British Journal of Education, 14 (4),82-93, 2026
Online ISSN: 2054-636X (Online)
Print ISSN: 2054-6351(Print)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development-UK

Validation of Social Studies Achievement Test developed by Ekiti State Ministry of Education for Junior Secondary School Examination

Ademola Amos Adegboye
Department of Social Science Education
Bamidele Olumilua University of Education, Science and Technology
Ikere-Ekiti, Nigeria

Olawale Omoya
Department of Counseling Psychology,
Bamidele Olumilua University of Education Science and Technology, Ikere-Ekiti

doi: https://doi.org/10.37745/bje.2013/vol14n48293 Published May 10, 2026

Citation: Adegboye A.A. and Omoya O. (2026) Validation of Social Studies Achievement Test developed by Ekiti State Ministry of Education for Junior Secondary School Examination, *British Journal of Education*, 14 (4),82-93

Abstract: *This study aimed to validate the Social Studies Achievement Test developed by Ekiti state ministry of education for Junior Secondary Three (JS3) students. It was guided by two research questions and adopted an instrumentation research design. The study population consisted of 99,840 JS3 Social Studies students in public secondary schools across the state, from which a sample of 650 students was selected. The instrument validated was the Social Studies Achievement Test for JS3 students developed byEkiti state ministry of education Its reliability coefficient was determined to be 0.76 using the Kuder-Richardson Formula 20 (KR-20). Initially, 70 items were validated; however, following validation These items were further analyzed to determine their difficulty levels, discrimination indices, and the effectiveness of distracters. Findings revealed that 58 items had acceptable difficulty indices, all 70 items demonstrated positive discrimination indices, and all items functioned effectively in terms of distractors. Mean and standard deviation were used to address the fifth research question, showing that the test items aligned with the core curriculum. Overall, the developed Social Studies Achievement Test exhibited strong psychometric properties, particularly in terms of item difficulty, discrimination, and reliability. Based on these findings, it was recommended that the test could serve as a model for constructing achievement tests in Social Studies and other subjects.*

Keywords*:* validation, reliability, social studies, test, achievement.

INTRODUCTION

In contemporary Nigeria, the significance of education is paramount, as it substantially contributes to the nation's holistic welfare and advancement. To fulfill educational objectives, a diverse array of subjects has

been integrated into the academic curriculum and is imparted at various tiers of the educational framework. These subjects are meticulously crafted to furnish learners with the essential knowledge, competencies, and ethical values requisite for the progressive development of Nigeria. At the secondary educational level in Nigeria.

Social Studies is imparted as a pivotal subject for the comprehensive development of students. Its relevance is particularly pronounced in addressing fundamental inquiries regarding human existence, as it synthesizes insights from multiple disciplines that scrutinize diverse dimensions of life. Through engagement with Social Studies, students evolve into enlightened citizens by acquiring an understanding of the "socio-political and economic circumstances of their nation and the global context. They also become aware of their rights and obligations as societal members, which fosters their development into responsible individuals. Furthermore, the subject introduces learners to human behavior and cultural norms, facilitating their comprehension of variations among individuals in terms of social stratification, economic situations, and lifestyles. The objective of Social Studies is to cultivate informed individuals who can effectively function as responsible citizens within their nation. The pedagogical approach of this subject augments students' awareness and promotes ethical conduct.

The effectiveness of students within society is significantly influenced by the quality of Social Studies instruction and the degree to which its educational aims are realized. Consequently, to ascertain whether these educational outcomes have been achieved, it is imperative to evaluate students through assessments and various educational measurement techniques.

Assessments are instrumental for educators to gauge the degree to which learners have attained the educational goals of a subject or program. Adebule and Oluwatayo (2011) characterized a test as a collection of tasks, exercises, or inquiries devised to measure learners' skills and intellect. A test is intended to provoke specific behaviors and functions as an evaluative tool utilized to appraise knowledge, capabilities, aptitudes, attitudes, interests, concentration, motivation, competence, and other related characteristics; consequently, it can manifest in various formats.

In a similar vein, Asuru (2015) described an achievement test as an evaluative tool employed to ascertain what an individual has assimilated as a result of prior instruction or training, typically administered within an educational environment. It is explicitly crafted to ascertain the extent to which a learner has acquired, comprehended, or mastered specific knowledge or skills after undergoing instruction or training. Through the implementation of achievement tests, educators can evaluate students' advancement in a subject and assess whether the articulated educational objectives have been fulfilled.

