

Mathematics teachers' skills in the development of valid and reliable classroom-based tests in secondary schools

Basil Chinecherem Oguguo ^{1*} , Wisdom Kenneth Eze ¹ 

¹University of Nigeria, Nsukka, NIGERIA

*Corresponding Author: basil.oguguo@unn.edu.ng

Citation: Oguguo, B. C., & Eze, W. K. (2026). Mathematics teachers' skills in the development of valid and reliable classroom-based tests in secondary schools. *Journal of Mathematics and Science Teacher*, 6(2), em104. <https://doi.org/10.29333/mathsciteacher/18123>

ARTICLE INFO

Received: 11 Jun 2025

Accepted: 30 Dec 2025

ABSTRACT

Faulty designed test items can result in erroneous assessments of learning and give inaccurate information about both student performance and the efficiency of instruction. The study examined mathematics teachers' skills in developing tests that are classroom-based in secondary schools in Nigeria. An Expo-facto research design was adopted, involving (n = 120) mathematics teachers in which 49 are males and 71 are females. Teachers Test Development Skill Rating Scale (TTDSRS) instrument was used in collecting data. Three experts validated the instrument. The reliability coefficient of the instrument was estimated to be 0.88 using Cronbach Alpha method. The research questions were analyzed using mean and standard deviations and hypotheses tested using t-test statistics at 0.05 level of significance. The result showed that mathematics teachers rarely possess test development skills since the number of items disagreed by teachers is greater than the agreed response, teachers' gender has a slight influence on test development skills in mathematics in favour of the males. Most mathematics instructors lack test-development expertise and may be utilising matched assessments to gauge their students' mathematical proficiency. Mathematics teachers should be given an opportunity for intensive in-service training and workshops on how to improve their test development skills, for a better teaching and learning in schools.

Keywords: mathematics, teacher's variables, test development skills, secondary education

INTRODUCTION

The relevance of mathematics to mankind cannot be overemphasized. It is considered to be the source from which the knowledge of science and technology which are considered essential for societal, economic and national development originates from. Adequate knowledge of mathematics is needed in our contemporary world, which is driven by science and technology. In view of this, students must be provided with adequate foundation and knowledge of mathematics in order to help them sail through the challenges of present day's technological society. This is done by teaching mathematics as a subject in schools. As important as Mathematics is, the tremendous and persistence failure of Nigerian students in Mathematics has remained a major threat to its learning (Abakpa & Agbo-Egwu, 2008; Sanni & Ochepe, 2002; Uloko & Imoko, 2007). Most students appear to be of moderate level of achievement in mathematics while a few of them are either in the category of low or high achievers (Sanni & Ochepe, 2002). This makes one wonder if the since these students are usually provided with learning instructions uniformly. Perhaps, it is possible to think that a factor or some number of factors could be responsible for these observed differences in students' achievement in mathematics. This unsatisfactory state of students' achievement has raised a lot of questions among educators and researchers, but with hardly satisfactory answers (Kushwaha, 2014).

These poor achievements in Mathematics can be attributed to a lot of factors. Asikhia (2015) categorizes factors militating against good academic achievement to include physical, health and emotional as well as the child's interest. Tshabalala (2013) highlighted factors such as lack of qualified teachers, inadequate instructional materials as well as non-proper usage of test development skills to construct test items by the teacher are the reasons for poor achievement of students in mathematics. In the same vein, Kelly (2010) stated that the major reason for poor achievement is that mathematics teachers lack the necessary skills for constructing questions. Kelly further stated that Mathematics teachers choose to construct questions that were not taught or not contained in the scheme of work which were supposed to be taught. These could make the students answer questions that were not stated properly which invariably affect their achievement. Students' poor academic achievement in Mathematics has been attributed to teachers' lack of test development skills (Imo, 2012).

In order for a teacher to achieve the goal of teaching and learning effectively, such a teacher is required to have the skill of test development. Testing determines the effectiveness of teaching and learning. This can be achieved using reliable and valid test items. A test as defined by Ovat and Ofem (2017) refers to any instrument that is used to carefully collect responses from a testee

or an examinee so as to ascertain whether or not a particular ability is present. In essence, a test could be considered as an instrument used for making judgement on students' learning in school. It helps to measure the extent to which a learner has been able to master previously identified objectives. Such objectives may be concerned with cognitive attainment, personality traits such as interest, attitude, social adjustment, or psychomotor skills. There are basically four steps to be followed when constructing a test, these are identifying the purpose of the test, constructing a table of specification (TOS), selecting appropriate item types and writing or preparing items (Ovat & Ofem, 2017). Beyond these steps, scholars have identified other steps which are of paramount importance. These steps according to Olatunji and Onofeghara (2013) range from content analysis, purpose of the test, table of specification, item writing, face validation and item review. As the authors noted, in order to properly design an instrument that is valid and relevant for cognitive domain, test development processes, particularly, construction of table of specification must be followed. Test development skills requires the following steps of conducting a content analysis instructional objectives review, designing of table of specification, writing the items, subjecting the instrument to face validation, reviewing the items, pilot testing or trial testing, conducting an item analysis, selecting relevant items, assembling the items together, final administration of test/norming, designing test manual, final production of the test and marketing of the test as well as its manual (Nworgu, 2015).

Feedback gotten based on testing could facilitate decisions in education. This may be in terms of level of knowledge attainment of learners or the level of progress in learning programme that has been made. Regardless of the information required, valid and trustworthy measures must be used to provide the decision-makers with the information they need. Achievement tests are useful for assessing crucial elements of discipline, appropriately reflect the attention placed on crucial elements of education, and gauge the appropriate degree of a student's knowledge in a particular academic discipline (Frey, 2010). These exams can be either classroom-based (teacher-made tests), which are created by a specific instructor specifically for his or her class, or standardized (tests created by test experts for more general reasons). The majority of accomplishment tests, however, are conducted in the classroom and, when correctly designed, give teachers accurate and helpful information on the knowledge that their pupils have retained in certain academic topics (Agu et al., 2013).

