka: University Trust Publishers, 2006.
[16] Research Advisor, "Sample Size Table," *The Research Advisor*, 2006. [Online]. Available: http://www.research-advisors.com/documents/samplesize-web.xls. [Accessed: 26-Jan-2024].
[17] G. F. Kuder and M. W. Richardson, "The Theory of the Estimation of Test Reliability," *Psychometrika*, vol. 2, no. 3, pp. 151–160, Sep. 1937.
[18] E. S. Evroro, "Item Analysis of Test of Number Operations," *Asian J. Educ. Res.*, vol. 3, no. 1, pp. 18–25, 2015.
[19] P. U. Osadebe and M. O. Jessa, "Development of social studies achievement test for assessment of secondary school students," *Eur. J. Open Educ. E-Learning Stud.*, vol. 3, no. 1, pp. 104–124, 2018.
[20] P. U. Osadebe, "Construction of Economics Achievement Test for Assessment of Students," *World J. Educ.*, vol. 4, no. 2, pp. 58–64, Apr. 2014.