In any educational framework, it is essential to establish clear criteria for measuring success or achievement, and these criteria should be consistent and standardized. Within the Nigerian context, there is a pronounced emphasis on the necessity for assessment materials to minimize inaccuracies during evaluation. Tools employed to assess achievement must possess requisite psychometric attributes—such

as suitable item difficulty, discrimination, and effectiveness of distractors—to achieve their designated purpose.

Locally developed, teacher-created assessments are the most prevalent instruments for both evaluating students and determining promotion to subsequent academic levels. However, research has shown that many of these tests are often poorly designed, lacking adherence to test blueprints or tables of specification, and do not possess the essential psychometric qualities of validity, reliability, proper difficulty levels, discrimination, and effective distractors (Ugwu, 2012; Chime, 2012; Adams, 1981; Inomiesa, 1988; Nkpone, 2001). Observations of Social Studies test items confirm these concerns: questions are often unevenly distributed across topics, some items are repeated, certain papers are handwritten, and many questions fail to reflect the actual content of the Social Studies curriculum. These issues clearly indicate problems of test invalidity.

According to Ubolom et al. (2011), the reliability of a test refers to the consistency of results obtained when the same individual is assessed with the same instrument at different times or with equivalent forms of the test. In other words, test reliability measures the extent to which similar scores can be reproduced when an individual is retested using the same or equivalent instrument (Obilor, 2018). The closer the reproduced scores, the higher the correlation coefficient, and consequently, the more reliable the test.

Usability, or employability, of a test depends on practical considerations regarding whether the test can be effectively applied. Key factors include the availability of equivalent forms of the test, clarity and simplicity of instructions, ease of administration, simplicity in scoring, ease of interpretation and application of test scores, and overall test economy

Most assessments crafted by educators fail to guarantee the intrinsic and extrinsic dimensions of validity, reliability, usability, administration, and scoring. This research highlights the essential phases in the construction of tests, which encompass test planning, item formulation, trial examination, item analysis, item selection, validity estimation, reliability computation, and the timing and printing of the final test version.

Test planning incorporates all preparatory actions in the construction of assessments, representing the initial phase in developing achievement tests. These actions include articulating and defining objectives, delineating the instructional content, and formulating a test blueprint or specification table. The behavioral articulation of objectives is pivotal in the educational process, as they signify the anticipated behavioral transformations of learners resulting from the instructional experience. They serve to direct both the teaching and evaluative practices. The primary purpose of articulating objectives is to ascertain the extent of the instructional goals to be accomplished. Objectives should be articulated in specific behaviors that students are expected to manifest at the conclusion of the lesson, often employing action verbs that clearly denote direct and observable behaviors.

Outlining the instructional content encompasses the diverse areas, units, topics, and sub-topics that

comprise the course or subject as delineated in the course outline or syllabus" (Asuru, 2015). To ensure that a test adequately samples the topics and sub-topics encompassed in the content, an outline must be constructed. This involves segmenting the content into smaller units. For example, a "test addressing part one of educational measurement and evaluation for undergraduate students could be subdivided into the following units: fundamental concepts, test types, characteristics of effective tests, test construction, administration and scoring, continuous assessment, and soft skills.

The Table of Specification is a two-dimensional chart that presents a list of instructional objectives, content areas, and item types across its dimensions (Obilor, 2019). It also delineates the proportion of questions assigned to each behavioral objective and topic within the content. The preparation of the table of specification comprises four primary steps: (a) identifying which instructional objectives to incorporate; (b) determining the relevant content areas; (c) selecting the item types to include; (d) preparing the two-way chart (Table of Specification).

The subsequent stage of test construction is item writing, which involves framing the items in the requisite language and format. This aspect constitutes a critical component of test construction. Item writing necessitates both practical experience and professional discernment. It is noteworthy to mention that 50% more items than required should be created. Following item writing, the resultant pool of items should undergo review by another expert in the relevant subject area to identify any ambiguities or grammatical errors. Subsequently, the refined pool of items should be produced for trial testing.

Trial testing or the trial run of the item pool represents the third phase in the construction of tests. This process is comparable to the preliminary assessment of a newly constructed industrial facility aimed at verifying whether the installation meets regulatory standards prior to the commencement of actual production activities. Similar to industrial contexts, pilot testing represents a critical phase in the development of assessments, serving as a quality assurance mechanism that ultimately ascertains the appropriateness of each item and the overall test. This evaluation is executed through both qualitative judgment and statistical methods during the item analysis phase.

