PAEDAGOGIA
Jurnal Penelitian Pendidikan
http://jurnal.uns.ac.id/paedagogia

DOI : 10.20961/Paedagogia.v28i1.99107
Vol. 28, No. 1, 2025, pp. 24-33
p-ISSN 0126-4109
e-ISSN 2549-6670

Achievement Testing Competence of Secondary School Teachers in Osun State, Nigeria

Olawale Ayoola Ogunsanmi[1], Victor Abiodun Lijofi[2], Abiodun Enitan Gbolagun[3]
[1,3] Department of Educational Foundations and Counselling, Obafemi Awolowo University, Ile-Ife.
[2]Department of Educational Foundations and Counselling, University of Ilesha, Osun State

Keywords: *Assessment, Achievement Test, Teacher's Competence, Middle and High Schools*

Article history
Received: 1 February 2025
Revised: 8 February 2025
Accepted: 23 February 2025
Published: 27 February 2025

*Corresponding Author Email:
olawaleogunsanmi82@gmail.com

Doi: 10.20961/paedagogia.v28i1.99107

Abstract: Properly designed tests in the classroom can provide teachers with accurate and helpful information about their students. The study determined the content validity of an achievement test used by teachers and examined the test construction skills of the teachers in Osun State secondary schools. A descriptive research design was adopted for the study. The population consisted of public secondary school teachers in Osun state. A sample of 120 teachers was selected purposively from six Local Government Areas based on five core subjects offered in each school, and 12 schools were selected using a simple random sampling procedure. Two research instruments, The test Construction Skills Inventory (TCSI) and Observation Records Sheet (ORS), were used for data collection. The TCSI possessed reliability estimates of 0.85. Data were analysed using frequency count and content validity index (CVI). The results showed that the achievement tests used by teachers in high schools are (CVI = 6.2, mean topic covered = 3.7, 59.7%) while middle schools are (CVI = 6.0, mean topic covered =3.4, 58%). This implied that both achievement tests used by secondary school teachers have moderate content validity. The results also revealed that the level of test construction skills is high (66, 55.0%), and a mean competence value of 3.2 for both middle and high school teachers is suitable enough to ensure proper achievement testing competence. The study concluded that the test construction skills of Osun State secondary school teachers are highly sufficient to provide adequate achievement testing competence.

How to cite: Ogunsanmi, O. A., Lijofi, V. A. & Gbolagun, A. E. (2025). Achievement Testing Competence of Secondary School Teachers in Osun State, Nigeria. *PAEDAGOGIA*, *28*(1), 24-33. doi: 10.20961/paedagogia.v28i1.99107

INTRODUCTION

Educational decisions depend upon valid and reliable measures. These measures are used to make informed decisions because they provide different alternatives (Gareis & Grant, 2015). The main goal of classroom testing is to obtain valid, reliable, and valuable information concerning the testee's learning achievement. In other words, classroom tests reveal the latent ability of examinees and serve as a determinant of what is to be measured, using test items constructed to measure intended instructional objectives known as achievement tests (Rao & Banerjee, 2023). An achievement test measures an individual's proficiency in knowledge or skill. Typically administered in school settings, achievement tests are designed to inform how well test takers have learned the material introduced in school. The tests are standardised, and an individual's performance is usually determined by comparing it to the norm, the performance of a national group of students in the individual's grade or age level who took the same test. These tests are given out daily, weekly, monthly, quarterly, semi-annual or end-of-term to students in all educational levels. The primary goals of achievement tests are to evaluate students' cognitive capacities in the relevant course and ascertain if the previously established learning objectives were successful in that particular course (Duckworth & Yeager, 2015).

Achievement tests also measure students' mastery of a particular instructional domain and are used to decide the advancement or competency of the students. The content of achievement tests can be derived from three sources: textbooks, course syllabuses, and class objectives (Mahajan, 2015). Furthermore, it accurately reflects the emphasis on essential aspects of instruction and measures the appropriate level of students' knowledge of a school subject (Kunter et al., 2013). These tests can be

either classroom-based, produced by an individual classroom teacher for their class (teacher-made tests) or standardised, developed by test experts or examination bodies for more global purposes (Sasu, 2017). The most instructional-relevant achievement tests, however, are classroom-based and, if carefully constructed, provide teachers with accurate and helpful information about the knowledge retained by their students in particular subjects (Nkemka, 2022)