The majority of the confusing and deceptive questions on the tests used for end-of-term exams and classroom assessments in secondary schools may be the cause of some students' failures on these exams. Assessment is a difficult task; as a result, the knowledge on the approaches and mastery of assessment strategies is relevant for an efficient classroom assessment. Therefore, skillfulness and sound training is required on the part of teachers in order to ensure smooth and proper application of classroom assessment (Oguguo et al., 2021). This suggests that most teachers lack test-development expertise and may be utilizing poor questions to gauge students' progress in various academic topics (Gullickson, 2015). The teachers and school administrators will not be able to give educational opportunities and support each individual student's requirements if the accomplishment levels of the students are not accurately evaluated and interpreted. This poor knowledge of the test development skills in mathematics could be attributed to some teacher variables such as gender, qualification and experience. This study assessed the influence of teachers' variables on test development skills in developing classroom-based mathematics tests. This study is very important and will be of great value to the mathematics teachers and researchers across the globe. This is because it will add to the existing body of educational research literature which will be harnessed for future research in this area, as well as providing empirical data for future research in related areas. The study addressed the following research questions and hypotheses:

RQ1 What are the test development skills possessed by mathematics teachers?

The following hypotheses were tested at 0.05 level of significant.

H1 Gender does not significantly influence mathematics teachers test development skills.

H2 Mathematics teachers' year of experience does not significantly influence test development skills.

H3 Mathematics teachers' qualification does not significantly influence test development skills.

LITERATURE REVIEW

The majority of achievement assessments used in secondary schools in Nigeria are created by teachers. Therefore, in order to create test items that would provide reliable feedback on students' success, teachers must apply a certain level of acceptable test development abilities. Skills in test development include those required to create high-quality test items based on predetermined test development standards. Ujah (2010) identifies these abilities as objectivity, communication, item validation, and the ability to use appropriate procedures to determine the validity of test instruments. According to Koksai (2004), a teacher needs the skill to create assessments with accuracy, acceptable language use, objectivity, and appropriate grading systems. To create valid and accurate tests, teachers do not necessarily need to be experts in educational measurement and evaluation, but there are several fundamental test building abilities that every teacher should have. These talents support teachers in creating tests that are acceptable for students of varied ages, abilities, and genders, structuring things to elicit clear and succinct answers from students, setting exams so that students finish within the allotted time and do not become test-phobic. There are issues with how teachers test their students, according to observations (Osadebe, 2013). According to the authors, erroneous teacher-made test items that are hastily set and given to students without following the proper test development and validation procedures are frequently used to test pupils. It is obvious that a good number of the items of these tests end up not measuring what they are designed to measure. In essence, majority of tests designed by the classroom teachers are not valid.

Teachers who lack test-development abilities may give students' accomplishments a wrong assessment. According to some academics, such as Agu et al. (2013), this lack of test-development expertise among teachers is a significant contributor to cheating on school exams by both teachers and pupils in Nigerian secondary schools. Several studies have claimed that inaccurate test

items impair students' comprehension and ability to respond accurately to the questions, and that this compromises the inference made about what a student knows and understands (Agu et al., 2013).

The effectiveness of a test administered by a teacher is directly related to its capacity to deliver the necessary data regarding students. A well-written test enables the teacher to precisely and consistently assess the level of student understanding of a particular subject matter covered in class. These test results give teachers a partial indication of how effective their lessons have been. On the other hand, badly designed test items can result in erroneous assessments of learning and give inaccurate information about both student performance and the efficiency of instruction. The usefulness of a test item is diminished by any features that draw the test taker's attention away from the main idea or subject (Frey, 2010). Classroom teachers should have the competence in the listed areas to create quality test items (Sireci, 2013, p. 23); figuring out the purpose of each testing exercise, stating clear, quantifiable educational objectives, creating a strong content outline, creating a test plan that will guide the development of test items, selecting appropriate test item formats, developing clear, precise, and unambiguous items, developing items that focus a group of students' attention on a single concept even though they have a wide range of varying background experiences, developing items with appropriate difficulty, and developing items with discriminative indices; providing clear instructions on how the test should be administered and taken, developing a marking guide appropriate for the test, performing item analysis of their test items, developing tests that are economical with time and resources, and reviewing the test to correct any errors made during item development.

In addition to the list of abilities, Koksai (2004) pointed out that it is crucial for the teacher to inquire about the task's clarity while creating a test. Can test-takers come up with the right answer even when they don't possess the talent that is supposed to be tested? Have participants ample time to complete the task(s)? According to Quansah and Amoako (2018), test development is the main cause of concern for many instructors in Nigerian schools, particularly those who are inexperienced. This concern primarily results from these teachers' inability to construct tests. This is unexpected considering that the least qualified teacher now employed in secondary schools in Nigeria holds a Nigerian Certificate in Education (NCE). After three years of study in a college of education, where teachers are trained in various education courses, including educational testing and measurement, one can earn this NCE qualification. This insufficiency could be caused by a variety of factors. It could be because there was not enough emphasis placed on assessment during the teachers' pre-service training, or it could be because there are gender differences in test development issues, or it could be because the majority of the teachers failed to acquire the test development skills necessary for producing high-quality test items during their pre-service training. It is important to evaluate the test development abilities of secondary school teachers.