The test items should be administered to a sufficiently representative cohort of participants who mirror the intended test-takers concerning the content and objectives of their curriculum. The primary purpose of pilot testing is to collect empirical data regarding the individual effectiveness of each item, thereby facilitating informed item analysis. As articulated by Wiggins (1998) and Riaz (2008), item analysis assesses the difficulty level of an item and its efficacy in distinguishing between high and low-performing students. In essence, item analysis yields quantitative evaluations of item difficulty, discrimination, and the performance of distracters. It furnishes objective, external, and empirical evidence regarding the quality of the test items. The aim of item analysis is to pinpoint flawed or ineffective items that may confuse participants or lack a definitive correct answer, with distracters potentially competing with the keyed response. Item analysis encompasses item difficulty, discrimination, and distracter effectiveness.

The fourth stage in the assessment development process is item analysis. This phase entails testing the

item to verify whether it accurately measures the construct that the overall assessment is designed to evaluate (Obilor, 2019). Item analysis commences post-administration and scoring of the test. It incorporates a thorough and systematic evaluation of test-takers' responses to each item to ascertain the item's level of difficulty and discriminative capacity. The procedure for item analysis includes organizing the scores in ascending order from highest to lowest. The scores of the top 27% and bottom 27% of test-takers are selected for analysis. For example, if the assessment was trialed with 450 students, the scripts will be sorted in descending order of scores. One hundred and twenty-one top scores (27% of 450) will constitute the high scorers, while one hundred and twenty-one bottom scores (27% of 450) will represent the low scorers. The remaining scripts will be excluded. These highest 121 and lowest 121 scorers will be utilized for item analysis.

Item difficulty is defined as the percentage of test-takers who correctly respond to an item, which corresponds to the mean item score. The index of item difficulty ranges from 0 to 100; a higher value indicates that the item (question) is easier. Item discrimination pertains to an item's capacity to differentiate among students based on their mastery of the assessed material. Ideally, a larger proportion of individuals in the upper performance group should correctly answer each item compared to those in the lower performance group, while the effectiveness of distracters pertains to their ability to engage students with misconceptions or flawed reasoning, typically those with lower overall competency. It is anticipated that each distracter within an item should function effectively..This indicates that every distractor must be selected by at least one individual from both the upper and lower groups.

The fifth stage of test development is the selection of items, which involves determining the required quantity of items. Initially, items should be organized in descending order based on the extent of their discriminative efficacy. Items exhibiting zero or negative discriminative values should be unequivocally dismissed. Regarding item difficulty, no definitive value is mandated for selection; it is contingent upon the intended application of the assessment. Typically, in achievement evaluations, difficulty levels ranging from 40% to 60% (i.e., .4 to .6) are chosen to accommodate a broad spectrum of abilities and achievements. The rationale for generating more items than necessary is predicated on the necessity to eliminate certain items.

The sixth stage involves estimating the validity and reliability coefficients of the assessment. In this phase, the researcher may utilize various methodologies: for validity, these include content, face, construct, and criterion validity; for reliability, the options encompass Test-Retest, Split-Half, Cronbach Alpha, Kuder Richardson, and Parallel Form methods. The final phase in test construction pertains to timing and formatting the assessment, which should be conducted with consideration for the age, gender, temperament, geographical location, and educational background of the examinees. Additionally, the purpose of the assessment must be taken into account when determining the timing of the test.

Statement of the Problem

The educational framework mandates that various assessments be conducted throughout the instructional process to facilitate effective pedagogy, learning, and to evaluate students' mastery levels. Among these

assessments are routine weekly tests, end-of-term evaluations, and/or concluding session assessments. Regrettably, a significant number of educators lack the knowledge necessary to formulate valid and reliable assessments. Observations by the researcher, as well as findings from other scholars (Chime, 2012; Ugwu, 2012; Obilor, 2019), indicate that teacher-generated assessments frequently lack essential psychometric properties, rendering them unsuitable for accurately evaluating student performance. This deficiency may stem from multiple factors, including educators" insufficient expertise in test development. To address this issue, this study aims to establish a comprehensive test construction framework that will empower educators and relevant stakeholders to develop and implement high-quality assessment items.