However, most achievement tests are considered criterion-referenced as students' scores are compared with the level of mastery achieved rather than with other students (Kaftandjieva, 2010). Therefore, students' performance is measured according to an agreed-upon criterion and standard. To achieve instructional objectives and the purpose of testing, there is a need for a competent teacher who can easily construct classroom-based tests. Similarly, every classroom teacher must possess and apply requisite skills in creating good items for class assessments. In light of these, validity and reliability are essential characteristics of every test (i.e., a good test item must be valid and reliable) (Bowes et al., 2017). A test is valid if it is suitable for the intended purpose. On the other hand, a test is reliable if it measures what it is supposed to measure consistently under all conditions. The various validity evidence of a test include face, content and construct validity (Snow, 2012). Content validity is the extent to which a test measures the subject matter content and the instructional objectives designed for a given course. It also refers to item sampling adequacy because it allows the content domain to be shown through a specific selection of items (DeVellis & Thorpe, 2021).

Kissi (2020) states that testing competence is "an integrated set of an individual's knowledge, skills, attitudes, and values in/on the assessment that is acquired, modifiable, and unobservable but demonstrable ability". Competence in testing pertains to the capacity of an individual to apply or exhibit the knowledge and abilities they have gained from assessment training to evaluate the learning of students (Kissi, 2020). Test construction competence and quality are vital ingredients any classroom teacher needs if pedagogical goals are to be achieved. The importance of tests in a school system cannot be overemphasised since it is the process of attaining educational goals. These competencies are outlined by Carpenter (2018) as objectivity, communication, item validation skills, and skills for applying appropriate strategies for ascertaining the reliability of test instruments. According to Agu et al. (2013), skill in test construction enables a teacher to construct tests with precision, appropriateness of language use, objectivity, and good grading scales. Teachers need not be educational measurement and evaluation experts to build valid and reliable tests. Still, there are basic test construction skills that every teacher ought to possess to construct quality tests. These skills help teachers structure items to elicit clear and concise answers from students, construct tests appropriate for learners of different ages, abilities and genders; and set tests so that students finish within the allotted time and do not grow scared of tests.

In education, it is becoming increasingly crucial for teachers to comprehend and apply classroom-based tests to enhance students' learning (Guskey & Brookhart, 2019; Kissi et al., 2023). As they have historically been in curriculum and instruction, teachers must be skilled and competent in assessment (Gareis & Grant, 2015). The body of research on the relationship between achievement testing competencies and the quality of assessments created by teachers has demonstrated that this relationship exists (Blasi, 2005; Agu et al., 2013; Kinyua & Okunya, 2014; Kissi et al., 2023). Accordingly, producing high-quality test instruments directly correlates with a teacher's proficiency in creating test items (Hamafyelto et al., 2015). The quality of tests developed by classroom teachers is negatively impacted when their skills in test construction are inadequate.

Research has shown that classroom teachers generally write poor items and that a typical classroom teacher in a secondary school cannot construct good multiple-choice test items, which in turn may affect the performance of students (Baffoe, 2021; DiBattista & Kurzawa, 2011; Kissi, 2020). Teachers ' lack of test construction skills might result in false assessments of students' achievements. Some researchers (Adeosun, 2024; Petters & Okon, 2014) revealed that incompetency in test construction by teachers is a significant cause of malpractice in school examinations by teachers and students in Nigerian secondary schools. Classroom-based achievement tests have been extensively used in Nigerian secondary schools. Over the years, these achievement tests have been criticised for lack of proper psychometric properties. The issue of validity and reliability of classroom-based achievement tests in Nigerian secondary schools has engaged the attention of researchers (Blasi, 2005; Adeosun, 2024; Kissi

et al., 2023). These researchers observed that most of these classroom-based tests in Nigeria lack validity and reliability because teachers lack test construction skills and thus cannot construct good achievement tests. More so, tests used for continuous assessments and end-of-term examinations in secondary schools sometimes contain ambiguous and misleading questions, which may be why some students fail (Durowojum, 2014). The implication is that most teachers lack competencies in test construction and may be using poorly constructed tests to measure student's achievements in various school subjects. When students' achievement levels are not adequately measured and interpreted, the teachers and school administrators cannot provide educational opportunities and support each student's needs.

