www.unibulletin.com
UNIVERSITEPARK BULTEN | BULLETIN
ISSN 2147-351X | e-ISSN 2564-8039

UNIVERSITEPARK bülten bulletin

RESEARCH ARTICLE

Secondary School Teachers' Knowledge on Procedures for Constructing Quality Classroom Tests in Tanzania

Jaquiline Amani · **Septimi Kitta** · **Orestes Silverius Kapinga** **Christina Mbilinyi**

Received *February 12, 2021*
Accepted *June 13, 2021*
Published Online *June 30, 2021*

CORRESPONDENCE
Jaquiline Amani
jacque.amani@gmail.com
Mkwawa University College of Education, PO BO 2513, Iringa-Tanzania.

AUTHOR DETAILS
Additional information about the authors is available at the end of the article.

To cite this article: *Amani, J., Kitta, S., Kapinga, O.S., & Mbilinyi, C. (2021). Secondary School Teachers' Knowledge on Procedures for Constructing Quality Classroom Tests in Tanzania. Üniversitepark Bülten, 10(1): 40-54.*

ABSTRACT

Background/purpose – Classroom assessment practices, being either formative or summative, form a fundamental part of the teaching and learning process. This paper presents findings on secondary school teachers' competences for constructing quality classroom tests. In particular, the study examined teachers' awareness of skills and procedures for constructing quality classroom tests, established the type of professional support teachers need for constructing quality tests based on the identified areas of deficiency, and determined the influence of experience in the teaching profession on teachers' competences for test construction.

Materials/methods – This study was conducted with a convenient sample of 246 secondary school teachers who were drawn from four regions in Tanzania, namely Lindi, Mtwara, Kilimanjaro, and Arusha. The study employed a quantitative research approach with the use of semi-structured questionnaires as the data collection tool. Data were analyzed using IBM's Statistical Package for Social Sciences Version 22.0 to compute frequencies, percentages, and other relevant statistical tests.

Results – The findings show that the majority of the participant teachers lacked competences for preparing quality classroom tests, particularly on the use of Table of Specification and test-item analysis. The results showed that more than 70% of the teachers had never received inservice training on the subjects of assessment and testing. It was further found that the teachers lacked professional support on how to prepare matching items, short answers, and multiple-choice test items.

Conclusion – Based on the findings, the authors recommend strengthening initial teacher education in view of competence-based assessment.

Keywords – Teachers' test construction skills, classroom-based tests, secondary schools, Tanzania.

To link to this article–https://dx.doi.org/10.22521/unibulletin.2021.101.3

1. INTRODUCTION

Over the past decade, the quality of education in Tanzania has been a major growing concern amongst researchers and the general public. Efforts to increase access to education at the national level have been evident over recent years. For example, net enrollment in Forms I to IV has increased by 5.5%, up from 29.1% in 2009 to 34.6% in 2018 (United Republic of Tanzania, Ministry of Education, Science and Technology, 2019; United Republic of Tanzania, Office of the President, 2017). This increase reflects the concerted efforts made by the Government of Tanzania to expand secondary education over the past 10 years by ensuring that all primary school leavers who pass the Primary School Leaving Examination (PSLE) go on to attend secondary education (United Republic of Tanzania, Ministry of Education, Science and Technology, 2017). However, the enrollment increase has not matched with the quality of learning outcomes. For example, the pass rates in the Certificate of Secondary Education Examination (CSEE) have fluctuated annually, with a significant downward trend from 91.5% in 2004 to 42.9% in 2013 (United Republic of Tanzania, Office of the President, 2017). Similarly, the pass rates show a rapid decline from 83.6% in 2008 down to 43.1% in 2012 (United Republic of Tanzania, Ministry of Education, Science and Technology, 2017), although this has recovered somewhat to 77.09% in 2017 (UnistoreTz, 2018) and 78.38% in 2018 (Mtema, 2019).