The purpose of the study

The purpose of this study was to valid the achievement test of 2024/2025 developed by the ministry of education Ekiti State specifically the study;

Established item difficulty, discrimination indices and effectiveness of distracters of the items of achievement test of 2024/2025 developed by the ministry of education Ekiti state

Research Questions

1.To what extent are the achievement test of 2024/202 developed by the"Ministry of education Ekiti State valid in terms of content of the Social Studies Curriculum.?

2 What are the item difficulty, discrimination indices and effectiveness of distracters of the items achievement test of 2024/2025 developed by the ministry of education Ekiti state?

METHODOLOGY

The study adopted the instrumentation research design. Instrumentation research design was used because the study involved the development and validation of Social Studies Achievement Test (SSAT)) for evaluating the cognitive learning outcomes of junior secondary students in Social Studies. The population of the study was 99840 junior secondary school distributed as follows: Ekiti East – 55647; Ekiti South-East – 21305; and Ekiti West – 22888 (Ekiti State Universal Basic Education Board, 2018). The sample size comprised all the 650 Junior Secondary 3 (JS3) students offering Social Studies in Ekiti State made up of 280 students (Ekiti East), 178 students (Ekiti South-East), and 192 students (Ekiti West).

The instrument used for data collection was the researcher-developed SocialStudies Achievement Test (SSAT) with a reliability coefficient of 0.76 computed using Kuder-Richardson Formula 20 (K-R20). The test was constructed based on the JS3 Social Studies Curriculum. The topics covered were: contents of Social Studies, family life education, roles of extended family members in child development, human trafficking, harmful traditional practices in Nigeria, promoting peaceful living in our society, social conflict, conflict management and resolution strategies, controlling cultism in our society, preventing drug trafficking, common crimes and associated punishment, and Crimes and National Security.

British Journal of Education, 14 (4),82-93, 2026
Online ISSN: 2054-636X (Online)
Print ISSN: 2054-6351(Print)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development-UK

RESULTS

Research Question 1: To what extent are the developed items valid in terms of content of the Social Studies Curriculum?

To answer this Research Question social studies achievement test developed by the researcher (SSAT) was submitted to 47 experts in Social Studies, Measurement and Evaluation, and the teachers of Social Studies from all the sampled junior secondary schools in Ekiti State, for detailed editing, careful and critical review of the test items. This was done toavoid the inclusion of irrelevant items. To further ensure content validity of the SSAT, a table of specification (test blueprint) was used as a guide in constructing the items. The said table of specification was also made available to the Social Studies experts for the validation exercise. Table 1 displays the Table of Specification for the SSAT for JS3in Ekiti State

Table 1: Table of Specification for Social Studies Achievement Test for JS3

Contents	**Know.4%**	**Comp.20%**	**Appl.16%**	**Anal.10%**	**Synth.10%**	**Eval.0%**	**Total100**	
Meaning of social studies	**4**	**1**	**NA**	**1**	**NA**	**NA**	**6**	
Role of family	**4**	**2**	**NA**	**NA**	**NA**	**NA**	**6**	
Human trafficking	**4**	**2**	**2**	**2**	**NA**	**NA**	**10**	
Harmful traditisonal practices	**4**	**1**	**1**	**1**	**1**	**NA**	**8**	
Preventing human trafficking (8%)	**3**	**2**	**1**	**NA**	**1**	**1**	**8**	
Promoting coexistencesociety(8%)	**3**	**2**	**2**	**1**	**NA**	**NA**	**8**	
Social conflict	**4**	**2**	**2**	**1**	**NA**	**NA**	**9**	
Managing and resolving issues	4	2	1	NA	**NA**	**NA**	**7**	
Managing and and resolving	4	2	1	NA	**NA**	**NA**	**7**	
Controllingcultisminour	4	1	2	1	**NA**	**NA**	**8**	
society(10%)								

Research Questions2: What are the item difficulty, discriminating indices, and effectiveness of distracters of the items?