Some of the factors that affect teachers' skills in test development are: Gender (Akpochofo, 2015), qualification (Fajonyomi, 2011), experience (Fehintola, 2014). Gender is one of the key elements that affects teachers' abilities to construct tests, according to Akpochofo (2015). It is possible to describe gender as the socially and culturally produced traits, features, behaviors, and roles that various civilizations assign to males and females (Imogie & Eraikhuemen, 2011). According to Eze and Okoro (2010), gender can be viewed as the socially dictated variations in the mental and emotional traits of the two sexes. Gender can be viewed in the context of this study as a socially constructed set of obligations and roles that come with being male or female. Teachers' professional behavior, effectiveness, and job performance may be impacted by their gender. It appears that most male teachers chose to teach because it is the most common profession. It has been noted that their responsibilities as the family's earner and head may have an impact on their teaching abilities. It appears most primary and secondary school instructors occasionally work on other projects that interfere with the time they should be using to create effective tests. This therefore affects the efficiency or effectiveness of their instruction.

The debate over teacher certification has been ongoing for many years. According to Ngada (2010), the success or failure of any educational program largely depends on the availability of sufficient numbers of qualified (professional), capable, and committed teachers. According to Akpochofo (2015), a teacher's capacity to instruct students is based on their exceptional pedagogical talent rather than only their academic background. He also argued that the quality of the instructor will unavoidably be reflected in the inhabitants of tomorrow. When commenting on the quality of instructors, Ngada (2010) noted that more than 80% of survey participants believed that teachers were bearers of flaws. These flaws include, among others, a lack of professional training, poor classroom management and control, and a lack of depth in the subject area.

There aren't many qualified teachers working in schools. The causes might not be absurd. Regardless of their area of specialization, teaching is viewed as a dumping ground for any unemployed school graduates. Thus, this group of physically fit young people handles the work as a side project to achieve their goals. As a result, they would contribute very little to the work because they lack the commitment required for the position (Ngada, 2010). The handful who appears to be uncommitted lack the necessary technical expertise to instruct because they never received on-the-job training. The subsequent impact on the students' academic performance is disastrous. The main harm caused by this is pupils who frequently perform poorly on examinations and have mediocre understanding.

Years of teaching experience is a factor that is connected to instructors' gender and qualifications. Learning by doing effect, in particular, is particularly pronounced in the early years of teaching, according to several studies that have demonstrated a favorable relationship between experience and teacher effectiveness. Experienced teachers should be more successful than inexperienced teachers in creating better test items if teachers' learning accumulates over a longer period of teaching practice (Ngeche, 2017). The adage "experience is the best teacher" has long been held true, and numerous empirical investigations have demonstrated a large and favourable correlation between teachers' years of experience and test development. In order to identify which of the predictor variables were statistically significant, Adesina et al. (2016) looked at the association between a few different characteristics and students' academic achievement in Oyo State, Nigeria. The authors discovered a strong correlation between teacher experience and student achievement. According to research by Ewetan and Ewetan (2015), Fehintola (2014), Yara and

Wanjohi (2011), and Adeyemi (2005), instructors' experience and educational background are the best indicators of how effective they will be at developing tests. However, Brookhart (2011) discovered that the effectiveness of teachers was not significantly influenced by their educational background or prior teaching experience. Therefore, its inclusion in the current study is necessary due to the inconclusiveness of the predictive and related effect of teachers' years of teaching experience.

The creation of tests needs the application of abilities that can help teachers create tests with accuracy, acceptable language use, objective communication, item validation, and appropriate grading scales (Ujah, 2010). To ensure that items are constructed to elicit clear and bring replies suited to the learner's age, abilities, and other observable distinctions, teachers must develop general test creation skills. A poor and inaccurate assessment of student achievement may come from this lack of test development expertise. According to Agu et al. (2013), the poor test design has encouraged academic dishonesty and examination fraud in the majority of Nigerian secondary schools. To prevent systematic errors, the instructor alone is responsible for creating and using this instrument to make sure that what is taught and the test questions are in line with the course materials. Chan (2009) observed that schoolteachers give little thought to the creation of trustworthy evaluation tools. Many have theorized that the wrong production of a test instrument is caused by teachers' ignorance or lack of awareness, their lack of experience in creating a test blueprint, and gender inequalities in concerns over tables of specification.

Various research has shown consistent poor academic achievement among secondary school students in mathematics as a result of lack of test development skills by mathematics teachers. This has become a major point of concern to students, teachers, educational administrators as well as other stakeholders in the field of education. Unfortunately, test development is a major source of anxiety among mathematics teachers in secondary schools. These anxieties are because of poor knowledge of the relevance of the skills involved in the development of tests by the mathematics teachers. This has resulted in invalid and unreliable instruments for testing students and could lead to poor performance on the subject. This poor knowledge of the test development skills in mathematics could be attributed to some teacher variable factors such as gender, qualification and experience. The research findings can inform educational stakeholders with regard to teachers' professional development on assessment.

Theoretical Framework

This study was anchored on Classical Test Theory (CTT). The theory was propounded by Charles Spearman (1863-1945). Classical Test Theory (CTT), sometimes called the true score model, is the mathematics behind creating and answering tests and measurement scales. The aim of CTT is to improve tests, particularly the reliability and validity of tests. The equation for the classical test model is $X = T + E$. This is the starting point of classical test theory. CTT postulates that an observed score (X) is decomposed into a true score (T) and a single undifferentiated random error (E). It is a deterministic model for minimizing the error of measurement of a test. The following assumptions underline CTT:

- True score and error score are uncorrelated
- The average error score in the population of examinees is zero
- Error scores in parallel tests are uncorrelated

The findings of this study would hopefully be theoretically useful in the sense that the theoretical information of CTT which revealed that everyone has a true score which would be obtained if there were no errors in measurement may be better understood when used in line with this study. This ensures that items are not only valid and reliable but also fair to across the subgroup of examinees. In relation to the present study, Classical Test Theory was of the view that constructing test using the test development skills makes the test standard. This implies that if mathematics teachers employ the use of test development skills in developing test items, the test will be standardized, and its validity will be certain. The findings of this study have strengthened the classical test theory because mathematics teachers will get empirical information that will help them in test development skills and construct mathematics test items for better student academic achievement.