Several factors have been attributed to the poor quality of education in Tanzania. Mosha (2011) classified these factors into two main categories, namely: Contextual factors and input factors. “Contextual factors” include political, economic, demographic, cultural, and international conditions, while “input factors” include poor institutional leadership, poor funding, poor infrastructure, poor quality of teachers, and curriculum problems. Other researchers have linked the quality of education in Tanzania with teacher supply and competence, teachers’ commitment to teach, inadequate learning materials, and low levels of parental and community involvement in education (Irvine, 2010; Mkumbo, 2013; Mosha, 2011; Oduro et al., 2008). Most of the factors investigated in the literature have focused upon teaching methods, low parental and community engagement, poor quality of teaching and learning infrastructure, teachers’ qualifications, and teachers’ commitment to teaching. However, little has been published on how classroom assessment practices affect students’ learning, which eventually leads to poor learning outcomes.

Classroom assessment practices serve both formative and summative purposes depending on the intended learning goals. Formative assessment involves gathering and analyzing information about the students’ understanding of learning goals in order to provide instructional feedback for both teachers and students (Black & Wiliam, 2009; Perrigrino, 2014). It specifically intends to determine knowledge and skills to be taught and assessed (Ashfrod & De Stobbeleir, 2013), whilst making inferences and decisions based on students’ levels of mastery (Moon, 2005; Perrigrino, 2014). On the other hand, summative assessment intends to measure what students have learned at the end of a unit, a chapter, or a particular learning experience in order to make judgement about the quality of learning based on an established standard for various decision-making purposes (Earl, 2013; Isaacs et al., 2013); for example, selection of students for further education or employment, certification of school completion, or to enter certain occupations or professions (Ashfrod & De Stobbeleir, 2013; James & Lewis, 2012; Van den Berg et al., 2018). Reasonably, it makes sense to assume that both assessments *“for”* and *“of”* learning are critical in affecting students’ learning and how they perceive learning outcomes. As such, whatever form it may

take, since assessment has a role to play on students' learning and their future, then understanding assessment literacy among teachers is imperative.

Teachers at all levels of education construct and administer classroom tests and examinations. Research demonstrates that using appropriate classroom tests improves students' learning, motivation and achievement, as well as students' attitudes toward tests, reduces test anxiety, and improves teachers' instructional methods (James & Lewis, 2012; Jones, 2005; Rahman & Majumder, 2014; Umar, 2018; Vlachou, 2018). In order for a test to serve its purpose, it should possess, among others, two important qualities; reliability and validity. Reliability means that the test measures with consistency what it is supposed to measure, whereas validity is the extent to which a test serves its purpose with respect to the content and group for which it is intended (Airasian, 2004; Miller et al., 2009). The quality of tests given by teachers is closely linked with its ability to provide the kind of information needed regarding students' performance. A well-written test also allows teachers to accurately and consistently measure students' mastery of specific content taught in the classroom. On the contrary, poorly designed test items can lead to the inaccurate measurement of learning outcomes, and in presenting false information regarding students' performance as well as instructional effectiveness.

Previous studies on the subject of academic assessment have revealed several challenges with regards to school-based continuous assessment practices. For example, Byabato and Kisamo (2014) found that the implementation of school-based continuous assessment (CA) is not always properly applied due to a lack of integrity of some teachers, and a lack of uniformity in both the assessment tools employed and the procedures applied for both CA recording and reporting. This not only brings into question the content validity of CA testing, but also teachers' competences in their construction of quality classroom test items. The major reason put forward for this can be attributed to teachers' deficiencies in their basic professional skills, including the required skills and procedures used in the construction of quality classroom tests (Chonjo et al., 1996). Moreover, Hamafyelto et al. (2015) revealed a significant relationship between the competence of commerce teachers and content validity. In their study, it was shown that teachers concentrated on the lower levels of the cognitive domain. Deluca and Klinger (2010) reported further that low levels of teachers' assessment knowledge was more evident in newly qualified teachers and which lowers the quality of education (Popham, 2009). In another study, Ololube (2008) evaluated the competencies of professional and non-professional teachers in Nigeria and reported that professional teachers tend to construct effective evaluative instruments more than non-professional teachers.