Aina (2016) and Ogunyinka et al. (2015) asserted that there was great employment of untrained and unqualified teachers in the teaching profession in the Nigerian school system. Therefore, teaching became a means to an end for many as it was used as a stepping stone to greener pastures, which resulted in an influx of incompetent teachers in the school system, leading to student failure in public examinations. This may be one of the factors that contributed to the more significant influx of incompetent teachers into the education sector, which has a feedback effect on the testing competence of the teachers. Moreover, students' performance in achievement tests is a function of many factors, including item quality and the nature of test administration. However, the lack of coherence in the content examined and material assessed could lead to a test that fails to provide evidence from which teachers could make valid judgments about students' progress.

The ability of a teacher to construct test items of good quality and not ambiguous will enable the test to measure the intended content and create it accurately and adequately (Kissi et al., 2023). Hence, a teacher with such ability is considered competent in test construction. Conversely, when a teacher lacks all the aforementioned attributes, it may affect the quality of test items constructed and decisions made from such tests. This may result in a test having a consequential validity issue. Moreover, students' test performance may be linked to the quality of the test items and the nature of the test construction (Ward & Bennett, 2012). Teachers-made test items employed for classroom testing may possess inadequate psychometric characteristics, affecting the quality of their constructed test questions (Kubiszyn & Borich, 2024; Voss et al., 2011). The problem of teachers constructing poor tests is a significant issue in education that requires special attention. This calls into question the test construction competence of the teachers. Hence, there is a need to investigate the achievement testing competence of teachers in Osun State secondary schools. Based on the study's objective, the following research questions were developed to guide the study.

1. To what extent does the achievement test used by middle and high school teachers in Osun State possess content validity?
2. What is the level of test construction skills of Osun State secondary school teachers?

METHOD

The study adopts the descriptive survey research design. The population of the study comprises a total of 5204 public Secondary School teachers in Osun State. The total number of teachers shall consist of 3147 males and 2057 females (Ministry of Education, Science and Technology, Osun state, 2016). The teachers in the public secondary schools are either graduate teachers or those with Nigerian Certificate of Education (NCE), who are employed to teach in the secondary schools (known as either middle schools or high schools in the study area). The study sample was selected using the multistage sampling technique. From the three senatorial districts in Osun State, two Local Government Areas (LGAs) were randomly selected out of the ten LGAs in each district. The two randomly selected local governments in each district. From each of the two chosen LGAs, two public schools were selected using a stratified random sampling technique with school type (middle and high) for stratification, making a total of twelve schools. Ten teachers were selected in each school using a purposive sampling technique. The selection was based on five core subjects offered in each school (middle and high), totalling 120 teachers. The choice was based on the five core subjects offered in each school (middle and high), namely English, Mathematics, Civil Education, Basic Science, Business Studies for the middle schools and English, Mathematics, Civil Education, Biology, and Economics for the high schools. The researcher adapted an instrument titled Test Construction Skills Inventory (TCSI), containing forty-one items adapted from

similar scales. The instrument is divided into two sections. Section A consists of the demographic variables of the respondents, including age, qualification, gender, teaching experience, etc. In contrast, section B measures respondents' competence, skill, and degree of test administration/supervisory behaviour. The researcher developed the observation checklist to determine the level of content validity of the test items developed by classroom teachers. The instrument Test Construction Skill Inventory (TCSI) initially consisted of 41 items, which was reduced to 40 by experts' judgement. Two experts reviewed the TCSI and are experts in tests and measurement. Participants' responses from the pilot study were used to determine the reliability coefficient. The internal consistency reliability of the TCSI was established using Cronbach Alpha, and it yielded 0.85, an acceptance coefficient for a reliable instrument. An Observation Record Sheet (ORS) was developed and scrutinised by the supervisor to ensure its appropriateness in collecting and recording the relevant school records. Research question one was analysed using content analysis. Research question two was analysed using frequency count (mean) to determine the mean competence.

RESULT AND DISCUSSION

Research Question 1: To what extent does the achievement test used by middle and high school teachers in Osun State possess content validity?