British Journal of Education, 14 (4),82-93, 2026
Online ISSN: 2054-636X (Online)
Print ISSN: 2054-6351(Print)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development-UK

Table2: Item Analysis Showing the Difficulty, Discriminating Indices and Distracters

S/No.	Items	Keys	Upper27%	Lower27 %	ifficulty Index	riminating Index	Distracters
1	1	D	160	82	0.69	0.44	A,BandC
2	2	C	158	49	0.59	0.62	A,BandD
3	3	C	104	57	0.46	0.27	A,BandD
4	4	D	147	59	0.59	0.50	A,BandC
5	5	C	163	54	0.62	0.62	A,BandD
6	6	A	158	62	0.63	0.55	B,CandD
7	7	B	148	54	0.58	0.54	A,CandD
8	8	B	85	40	0.36	0.26	A,CandD
9	9	A	138	60	0.57	0.44	B,CandD
10	10	A	145	72	0.62	0.42	B,CandD
11	11	A	153	56	0.58	0.55	B,CandD
12	12	B	150	55	0.59	0.54	A,CandD
13	13	B	83	47	0.37	0.21	A,CandD
14	14	B	138	51	0.54	0.50	A,CandD
15	15	A	141	53	0.55	0.50	B,CandD
16	16	A	143	56	0.57	0.50	B,CandD
17	17	C	157	66	0.64	0.52	A,BandD
18	18	C	140	61	0.57	0.45	A,BandD
19	19	D	139	48	0.53	0.52	A,BandC
20	20	A	128	49	0.51	0.45	B,CandD
21	21	C	113	48	0.46	0.37	A,BandD
22	22	D	118	47	0.47	0.41	A,BandC
23	23	A	123	51	0.50	0.41	B,CandD
24	24	D	142	70	0.61	0.41	A,BandC
25	25	D	152	57	0.60	0.54	A,BandC
26	26	A	114	62	0.50	0.30	B,CandD
27	27	D	80	35	0.33	0.26	A,BandC
28	28	D	128	51	0.51	0.44	A,BandC
29	29	A	133	60	0.55	0.42	B,CandD

Table 2 continues:

30	30	A	138	74	0.61	0.37	B,CandD
31	31	A	140	63	0.58	0.44	B,CandD
32	32	C	155	58	0.61	0.55	A,BandD
33	33	A	141	68	0.60	0.42	B,CandD
34	34	A	114	67	0.52	0.42	B,CandD
35	35	D	133	52	0.53	0.46	A,CandC
36	36	B	156	52	0.59	0.59	A,CandD
37	37	D	150	51	0.57	0.56	A,BandC
38	38	C	154	53	0.59	0.57	A,BandD
39	39	D	151	60	0.60	0.52	A,BandC
40	40	D	90	48	0.39	0.24	A,Bandc
41	41	A	153	56	0.58	0.55	B,CandD
42	42	A	163	54	0.62	0.62	B,CandD
43	43	D	158	49	0.59	0.62	A,BandC
44	44	A	124	58	0.52	0.37	B,CandD
45	45	B	150	5	0.59	0.53	A,CandD
46	46	C	155	58	0.61	0.55	A,BandD
47	47	B	136	56	0.55	0.48	A,CandD
48	48	D	152	60	0.61	0.53	A,BandC
49	49	B	154	58	0.61	0.55	A,CandD
50	50	D	150	55	0.59	0.31	A,BandC
51	51	C	120	50	0.48	0.40	A,BandD
52	52	A	120	70	0.54	0.28	B,CandD
53	53	D	80	35	0.33	0.26	A,BandC
54	54	A	85	40	0.36	0.26	B,CandD
55	55	A	100	52	0.43	0.27	B,CandD
56	56	C	136	56	0.55	0.48	A,BandD
57	57	B	152	60	0.61	0.53	A,CandD
58	58	D	97	60	0.49	0.21	A,BandC
59	59	C	83	47	0.37	0.21	A,BandD
60	60	C	104	57	0.46	0.27	A,BandD
61	61	A	147	59	0.59	0.50	B,CandD

Table 2 continues:

62	62	C	91	48	0.39	0.24	A,BandD
63	63	D	128	49	0.51	0.45	A,BandC
64	64	A	113	48	0.46	0.37	B,CandD
65	65	C	98	40	0.39	0.33	A,BandD
66	66	A	123	51	0.50	0.41	B,CandD
67	67	A	142	70	0.36	0.41	B,CandD
68	68	B	88	40	0.36	0.27	A,CandD
69	69	D	114	62	0.50	0.30	A,BandC
70	70	A	85	41	0.36	0.25	B,CandD

Table 2 presents the item difficulty and discrimination indices for each item in the instrument. Regarding item difficulty, 58 out of the 70 items have indices ranging from 0.43 to 0.61, which fall within the acceptable range. However, 12 items have difficulty indices between 0.33 and 0.37, indicating that they may not be suitable for inclusion.In terms of item discrimination, all 70 items show indices ranging from 0.21 to 0.62, which are within the acceptable limits. Additionally, all 70 items demonstrated effective distractors.