To this end, based on the literature, it has become axiomatic that teachers' competence in classroom-based assessment is critical to the attainment of educational quality. Thus, examining assessment literacy amongst teachers for the purpose of improving students' learning outcomes is deemed worthwhile. More significantly, since learning and teaching are gradually becoming more competence-based, the call for assessment methods that adequately and effectively determine the required competences amongst learners has been equally emphasized in the literature. However, little has been done in Tanzania with regards to teachers' knowledge of test construction procedures, the status of professional support on test construction, and areas of deficiency in teachers constructing quality tests with a view to recommending potential counter strategies as far as a competence-based curriculum

is concerned. Based on this knowledge gap, the current study was designed to address three objectives.

- To determine teachers' awareness of existing procedures for constructing quality classroom tests.
- To establish the type of professional support needed to construct quality tests based on the identified areas of deficiencies in preparing examinations.
- To examine the influence of experience in the teaching profession on teachers' knowledge of test construction procedures.

Competencies for Test Construction: A Conceptual Framework

The changing theories and methods of educational assessment have been the focus of education systems in many countries throughout the world. Critics of contemporary education argue that no real change will take place in schools if traditional examinations remain unchanged, in that they exert a constraining influence on how teachers and students approach the curricula (Wolf, 1995). In 2005, the Government of Tanzania applied changes to the secondary education curriculum, which represented a paradigm shift from a content-based curriculum to one that is competence-based. These changes were guided by the Philosophy of Education for Self-reliance, Education, and Training Policy (United Republic of Tanzania, Ministry of Education and Culture, 1995), as well as other national programs such as the Growth and Poverty Reduction Strategy, the Tanzania Development Strategy of 2025, and the Education Sector Development Program (ESDP) (United Republic of Tanzania, Ministry of Education, Science and Technology, 2017). Curriculum improvement was one of the goals of the national program for the Secondary Education Development Plan (SEDP) and the Teacher Development Management Strategy (TDMS), of which some goals were aimed at improving the curriculum at the relevant level of education (Tanzania Institute of Education, 2013). As a result of the curriculum change, testing strategies are also expected to reflect the expected learning outcomes.

Test construction has long been a crucial element for measurement and evaluation in the educational setting, irrespective of whether its purpose is teaching or training. This means that both teachers and trainers are expected to have a certain level of mastery in the construction of reliable and valid tests in order to adequately assess learner performance. As teachers are primarily responsible for the assessment of students' learning, widespread concern exists about the quality of their assessment tools. The consensus has been that teachers use a variety of assessment tools, even though they may be inadequately trained in certain areas of educational assessment (Zhang & Burry Stock, 2003). Studies have also shown that teachers who only very minimally participate in professional development programs may not be adequately prepared to meet the demands and rigors of classroom assessment (Mertler, 2009; Schafer, 1991; Volante & Fazio, 2007).

This all implies that a key to quality classroom test construction is based on teachers' competences. Chidolue (1999, p. 36) highlighted that in order for a teacher to be able to construct high quality test items, they should possess a certain level of competency, including: determining the purpose of each testing exercise; stating specific and measurable educational objectives; establishing a good content outline; preparing a test plan which guides item construction; choosing an appropriate test-item format; constructing clear, precise, and unambiguous items; constructing items that focus on the attention of a group of students, often with widely varying background experiences, on a single idea; constructing

items with appropriate difficulty and discriminative indices; developing a marking guide specifically suited for the test; performing item analysis of test items; developing tests that are economical in both time and money; giving clear directions on how tests should be administered and applied; and, reviewing tests in order to correct any errors made during item construction. Moreover, Koksal (2004) added that for whatever purpose the test intends to achieve, the teacher needs to be able to provide answers to the following questions:

- Is the task perfectly clear?
- Is there more than one possible correct answer?
- Can test-takers arrive at the correct response without having the skill supposedly being tested?
- Do test-takers have adequate time available to perform the task(s)?