Content validity is determined by the representativeness and adequacy of the items included in the final test. The sample of 120 teachers made tests that included various test items and formats. Each participating classroom teacher was asked to provide a copy of their most recently administered teacher-made tests for a subject specifically in the 2015/2016 session. This procedure collected 111 (92.5%) usable teacher-made tests recently used in a regular classroom setting and subject diaries. The content of the teacher-made tests of each subject was checked alongside the expected topics to be covered during such period as stated in the curriculum and written in the subject diaries in their respective schools. Furthermore, the average topics covered for each subject across the selected classroom teachers and their percentages were also determined. The percentages were categorised in terms of content coverage as "0% -20%" as "very low", "21% - 40%" as "low", "41% - 60%" as "moderate", "61% - 80%" as "High" and "81% - 100%" as "very high". The various percentages represent the extent to which their tests are content valid concerning the item's content relevancy and coverage. The results are presented in Tables 1 and 2 below for High and Middle schools, respectively.

Table 1. Content Validity of Achievement Test Used by Teachers (High School)

Subject	Number of teachers made test	Expected Topics	Average Topics Covered	% Average Topics covered	Decision
Mathematics	12	8	4.5	56.3	Moderate
English	12	6	5.6	93.3	High
Biology	10	6	2.5	41.6	Low
Civics Education	12	5	3.4	68	High
Economics	12	6	2.3	38	Low
Average Content Validity Indeks:		6.2	3.7	59.7%	Moderate

Table 1 shows the content validity of teacher-made tests concerning content coverage for adequacy and relevancy. The results show that Mathematics has a moderate content validity (56.3%), with English and Civic Education having a high content validity of 93.3% and 68%, respectively. However, Biology and Economics have low content validity of 41.6% and 38%, respectively. Finally, the results indicated a moderate average content validity index of 59.7% for the achievement tests used by High school teachers in Osun state secondary schools. Moreover, the results of the extent to which the achievement tests used by middle school teachers are content valid are also presented in Table 2.

Table 2. Content Validity of Achievement Tests Used by Teachers (Middle School)

Subject	Number of teachers made test	Expected Topics	Average Topics Covered	% Average Topics covered	Decision
Mathematics	12	6	3	50	Moderate
English	11	5	4	80	High
Business	12	5	3.4	68	High
Civic Education	10	6	2.5	41.6	Low
Basic science	8	8	4.5	56.3	Moderate
Average Content Validity Index:		6	3.48	58%	Moderate

Table 2 shows the content validity of teacher-made tests concerning content coverage for adequacy and relevancy. The results show that Mathematics and Basic Science have a moderate content validity of 50% and 56.3%, respectively, with English and Business Studies having a high content validity of 80% and 68%, respectively. However, Civic Education has a content validity of 41.6%. Finally, the results indicated a moderate average content validity index of 58% for the achievement tests used by Middle school teachers in Osun state secondary schools.

Research Question 2: What is the level of test construction skill of Osun State secondary school teachers?

To answer this research question, the results of the mean, standard deviation, and standard error of the mean of teachers' competence rating in Osun State secondary schools are presented in Table 3. Furthermore, the teachers' level of testing competence was then subjected to descriptive statistics, and the results are shown in Table 4.

Table 3. Mean Competence Rating of Teachers in Test Construction Skills

No.	Skill	Mean	S.D.	Std. Error of mean
1	Outline the content covered for the term before setting tests for students.	3.28	0.96	0.08
2	I consult previous test questions before setting questions.	2.95	0.83	0.07
3	Prepare a test blueprint or specifications table as a test construction guide.	2.63	0.91	0.08
4	Consult the standard textbook on the subject for guidance.	3.20	0.88	0.08
5	Organise test items in a logical manner	3.09	0.79	0.07
6	Give clear instructions to guide the test items.	3.39	0.66	0.06
7	Write tests so that both high and low achievers can understand	3.60	0.57	0.05
8	Subject test items to item analysis	3.01	0.77	0.07
9	Keep a resource bank of questions that can be referred to when setting tests.	2.96	1.05	0.09
10	Set tests with due regard to the time available for testing.	2.93	0.91	0.08
11	Add enough test items to cover all the requisite levels of the cognitive domain.	2.93	0.78	0.07
12	Ascribe scores for each test item	3.28	0.83	0.07
13	Ensure that the items measure the determined objectives for teaching the topic.	3.28	0.69	0.06
14	Set essay items that elicit creative and imaginative answers from the students.	3.19	0.70	0.06
15	Prepare a marking guide while constructing the test.	3.43	0.76	0.06