MATERIAL AND METHOD

Study Design

An Ex-post facto research design was adopted for this study. In this design, as Nworgu (2015) state, the researchers do not have the ability to manipulate variables of interest as events had already taken place before data are collected. This design is appropriate for this study because the teacher's variables (gender, qualification and experience) cannot be manipulated and are already influencing teachers' test development skills. The Test Development skills covered in the study include Content analysis; review of instructional objectives; development of test blueprint or table of specification; item writing; face validation; item review; trial testing; item analysis; item selection; test assembly. The study investigated teachers' test development skills in mathematics. It is only a sample of teachers from the area under study that was used for the study.

Sampling and Participants

The participants consist of ($N = 120$) mathematics teachers from 49 public owned secondary schools in Sought Eastern, Nigeria which by gender (Male = 49 and Female = 71), by experience, 61 were between 0-15 years and 59 were above 15 years and by qualification, 48 were qualified while 72 were unqualified from all the 49 government secondary schools. There was no sampling because the entire mathematics teachers participated in the study. The study was conducted at the public owned secondary schools in sought eastern, Nigeria. The survey response data was retrieved in December 2023, and it lasted for two (2) weeks through questionnaire administered to each mathematics teacher face-to-face.

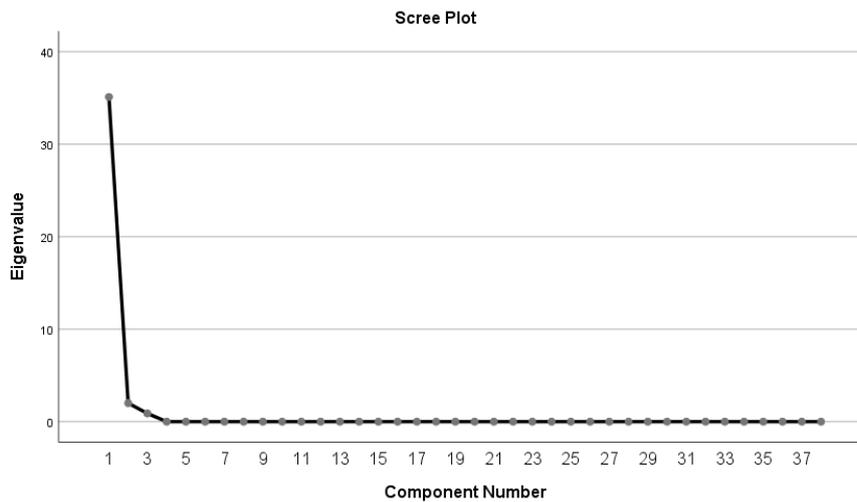


Figure 1. Scree plot of Varimax rotation method of factor analysis of the TTDSRS (Source: Authors' own elaboration)

Research Instrument

The researchers adapted the "Teachers Test Development Skill Rating Scale (TTDSRS)" developed by (Agu et al., 2013) for assessing teachers' skills in the development of classroom-based tests in mathematics. The original instrument contains 30 items with four (4) response options. In the process of adapting the instrument, 23 items were modified and retained. Seven (7) items were not included as they were considered irrelevant to the present study. Also, some item statements were added to the instrument so as to include other test development skills which were not captured in the adapted instrument. The modified instrument consists of 38 items for rating teachers' test development skills. The weight for each item was 4-point Likert type of rating scale with values of 4, 3, 2 and 1 respectively and there was reversal on negative sentences.

Validity and Reliability Procedures

TTDSRS was validated by experts. The experts were senior lecturers in the fields of Measurement and Evaluation, Mathematics Education, and Psychology of Education at the University in Nigeria, Nsukka. The experts were requested to validate the instrument in terms of wordings of the items, the language, clarity of the items and the appropriateness of the items to the purpose of the study. The constructive criticisms of the experts were used to produce the final version of TTDSRS for the study. Further, data collected on the instrument was subjected to construct validation by means of Varimax rotation method of Factor Analysis. The 38 items in the instrument loaded purely on Factor 1, see Scree Plot (**Figure 1**). The reliability of the instrument was determined by administering 30 copies of TTDSRS on 30 Mathematics teachers who were not part of the main study but shared similar characteristics that are necessary for the study. The internal consistency of the instrument was established using Cronbach's Alpha reliability method and the reliability index of 0.88 was obtained. Cronbach alpha was considered appropriate because the items were polytomously scored.

Data Analysis

The data generated through the research instrument was analyzed using SPSS version 23.00 to conduct the statistical analysis. This analysis involves the use of descriptive and inferential statistics [mean and standard deviation, t-test]. These analysis techniques were chosen because this study assessed mathematics teacher's skills in the development of classroom-based tests. The benchmark for each item of the questionnaire was based on a mean score of 2.50 and above. The t-test statistics were used for testing the hypotheses at a level of significance of 0.05. The null hypotheses are rejected if the associated exact probability is less than 0.05, otherwise do not reject.

RESULTS

Participants Statistics

Figure 2 shows the sample size distribution of the participants made up of 120 Mathematics teachers drawn from 49 public secondary schools in Sought East Nigeria. The chart categorised the participants into three (Gender, Experience & Qualifications). **Figure 2** shows that female mathematics teachers are more in number than male mathematics teachers, by experience 61 mathematics teachers have 0-15years of working experience while 59 mathematics teachers have 15years of experience and above and by qualification only 48 mathematics teachers were qualified while 72 teachers are not qualified based on the facts that they do not possess teaching qualifications.