Based on Chidolue's (1999) classification, the researchers in the current study adapted and propose a conceptual framework which guides the undertaking of this study in Tanzania (see Figure 1).

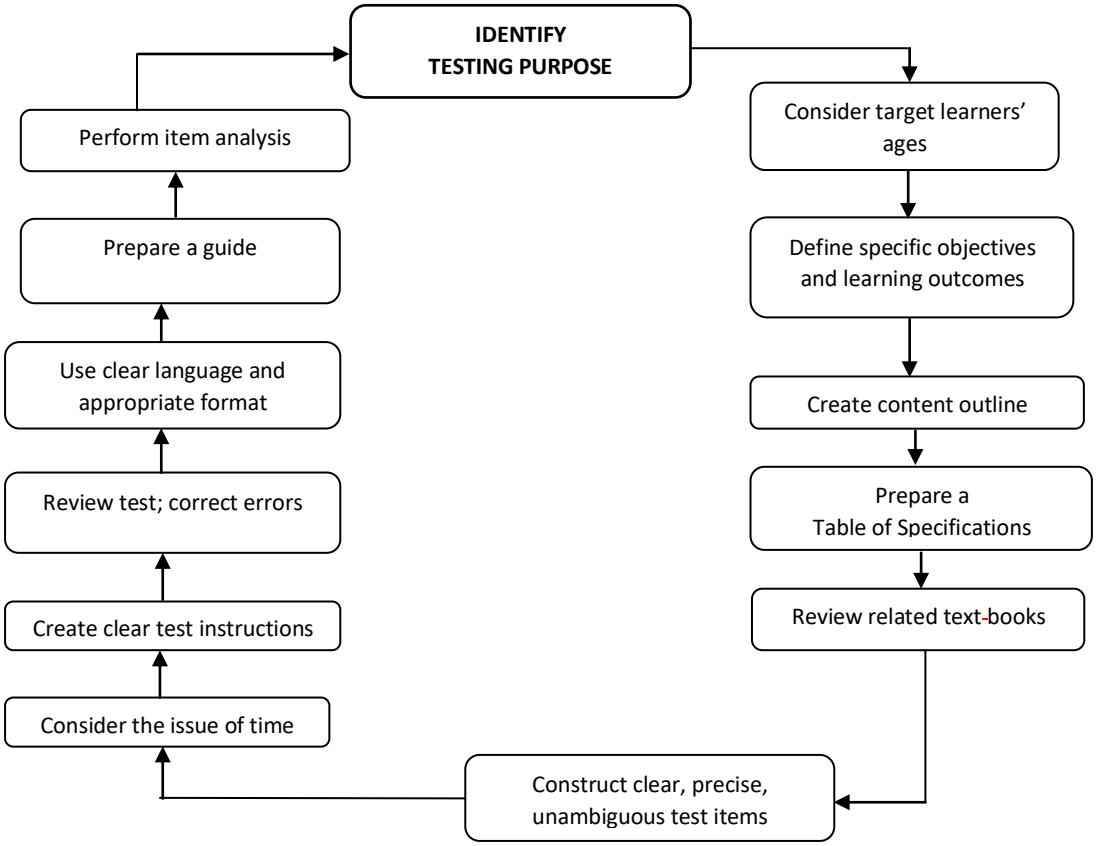

Figure 1. Teacher's competences in constructing quality test items (Chidolue, 1999)

2. METHODOLOGY

The current study was conducted in four different regions within mainland Tanzania, namely Lindi, Mtwara, Kilimanjaro, and Arusha. Form Four examination results from 2016 were used to rank the four regions in terms of both high and low levels of performance (Open Africa, n.d.). The population of the study consisted of ordinary level secondary school teachers working in the four selected regions. In each region, two low-performing and two

high-performing schools were selected based on the national examination results and relevant publically available statistics.