No.	Skill	Mean	S.D.	Std. Error of mean
16	Consider the age of learners during item writing.	2.85	1.01	0.09
17	Avoid gender stereotypes in the test items.	2.98	1.01	0.09
18	Add sufficient items to cover the appropriate instructional unit.	3.10	0.92	0.08
19	Submit items for vetting to the head of the department or principal.	3.06	1.01	0.09
20	Submit items meant for promotional examination for the expert's review.	2.74	0.97	0.08
21	Avoid the use of clues in multiple-choice questions.	2.86	0.93	0.08
22	The number diagram in the test clearly	3.07	1.13	0.10
23	Avoid the use of interlocking items.	2.75	0.91	0.08
24	Avoid items that measure opinion.	2.59	0.96	0.08
25	Review the test draft at least two times two days before the examination.	2.85	0.96	0.08
26	Limit essay tests to high-level objectives	2.74	1.00	0.09
27	Avoid overlapping alternatives in writing objective tests.	2.97	0.89	0.08
28	Use appropriate formats in writing tests.	3.07	0.88	0.08
29	Avoid too long questions in item writing.	3.21	0.82	0.07
30	Set test items that elicit information on one thing at a time.	2.96	0.80	0.07
31	Detecting when students cheat on a test	3.15	0.90	0.08
32	Supervision of peer instruction	3.03	0.94	0.08
33	Writing items on the affective domain of student's behaviour	2.67	0.97	0.08
34	Determining the reliability of measuring instrument	2.80	0.84	0.07
35	Reporting students' performance	3.23	0.88	0.08
36	Ensuring tests have content validity	3.08	1.06	0.09
37	Ability to analyse test results to identify areas of difficult	3.11	0.81	0.07
38	I am not comfortable giving students test items to mark.	2.65	1.03	0.09
39	I always give instructions before the commencement of the test/exam.	3.26	1.03	0.09
40	Attending workshops/seminars on test planning, construction, and administration	2.98	0.93	0.08
	Grand Mean	**3.02**		

Table 3 shows that item means (item 1 had a mean of 3.28, item 2 had a mean of 2.95, item 3 had a mean of 2.63. item 40 had a mean of 2.89), and the cluster mean of 2.89. To answer the research questions, the mean value of the scale, which is 2.89, was taken as the cut-off point. This represents the lower absolute limit of competence on the scale. The mean values of 2.89 and above were considered competent in that achievement testing skill. Teachers were found to prepare a marking guide after setting questions with the highest mean competence of 3.48, followed by an outline of the content covered before setting questions with mean competence of 3.28. However, they were found to have the lowest mean competence of 2.64 in preparing test blueprints as a guide when constructing test items. Hence, the teachers possess 29(72.5%) achievement testing skills from the 40 achievement testing skills identified. However, the grand mean of 3.02 was obtained for the average achievement testing competence of

teachers in both middle and high schools used for the study. This result indicated that most teachers possess specific skills to ensure proper achievement testing activities.
Furthermore, the teachers' level of testing competence was then subjected to descriptive statistics, and the results are shown in Table 4.

Table 4. Teachers' Level of Testing Competence

Level of teachers' competence	Frequency	(%)
Very low	17	14.2
Low	23	19.2
High	66	55.0
Very high	14	11.7
Total	120	100

Table 4 shows that 17(14.2%) of the teachers have a very low level of achievement testing competence, and 23(19.2%) have a low level of achievement testing competence. In comparison, 66(55.0%) and 14(11.7%) teachers have high and very high achievement testing competence, respectively. This implied that the level of test construction skills that Osun state secondary school teachers possess is high to ensure proper achievement testing competence. This is further visualised as shown in Figure 1.