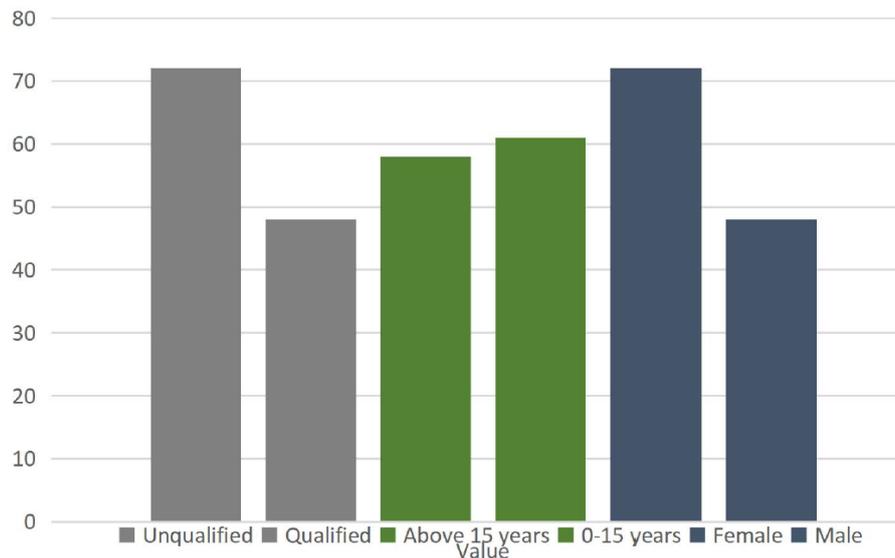


Figure 2. Analysis of the participants (Source: Authors' own elaboration)

Table 1. Test development skills possessed by mathematics teachers

S/N	Test development skills	Mean	Std D
1	Before designing a test, contents I have covered are often outlined	2.59	0.95
2	I do develop my test based on the content covered	2.63	0.94
3	I determine the objective(s) of my test before developing the test	2.54	0.92
4	Test items are often matched with the objectives from the syllabus	2.65	0.88
5	Table of specification is a guide I use in test development	2.33	0.92
6	I just keep my table of specifications for myself	2.16	0.88
7	I determine the test type to give before developing my tests	2.59	0.92
8	I make use of only one test type	2.51	0.89
9	I write more items than will actually be used while developing test	2.54	0.96
10	I use flamboyant and ambiguous words while writing my items	2.43	0.98
11	In developing my items, I give clues to the right answers	2.35	0.94
12	Items develop are usually vetted by the Head of Unit or Dean	1.98	0.71
13	Tests I prepare for promotional exams are often submitted to experts for review on time	2.53	0.90
14	I give my test to other colleagues to help me vet	1.58	0.51
15	I do review the items after face validation	2.51	0.96
16	I prepare a test draft before developing my test	2.43	0.95
17	I trial-test the draft before finally using it	2.40	0.91
18	I subject test items (objectives) to item analysis	1.98	0.89
19	I often consult tests I designed before and adapt questions from them	2.33	0.92
20	I prepare my test long before its administration to avoid hurrying	2.14	0.86
21	I have question bank that I can be referred to	2.48	0.93
22	I use text books related to the subject area as guide	2.50	0.89
23	I ensure test items are logically organized	2.14	0.96
24	I usually outline clear instructions for test takers	2.53	0.98
25	I write items to cover both low and high achievers	2.64	0.93
26	I consider time while designing tests available for testing	1.98	0.71
27	I ensure items covers all the levels of cognitive domain	2.52	0.90
28	Each item of a test is assigned a scores	2.57	0.49
29	I make sure the set objectives are measured by the items	2.50	0.95
30	I try to elicit creative response from students when I set essay items	2.42	0.94
31	While developing a test, I often prepare a scoring guide	2.64	0.92
32	I take note of learners' age when writing items	1.98	0.72
33	I cover instructional units adequately by adding sufficient items	2.33	0.91
34	I make sure test drafts are reviewed at least, twice before administration	2.14	0.86
35	Long question and phrases are usually avoided	2.49	0.92
36	I design test that the items elicit specific information at a time	2.51	0.89
37	I carry out reliability test of the test developed	2.16	0.95
38	I ensure that items of multiple-choice test don't give clue of answer	2.51	0.89
	Overall mean	2.37	

Test Development Skills Possessed by Mathematics Teachers

The result in **Table 1** shows the mean and standard deviation of the test development skills possessed by mathematics teachers. It shows that most of the test development skills possessed by mathematics teachers are items 1, 2, 3, 4, 7, 8, 9, 13, 15,

Table 2. Influence of teachers' gender on test development skills

Teachers' gender	N	\bar{X}	SD	t-val	Df	P	Dec.
Male	9	9.82	2.64	1.78	118	0.07	NS
Female	71	86.03	10.61				

NS = Not Significant

Table 3. Influence of teachers' year of experience on test development skills in mathematics

Years of experience	N	\bar{X}	SD	t-val	Df	P	Dec.
0-15 years	61	85.52	11.66	-2.00	118	0.04	S
Above 15 years	59	89.69	11.19				

S = Significant

Table 4. Influence of teachers' qualification on test development skills

Teachers' qualification	N	\bar{X}	SD	t-val	Df	P	Dec.
Qualified	48	89.61	11.19	1.15	118	0.04	S
Non-qualified	72	87.08	12.23				

S = Significant

22, 24, 25, 27, 28, 29, 31, 36 and 38, since their mean responses to the items are greater than or equal to 2.50 benchmark mean. It can also be seen that mathematics teachers disagree to item 5, 6, 10, 11, 12, 14, 16, 17, 18, 19, 20, 21, 23, 26, 30, 32, 33, 34, 35 and 37. This implies that mathematics teachers rarely possess test development skills since the number of items disagreed by teachers is greater than the agreed response. The overall mean is 2.37. Since it is less than the benchmark of 2.50, it implies that mathematics teachers generally do not possess test development skills.