The participant teachers' qualifications and their working experience levels were assessed. The study included both teachers with a Diploma in Education and also those with a Degree qualification. Through semi-structured questionnaires, the teachers' skills related to constructing quality tests were examined. Furthermore, the participant teachers were asked for their opinion on those areas of the test construction process with which they experienced the most difficulty and in which they would value additional support. Descriptive statistics were used in calculating the mean scores and frequencies so as to determine the degree of emphasis placed on the problems. Also, an independent sample *t*-test was employed in order to determine differences in the participant teachers' knowledge on the procedures for constructing quality test items between the low-performing regions, whilst One-way Anova was applied in order to determine the effect of the participants' teaching experience on their knowledge of test construction procedures.

3. FINDINGS AND DISCUSSION

Participants' Background Information

In total, 246 completed questionnaires were analyzed following data cleansing. From the participant teachers purposively selected from secondary schools (I-IV) in the targeted study areas, of those who responded to the survey, 52.8% were female and 47.2% were male. As previously mentioned, the four regions of Tanzania were purposively selected based on the examination performance of their schools, as ranked by the National Examinations Council of Tanzania for 2016 (add citation/reference), and which included an equal mix of both low-performing and high-performing regions. The selection of the study participants was based on the following variables: gender, teachers' qualifications, and length of working experience. The study revealed that the majority of the participant teachers had teaching experience of 1-5 years. In terms of their qualifications, participants holding a Bachelor's degree formed the largest group (69.1%), followed by those with a Diploma qualification (26%), and those with a Master's degree (4.9%) (see Table 1).

Table 1. Participants' demographic information

Variable	Categories	Frequency	Percentage
Teacher numbers by region	Lindi	46	18.7
	Mtwara	57	23.2
	Kilimanjaro	63	25.6
	Arusha	80	32.5
	Total	246	100.0
Gender	Male	116	47.2
	Female	130	52.8
	Total	246	100.0
Teaching experience	1-5 years	109	44.3
	6-10 years	76	30.9
	11-15 years	31	12.6
	16 years+	30	12.2

Variable	Categories	Frequency	Percentage
	Total	246	100.0
Highest qualification	Diploma	64	26.0
	Bachelor's degree	170	69.1
	Master's degree	12	4.9
	Total	246	100.0

Status of Teachers' Professional Development in the Area of Test Construction

The study investigated the status of teachers' professional development in the area of test construction. This was considered important in order to relate the teachers' knowledge and skills on test construction, and whether or not they had the opportunity to attend training in this area. The results indicated that 72.8% of the participants had not attended any training on test construction (see Figure 2). Most of the participant teachers had received general professional development for teachers, but without any emphasis on the skills and procedures required for developing quality classroom tests. This implies that the majority of the teachers lacked continuous professional development on how best to assess their students based on the curriculum intentions.