DISCUSSION OF FINDINGS

Findings emanated from the study indicated that 58 out of the 70 items had difficulty indices ranging from 0.43 to 0.61, which fall within the acceptable range. However, 11 items had difficulty indices between 0.33 and 0.37, suggesting that they are less suitable. Regarding item discrimination, all 70 items recorded indices between 0.21 and 0.62, which also meet acceptable standards. This suggests that the majority of the items are appropriate, as they demonstrate suitable difficulty levels and positive discrimination indices. These findings align with those of Inomiesa (1998), who developed and validated an upper primary science achievement test. In that study, 102 items had facility indices ranging from 0.30 to 0.70 and discrimination indices between 0.20 and 0.60. Similarly, Adams (1981), in constructing and validating an Integrated Science achievement test for Nigerian Junior Secondary School (JS1) students, reported psychometric properties including an average discrimination index of 46.33 and difficulty indices ranging from 18.5 to 82.5. The present results are also consistent with the findings of Chime (2012) and Adams (1981), both of whom concluded that distracters in achievement test items functioned effectively.

Furthermore, the developed Social Studies Achievement Test (SSAT) for junior secondary schools in Ekiti State demonstrated a high reliability coefficient of 0.76 using the Kuder–Richardson formula. This is comparable to the findings of Effiong (2006), Nkpone (2001), and Inomiesa (1988), who reported reliability coefficients of 0.99, 0.89, and 0.87 respectively, also using the Kuder–Richardson formula. Since these values are considered indicators of high reliability, the SSAT in the present study can likewise be regarded as highly reliable and, therefore, a valuable assessment instrument.

CONCLUSIONS

Based on the results obtained, the following conclusions were drawn:

1. The Social Studies Achievement Test (SSAT) was successfully developed by the researcher.
2. The validated items were aligned with the Social Studies curriculum.
3. The developed and validated SSAT demonstrated appropriate difficulty and discrimination indices, with items that functioned effectively in distinguishing learners' performance.
4. The test exhibited a high reliability coefficient of 0.76, calculated using the KR-20 method.

Recommendations

In light of the study's findings, the following recommendations are proposed:

i. The developed Social Studies Achievement Test should be adopted by Social Studies teachers" for assessment purposes.

ii. The SSAT items can serve as a model for developing similar achievement tests in Social Studies and other subjects at the junior secondary school level.

iii. Regular workshops, seminars, and conferences should be organized to equip teachers with the necessary skills for constructing valid and reliable assessment instruments.

iv. Educational inspectors should strengthen their supervisory roles through consistent, systematic, and objective inspections.

REFERENCES

Adams, D. (1981). *Construction and validation of an Integrated Science achievement test for Nigerian Junior Secondary School students*. Unpublished thesis.

Adebule, S. O., & Oluwatayo, J. A. (2011). *Educational tests and measurement principles*. Ado-Ekiti: Green Line Publishers.

Asuru, V. A. (2015). *Measurement and evaluation in education and psychology*. Port Harcourt: Minson Publishers.

Chime, C. (2012). *Assessment and evaluation techniques in Nigerian schools*. Enugu: Academic Publishing Company.

Effiong, U. M. (2006). *Development and validation of achievement test instruments in secondary schools*. Calabar: University Press.

Ekiti State Universal Basic Education Board. (2018). *Statistical records of junior secondary school enrolment in Ekiti State*. Ado-Ekiti: SUBEB Publications.

Inomiesa, E. A. (1988). *Development and validation of upper primary science achievement test*. Benin City: University of Benin Press.

Nkpone, H. N. (2001). *Educational measurement and evaluation for effective teaching and learning*. Port Harcourt: Paragraphics Publishers.

Obilor, I. E. (2018). *Fundamentals of test and measurement in education*. Port Harcourt: Harey Publications.

Obilor, I. E. (2019). *Test construction and validation techniques in educational assessment*. Port Harcourt: Harey Publications.

Riaz, A. (2008). *Educational assessment and item analysis methods*. New Delhi: Educational Publishers.

Ubolom, W. J., Enukoha, O. I., & Nwanekezi, A. U. (2011). *Measurement and evaluation in education*. Port Harcourt: Harey Publications.

Ugwu, A. N. (2012). *Teachers' competence in test construction and classroom assessment in Nigeria*. Nsukka: University of Nigeria Press.

Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco: Jossey-Bass.