Influence of Mathematics Teachers' Gender on Test Development Skills

Table 2 shows the influence of teachers' gender on test development skills in mathematics. The result showed that male teachers had a mean test development skills score of 9.82 with a standard deviation of 2.64 while their female counterparts had a mean test development skills score of 86.03 with a standard deviation of 10.61 in mathematics. The mean test development skills score for male teachers is slightly higher than that of their female counterparts. This implies that teachers' gender appears to have a slight influence on test development skills in mathematics in favour of the males. In addition, the results indicated that teachers' gender had a t-value of 1.78 and an associated probability value of 0.07 with regard to test development skills in mathematics. The null hypothesis one (H_{01}) was not rejected because the associated probability value of 0.07, when compared with 0.05 defined as the level of significance, was judged not significant because it was greater. As a result, it was determined that teachers' gender had no discernible impact on test development skills in mathematics.

Influence of Mathematics Teachers' Years of Experience on Test Development Skills

Table 3 shows the influence of teachers' years of experience on test development skills in mathematics. The result showed that 0-15years of teaching experience had a mean test development skills score of 85.52 with a standard deviation of 11.66 while above 15years of teaching experience had a mean test development skills score of 89.69 with a standard deviation of 11.19 in mathematics. The mean test development skills score for above 15years is slightly higher than that of 0-15years of teaching experience. This implies that teachers' years of teaching experience appears to have an influence on test development skills scores in mathematics in favour of the teachers with above 10years of teaching experience. Regarding the impact of instructors' year of teaching experience on test improvement skills in mathematics, the result likewise showed that a t-value of -2.00 with an associated probability value of 0.04 was achieved. The null hypothesis two (H_{02}) was rejected since the associated probability value of 0.04 was found to be significant when compared to the level of significance of 0.05. Therefore, the conclusion drawn was that there was a significant influence of teachers' year of teaching experience on test development skills in mathematics.

Influence of Mathematics Teachers' Qualification on Test Development Skills

Table 4 revealed the influence of teachers' qualification on test development skills in mathematics. The result showed that qualified (B.Ed, NCE, B.Sc with PGDE) teachers had a mean test development skills score of 89.61 with a standard deviation of 11.19 while unqualified (B.Sc/HND/OND) teachers had a mean test development skills score of 87.08 with a standard deviation of 12.23 in mathematics. The mean test development skills score for qualified teachers is higher than that of non-qualified teachers. This implies that teachers' qualifications appear to have a slight influence on test development skills scores in mathematics in favour of the qualified teachers. A t-value of 1.15 with an associated probability value of 0.04 was obtained with regards to teachers' qualifications on test development skills in mathematics. Since the associated probability value of 0.04 when compared with 0.05 set as the level of significance was found to be significant. For this reason, the null hypothesis three (H_{03}) was rejected. Therefore, the conclusion drawn was that there was a significant influence of teachers' qualification on test development skills in mathematics.

DISCUSSION

The result revealed that mathematics teachers rarely possess test development skills since the number of items disagreed by teachers is greater than the agreed responses. Among all the items in the instrument for the study, most of the test development

skills possessed by mathematics teachers are item 1, 2, 3, 4, 7, 8, 9, 13, 15, 22, 24, 25, 27, 28, 29, 31, 36 and 38. This implies that mathematics teachers do not possess test development skills on item 5, 6, 10, 11, 12, 14, 16, 17, 18, 19, 20, 21, 23, 26, 30, 32, 33, 34, 35 and 37. The outcome is in line with Gichuhi's (2014) research, which showed that secondary school teachers do not consistently use Bloom's cognitive levels of objectives when creating their exam items. These results support Amaechi and Wakjissa's (2016) findings that teachers have some familiarity with content-related evidence, coverage methodologies, and proper sampling of content and objectives but lack this expertise for other test development techniques. This indicates that test development skills are rarely possessed by teachers. Teachers should be urged to consistently seek out new information and put their development abilities to the test by attending workshops, seminars, and conferences.

The outcome showed that teachers' gender had no significant impact on test development in mathematics. This study supports Izard's (2005) assertion that the issue does not relate to a person's gender but rather to their capacity to acquire and apply test development abilities. This is the case because most teachers, regardless of gender, have little to no experience using this tool for cognitive assessment, which accounts for the inadequate development of learner evaluation instruments that result in poor performance and other associated assessment-related repercussions. Furthermore, Ovat and Ofem (2017) found no evidence of a gender effect on secondary school teachers' use of the test blueprint. This implies that there are no differences among male and female teachers' disposition to test development skills in mathematics.

Teaching experience have an influence on test development skills scores in mathematics in favour of the teachers with more than 10 years of teaching experience. The study further revealed significant influence of teachers' years of experience on test development skills in mathematics. This result was at odds with Chan's (2009) report, which claimed that actual instrument usage, rather than years of experience is what counts. The majority of teachers, according to Chan, are aware of the importance of test development skills and are informed about them, but the issue is how they see the development and use of the instrument at the proper time. Teachers view these test development processes as a waste of time and unmotivating (Downing, 2003). Because of this, even though they have the necessary experience, they never use it.

Teachers' qualifications have significant influence on test development skills in mathematics. The findings of this study have revealed that mathematics teachers with professional teaching qualifications in Education such as B.Ed and NCE possess test development skills than non-professional teachers without teaching qualifications. This suggests that the qualifications of teachers have a major impact on their mathematical test development skills. This is consistent with the findings of Wiki (2013), Abe and Adu (2013), who discovered that teachers' qualifications improved pupils' academic scores. This result is consistent with research by Van den Bergh et al. (2014), who found that students taught by graduates with teaching qualifications outperformed those taught by graduates without teaching qualification in terms of academic achievement. This is a change from the adage that no one can provide what they do not have; therefore, excellent instructors continually endeavor to have the training and information necessary to carry out their tasks successfully and efficiently.