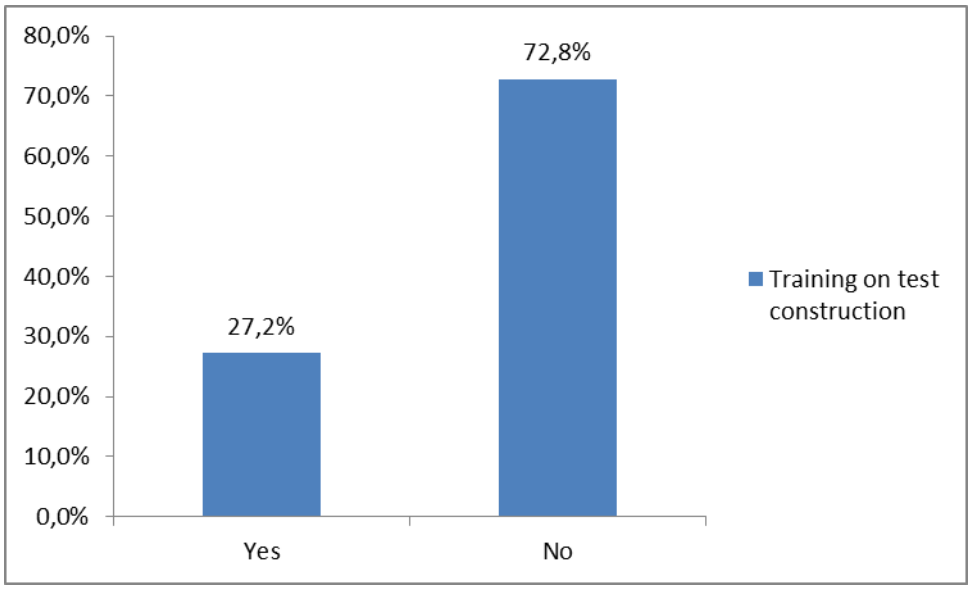

Figure 2. Status of teacher's professional development on test construction

Type of Test Items Preferred by Teachers

The study also found that when the teachers employed essay-type questions, 50.8% were said to have "very often" used true-false questions, whereas 18.3% said "rarely," and 15% said" or "never" in terms of assessing their students. The implications here may be that the teachers found it easier to prepare essay-type questions than matching items, short answers, or multiple-choice items. This was corroborated by the teachers' responses regarding the areas in which they most needed support so as to enable them to construct valid and reliable test items. The results illustrated in Figure 3 indicate that the teachers need support on how to construct multiple-choice questions (39.8%), essay-type questions (39.8%), matching items (39%), short answers (26.8%) and true-false items (23.2%).

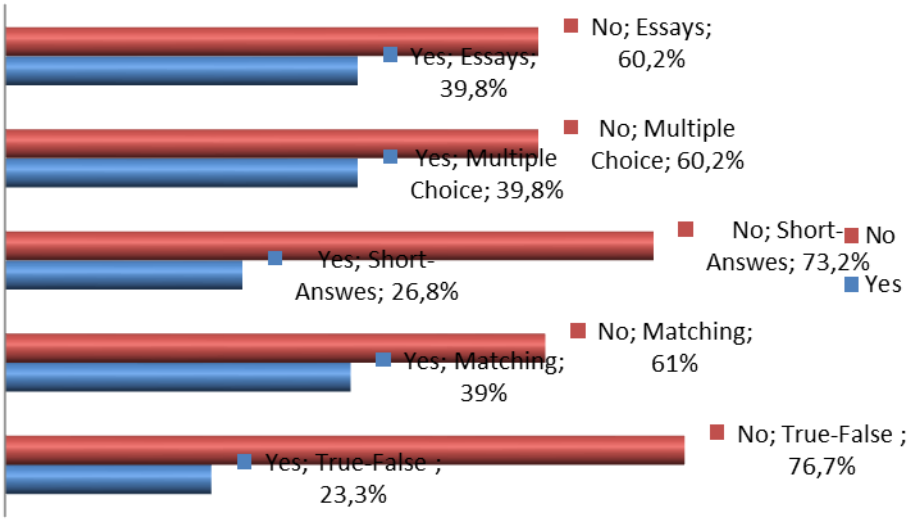

Figure 3. Professional support required by teachers on test-item construction

Teachers' Awareness Skills and Procedures for Constructing Quality Classroom Tests

This objective assessed knowledge of the participant secondary school teachers regarding the skills and procedures required for the construction of quality testing. In total, 20 items were developed based on the conceptual framework, and the relevant literature from the field of educational assessment and evaluation. The questions focused on both test preparation and test administration processes which are considered important for today's teachers. The results of the assessment are presented in Table 2.