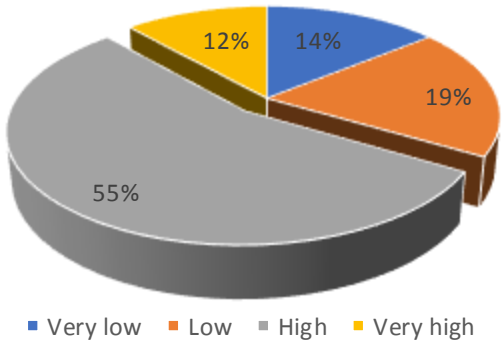

Figure 1. Teachers' Level of Testing Competence

Discussion of Findings

The findings from research question one indicated that the achievement tests used by teachers in middle and high schools in Osun State possess moderate content validity. It further revealed that among other core subjects, English Language possesses the highest content validity index for middle and high schools. This finding corroborated the findings of Musah et al. (2022) that teachers in the chosen schools were well-versed in connecting the curriculum with the assessment tool using process- and evidence-based approaches. Therefore, in most of their evaluation methods, the content validity of teacher-made test content was maintained. It's important to remember that teachers' alignment of their evaluations with the TOS suggests that teachers both in middle and high schools moderately create instruments for formative learning performance evaluation. The findings of research question two revealed the level of test construction skills possessed by Osun state secondary school teachers is high, indicating that they have specific skills required to ensure proper achievement testing competence. Findings showed that most teachers judged themselves as competent in constructing achievement tests. This agrees with the findings of Kissi et al. (2023), which state that secondary school teachers perceived themselves as skilled in constructing teacher-made multiple-choice tests. However, the findings contradict Blasi's (2005) and Agu et al. (2013) observation that teachers lack test construction skills.

Limitation of the Study

The study on the achievement testing competence of secondary school teachers in Osun State, Nigeria, employs a descriptive design survey approach. However, it has several limitations that can affect the findings and conclusions. These include sample size and representativeness, response bias, self-

reported data, limited depth of analysis, temporal constraints, contextual factors, instrument limitations, generalizability, and potential for misinterpretation. The sample size may not accurately represent the entire population of secondary school teachers in Osun State, leading to biased results. Response bias can skew the data, making it harder to assess teachers' competence in achievement testing accurately. Self-reported data can be subjective, leading to overestimating or downplaying teachers' competence. Descriptive surveys often provide quantitative data but lack qualitative depth, which may not fully capture the complexities of teachers' experiences or contextual factors influencing their competence. Temporal constraints may not account for changes in teachers' competence or the educational landscape over time, limiting the validity and reliability of the data collected. Contextual factors, such as school resources, professional development opportunities, and administrative support, may not be adequately addressed, limiting the understanding of the environment in which teachers operate. Instrument limitations, such as unclear questions or a lack of comprehensive coverage of all relevant aspects of achievement testing, could affect the validity and reliability of the data collected. Generalizability may not be possible due to different cultural, economic, and educational environments outside Osun State. Misinterpretation of the results may occur if conclusions are drawn without considering the limitations of the methodology. While the research provides valuable insights, it is crucial to acknowledge these limitations to contextualise the findings and guide future research efforts. Future studies could benefit from a mixed-methods approach, more extensive and diverse samples and a focus on longitudinal data.

CONCLUSION

The study concluded that the test construction skills of Osun state secondary school teachers are high enough to ensure proper achievement testing competence, which was found to be affected by their teaching experience and educational qualifications. The results of the study contribute towards increasing teachers' awareness and practice of testing competence, instigating learners towards every content of the textbook, and assessing whether the content areas to be tested are adequately reflected in the test content or not by reviewing the distribution of practical exercises among the content areas of the textbook, and increasing a closer relationship between teaching and testing. Based on the findings of the study, it is therefore recommended that classroom teachers should ensure that test items constructed are representative of course content to ensure content validity; a table of specification should be developed before constructing achievement tests to measure student's cognitive domains adequately, and in house seminars, workshops and conferences may also be held on teacher competence specifically on classroom testing. This will not only improve the present testing competencies of the teacher but may also enhance overall teacher competence, particularly in private and public secondary schools, local, state, and federal level, tertiary institutions, and most importantly, the policy maker. In this regard, the content validity of middle and high public school tests has been taken as a frame of reference. This study thus did not encompass other diversified areas or test qualities on which effectiveness could be assessed. This study strictly limited itself to evaluating the content validity of teacher-made middle and high public secondary school tests. Further study could be carried out to determine the content validity of private and public secondary school teacher-made tests.

REFERENCES

Adeosun, P. K. (2024). Teachers Competency in Test Construction in Secondary Schools in Ovia North-East Local Government Area Edo State. *Ilorin Journal of Education*, *44*(1), 237-250.

Agu, N. N., Onyekuba, C., & Anyichie, A. C. (2013). Measuring teachers' competencies in constructing classroom-based tests in Nigerian secondary schools: Need for a test construction skill inventory. *Educational Research and Reviews*, *8*(8), 431.

Aina, J. K. (2016). Employment of untrained graduate teachers in schools: The Nigeria case. *Journal of Studies in Social Sciences and Humanities*, *2*(2), 34-44.