Implications for Practice

Most mathematics instructors lack test-development expertise and may be utilizing shoddy assessments to gauge their students' mathematical proficiency. The teachers and school administrators will not be able to give educational opportunities and support each individual student's requirements if the accomplishment levels of the pupils are not accurately evaluated and interpreted. The marginal effect of teachers' gender on test development in mathematics, favoring males, was not statistically significant. This suggests that there are no gender disparities in the willingness of male and female teachers to assess students' mathematical development. Also, the findings of this study indicated that there was significant influence of teachers' years of teaching experience on test development skills in mathematics. This implies that students taught by more experienced teachers on test development skills in mathematics achieve higher levels, because their teacher has mastered the test development procedures. The qualifications of teachers have a major impact on mathematical test development skills. As a result, teachers must constantly look for ways to enhance their subject-specific knowledge, instructional strategies, and test-development abilities by participating in some kind of on-the-job in-service training.

Limitations

It is pertinent to note that caution must be taken when drawing absolute conclusions based on the findings of this study for some reasons. The study data relied on questionnaire which are self-report instruments, capable of generating subjective response. Thus, findings of this study may be further verified using interviews, open-ended questions and focused group discussions. The research was carried out in government owned secondary schools, and this may affect the generalizability of the findings of the study across other school types.

CONCLUSION

The findings of this study revealed that mathematics teachers rarely possess the necessary skills in developing classroom-based tests since the number of items disagreed by teachers is greater than the agreed response. Teachers' gender has a slight influence on test development skills in mathematics in favour of the males. Although, the corresponding null hypothesis revealed that teachers' gender does not significantly influence their test development skills in mathematics. Teachers' years of teaching experience appear to have an influence on test development skills scores in mathematics in favour of the teachers with more than 10 years of teaching experience. Furthermore, the corresponding null hypothesis revealed that there was a significant influence of teachers' year of teaching experience on test development skills in mathematics. Teachers' qualification has an influence on test development skills in mathematics in favour of the teachers with B.Ed/NCE qualification. Furthermore, the corresponding null

hypothesis revealed that teachers' qualifications have significant influence on test development skills in mathematics. There should be a reorientation of mathematics teachers to use test development skills in developing classroom-based tests. This method may help mathematics teachers to standardize students' tests and improve their test development skills. The study will add to the existing body of literature and assist the mathematics test developers in reviewing test items occasionally and utilize the test development skills to ensure a fair and standardized mathematics test.

Author contributions: BCO: Conceptualization, data curation, formal analysis, methodology, project administration, resources, validation, visualization, writing – original draft, writing – review and editing. **WKE:** Conceptualization, data curation, formal analysis, methodology, project administration, resources, validation, visualization, writing – original draft; writing –review & editing). All authors have agreed with the results and conclusions.

Acknowledgements: The researchers are very grateful to all their colleagues and the authors whose studies were made use of during the course of this study.

Funding: The authors did not receive any funding/support from any person or organization for the submitted work.

Ethical statement: The authors stated that the study was approved by the institutional ethics committee of University of Nigeria, Nsukka, Nigeria on 27th November, 2024 (UNNEC/EDU/SE/VII/275). Before the commencement of the study, the respondents were presented with informed consent forms to fill and sign. The informed consent forms were properly filled and signed. This research was carried out in accordance with the principles outlined in the ethical policy of academic publishing. The data collection from the entire sample was in accordance with the ethical standards of the researchers' institutional research committee and with the 1964 Helsinki declaration.

AI statement: The authors stated that, during the preparation of this work the author(s) did not use generative AI and AI-assisted technologies in the writing process of this manuscript.

Declaration of interest: The authors declare no conflicts of interest, both financial and non-financial interest regarding the publication of this paper.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

- Abakpa, B. O., & Agbo-Egwu, A. O. (2008). The effect of small group co-operative learning on students' achievement and retention in mathematics achievement tests. *Benue Journal of Research in Science and Science Education*, 1(1), 71-80.
- Abe, T. O., & Adu, E. I. (2013). Influence of qualification on development and assessment of computer programmed instructional package on energy concept in upper basic technology in Ekiti State. *Journal of Science and Technology*, 3(6), 611-618.
- Adesina, J. O., Raimi, S. O., Bolaji, O. A., & Adesina, A. E. (2016). Teachers' attitude, years of teaching experience and self-efficacy as determinants of teachers' productivity in teachers' professional development programme in Ibadan Metropolis, Oyo State, Nigeria. *Journal of Emerging Trends in Educational Research and Policy Studies*, 7(3). <https://hdl.handle.net/10520/EJC196815>
- Adeyemi, T. O. (2005). Teachers' teaching experience and students' learning outcomes in secondary schools in Ondo State, Nigeria. *African Journal of Educational Studies in Mathematics and Sciences*, 3(1), 89-99. <https://doi.org/10.4314/ajesms.v3i1.38607>
- 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), 431-439.
- Akpochofa W. P. (2015). Gender difference among social studies teachers' competences in the use of the inquiry method in South-South Nigeria. *Journal of Education and Practice*, 6(23), 19-24.
- Amaechi, U. C., & Wakjissa, S. G. (2016). What teachers know about validity of classroom tests: Evidence from a University in Nigeria. *Journal of Research & Method in Education*, 6(3)14-19.
- Asikhia, O. A. (2015). Students and teachers' perception of the causes of poor academic in Ogun State secondary schools [Nigeria]: Implication for counseling for national development. *European Journal for Social Sciences*, 13, 229-242.
- Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3-12. <https://doi.org/10.1111/j.1745-3992.2010.00195.x>
- Chan, K. K. (2009). Using test blueprint in classroom assessments: Why and how. In *The 35th International Association for Educational Assessment (IAEA) Annual Conference*. http://bo.io.gov.mo/bo/i/2006/52/lei09_cn.asp
- Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education*, 37(9), 830-837. <https://doi.org/10.1046/j.1365-2923.2003.01594.x>
- Ewetan, T. O., & Ewetan, O. O. (2015). Teachers' teaching experience and academic in mathematics and english languages in public secondary schools in Ogun State, Nigeria. *International Journal of Humanities Social Sciences and Education*, 2(2), 123-134.
- Eze, R. O., & Okoro, I. F. (2010). Comparative analysis of the performances of male and female student using problem-solving and expository method of teaching. *Journal of Curriculum Studies*, 17(2), 77-84.
- Fajonyomi, A. A. (2011). Reviving higher education relevance in Nigeria. In *The 3rd Faculty of Week and Annual National Conference, Faculty of Education*. University of Maiduguri.
- Fehintola, J. O. (2014). Teachers' characteristics as correlates of students' academic among secondary school students in Saki-west local government area of Oyo State. *Journal of Educational and Social Research*, 4(6), 459-468. <https://doi.org/10.5901/jesr.2014.v4n6p459>
- Frey, B. (2010). Teachers' classroom assessment practices. *Middle Grades Research Journal*, 5, 107-117.

- Gichuhi, C. (2014). Teachers' competence in tests construction within blooms taxonomy for effective learning assessment: A case study of Kikuyu District, Kiambu County [Research Project]. University of Nairobi, Kenya. <http://hdl.handle.net/11295/76048>
- Gullickson, A. R. (2015). Teacher perspectives of their instructional use of tests. *The Journal of Educational Research*, 77(4), 244-248. <https://doi.org/10.1080/00220671.1984.10885532>
- Imo, G. C. (2012). Effects of training in test construction on quality of teacher-made tests and students' physics achievement in secondary schools in Plateau State, Nigeria [Unpublished Ph. D Thesis]. University of Jos.
- Imogie, A. O., & Eraikhuemen, L. (2011). An inquiry into sex differentiation in admission and academic at the University of Benin, Nigeria. *Benin Journal of Gender Studies*, 1(1),1-13.
- Izard, J. (2005). Overview of test construction. In K. N. Ross (ed.), *Quantitative research methods in educational planning*. UNESCO International Institute for Educational Planning. <https://unesdoc.unesco.org/ark:/48223/pf0000214553/PDF/214553eng.pdf.multi>
- Kelly, T. R. (2010). The causes of poor in mathematics among senior secondary school students in Dutse Metropolis of Jigawa State, Nigeria.
- Koksal, D. (2004). Assessing teachers' testing skills and enhancing their professional development through distance learning. *Turkish Onlines Journal of Distance Education*, 5(1).
- Kushwaha, S. S. (2014). Trend in researches on mathematics achievement. *IOSR Journal of Research & Method in Education*,4(6), 53-62. <https://doi.org/10.9790/7388-04625362>
- Ngada, A. J. (2010). Impact of NCE programme on preparation and of NCE graduate teachers in Borno state Junior Secondary Schools. *Research in Curriculum*, 5(1), 190-195.
- Ngeche, T. N. M (2017). Student and teacher attitude as correlates of performance in mathematics in Cameroon secondary schools. *International Journal of Humanities Social Sciences and Education*, 4(12), 1-10. <http://doi.org/10.20431/2349-0381.0412001>
- Nworgu, B. G. (2015). *Educational measurement and evaluation. Theory and practice*. University Trust Publishers
- Oguguo, B. C. E., Ene, C. U., Dave-Ugwu, P. O., Ezechukwu, R. I., Okebanama, C. I., Ukala, G., Agugoesi, O. J., & Ugwu, F. C. (2021). Academic nurse educators' perceptions of classroom assessment practices as means for quality nursing education. *International Journal of Psychosocial Rehabilitation*, 25(2).
- Olatunji, S. O., & Onofeghara, N. (2013). *Psychological testing: Foreign and indigenous psychological test*. Lagos.
- Osadebe, P. U. (2013). Validity and reliability of evaluation techniques for quality education. *DELSU Journal of Educational Research and Development*, 12(1), 56-63.
- Ovat, S., & Ofem, U. S. (2017). Teachers' variables and application of test blueprints in learners' assessment in secondary schools in Cross River State. *International Journal of Scientific Research in Education*, 10(1), 112-118.
- Quansah, F., & Amoako, I. (2018). Attitude of senior high school teachers towards test construction: Developing and validating a standardized instrument. *Research on Humanities and Social Sciences*, 8(1), 25-30.
- Sanni, S. O., & Ochepe, I. A. (2002). Effect of practical discussion outside the classroom on students' performance in mathematics. *Abacus: Journal of the Mathematical Association of Nigeria*, 27(1), 45-52.
- Sireci, S. G. (2013). Agreeing on validity arguments. *Journal of Educational Measurement*, 50, 99-104. <https://doi.org/10.1111/jedm.12005>
- Tshabalala, T. A. (2013). Causes of poor performance of ordinary level pupils in mathematics in rural secondary schools in Nkayi District.
- Ujah, E. U. (2010). Development and validation of an introductory technology achievement test [Unpublished M.Ed. thesis]. University of Nigeria.
- Uloko, E. S., & Imoko, B. I. (2007). Effect of ethnomathematics teaching approach and gender on students' achievement in locus. *Journal of the National Association of Science, Humanities and Education Research*, 5(1), 31-36.
- Van den Bergh, L., Ros, A., & Beiijaard, D. (2014). Improving teacher feedback during active learning: Effects of a professional development program: Effects of a professional development program. *American Education Research Journal*, 51(4), 722-809. <https://doi.org/10.3102/0002831214531322>
- Wiki, M. (2013). *Teaching qualification*. Wikipedia. http://en.wikipedia.org/wiki/teaching_qualification_on_4/4/13
- Yara, P. O., & Wanjohi, W. C. (2011). Performance determinants of KCSE in mathematics in secondary schools in Nyamira Division, Kenya. *Asian Social Science*, 7(20), 107-112. <https://doi.org/10.5539/ass.v7n2p107>