Baffoe, J. (2021). *Effects of item sequencing in multiple-choice tests on senior high school student's academic performance in the Kumasi metropolis: the moderating role of gender* (Doctoral dissertation, University of Cape Coast).

Blasi, M. (2005). Standardised tests: a teacher's perspective. *Childhood Education*, *81*(4), 242-244.
Bowes, D., Hall, T., Petric, J., Shippey, T., & Turhan, B. (2017, May). How good are my tests? In *2017 IEEE/ACM 8th Workshop on Emerging Trends in Software Metrics (WETSoM)* (pp. 9-14). IEEE.
Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication methods and measures*, *12*(1), 25-44.
DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage publications.
DiBattista, D., & Kurzawa, L. (2011). Examination of the quality of multiple-choice items on classroom tests. *Canadian Journal for the Scholarship of Teaching and Learning*, *2*(2), 4.
Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational researcher*, *44*(4), 237-251.
Durowojum, E. O. (2014). *Effect of continuous assessment modes on students' learning outcomes in commerce in senior secondary schools in Ibadan* (Doctoral dissertation).
Gareis, C., & Grant, L. W. (2015). *Teacher-made assessments: How to connect curriculum, instruction, and student learning*. Routledge.
Guskey, T. R., & Brookhart, S. M. (2019). *What we know about grading: What works, what doesn't, and what's next*. ASCD.
Hamafyelto, R. S., Hamman-Tukur, A., & Hamafyelto, S. S. (2015). Assessing teacher competence in test construction and content validity of teacher-made examination questions in commerce in Borno State, Nigeria. *Education*, *5*(5), 123-128.
Kaftandjieva, F. (2010). Methods for setting cut scores in criterion-referenced achievement tests. *Cito, Arnhem: EALTA*, *170*.
Kinyua, K., & Okunya, L. O. (2014). Validity and Reliability of Teacher-Made Tests: Case Study of Year 11 Physics in Nyahururu District of Kenya. *African Educational Research Journal*, *2*(2), 61-71.
Kissi, P. (2020). *Multiple-choice construction competencies and items" quality: evidence from selected senior high school subject teachers in Kwahu-South District* (Doctoral dissertation, University of Cape Coast).
Kissi, P., Baidoo-Anu, D., Anane, E., & Annan-Brew, R. K. (2023, April). Teachers' test construction competencies in examination-oriented educational system: Exploring teachers' multiple-choice test construction competence. In *Frontiers in Education* (Vol. 8, p. 1154592). Frontiers Media SA.
Kubiszyn, T., & Borich, G. D. (2024). *Educational testing and measurement*. John Wiley & Sons.
Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: effects on instructional quality and student development. *Journal of educational psychology*, *105*(3), 805.
Mahajan, G. (2015). Construction and validation of achievement test in Economics. *International Journal of Humanities and Social Sciences Studies*, *1*(6), 54-60.
Musah, M. B., Al-Hudawi, S. H., Issah, M., Hussein, A. R., & Ibrahim, M. (2022). Testing content validity of teacher-made test: Profiling teacher perceptions and demographic variables. *International Journal of Evaluation and Research in Education*, *11*(2), 878-887.
Nkemka, A. (2022). *Educators' Descriptions of Academic Engagement Strategies for Recent Sub-Saharan Immigrants* (Doctoral dissertation, Capella University).
Ogunyinka, E. K., Okeke, T. I., & Adedoyin, R. C. (2015). Teacher education and development in Nigeria: An analysis of reforms, challenges and prospects. *Education Journal*, *4*(3), 111-122.
Petters, J. S., & Okon, M. O. (2014). Students' perception of causes and effects of examination malpractice in the Nigerian educational system: The way forward for quality education. *Procedia-Social and Behavioral Sciences*, *114*, 125-129.
Rao, N. J., & Banerjee, S. (2023). Classroom assessment in higher education. *Higher Education for the Future*, *10*(1), 11-30.
Sasu, E. O. (2017). *Testing practices of junior high school teachers in the Cape Coast Metropolis* (Doctoral dissertation, University of Cape Coast).
Snow, R. E. (2012). Construct validity and constructed-response tests. In *construction versus choice in cognitive measurement* (pp. 45-60). Routledge.

Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of educational psychology*, *103*(4), 952.

Ward, W. C., & Bennett, R. E. (2012). *Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment*. Routledge.