Table 2. Teachers' awareness of test construction procedures

No.	Statements	Responses (%)			
		Lindi	Mtwara	Kilimanjaro	Arusha
		SA+A	SA+A	SA+A	SA+A
		SD+D	SD+D	SD+D	SD+D
1.	I outline the content covered before setting a test.	93.5 6.5	96.4 3.6	98.4 1.6	93.8 6.2
2.	I prepare a table of specifications (test blueprint) as a guide in test construction.	9.1 90.9	8.9 90.1	39.7 60.3	42.5 57.5
3.	I consult standard textbooks in the subject as a guide.	84.8 15.2	98.2 1.8	90.6 9.4	96.3 3.7
4.	I organize test items in a logical manner.	93.5 6.5	92.8 7.2	96.9 3.1	97.6 2.4
5.	I give clear instructions to guide test-takers.	86.9 13.1	98.3 1.7	96.9 3.1	98.7 1.3
6.	I construct tests so that both high and low achievers can understand them.	89.1 10.9	94.6 5.4	93.7 6.3	88.8 11.2
7.	I subject the test items to item analysis.	10.1 80.9	18.9 81.1	29.7 70.3	32.5 67.5
8.	I keep a resource bank of questions that can be referred to	97.8 2.2	98.2 1.8	92.2 7.8	98.7 1.3

	when setting a new test.				
9.	I set tests with due regard to the time available for testing.	75.7 24.3	76.4 23.6	66.9 33.1	71.2 28.8
10.	I add enough test items to cover all the requisite cognitive domain levels.	63.4 36.6	61.1 38.9	63.8 36.2	62.6 37.4
11.	I assign scores for each test item.	71.3 28.7	76.4 23.6	85.3 25.7	86.2 23.8
12.	I ensure that the items measure the determined objectives.	87.8 12.2	89.3 10.7	88.2 21.8	90 1 0
13.	I prepare a marking guide while constructing the test.	51.3 48.7	61.1 38.9	55.4 44.6	56.3 43.7
14.	I consider the age of learners during item writing.	78.3 21.7	77.8 22.2	70.9 29.1	78 22
15.	I avoid gender stereotypes in the test items.	78.3 21.7	83.9 16.1	84.3 15.7	81.2 18.8
16.	I submit items for vetting to the Head of Department or the Head of School.	78.3 21.7	75.9 24.1	76.4 25.6	79 21
17.	I avoid giving clues in multiple-choice questions.	82.6 17.4	81.1 18.9	81.2 18.8	88.3 11.7
18.	I use appropriate numbering and lettering formats in writing tests.	97.8 2.2	96.4 3.6	93.8 6.2	91.3 8.7
19.	I avoid questions or phrases that are too long when item writing.	86.9 13.1	94.6 5.4	78.1 21.9	82.5 17.5
20.	I set test items that elicit information on one thing at a time.	75.8 24.2	76.3 23.7	75 25	77 23

From Table 2, it can be seen that, with regards to outlining the content covered for the specific term before test construction, over 90% of the respondents from all four regions indicated their awareness of this practice. On the preparation of a table of specifications which is used as a guide for test construction, very few of the respondents from the Lindi and Mtwara (9.1% and 8.9%, respectively) regions were aware of this tool; whereas, that percentage was far higher in Kilimanjaro (39.7%) and Arusha (42.9%). In addition, more than 80% of the respondents reportedly consulted standard textbooks in their respective subject areas as a guide during test preparation. The findings also revealed that more than 90% of the respondents organized their test items in a logical manner during the test construction process.

Regarding the giving of clear instructions to their test subjects (students), more than 95% of the respondents from Mtwara, Kilimanjaro, and Arusha reported having given clear instructions to guide their students, whilst that figure was 85% for teachers from the Lindi region. The results showed that more than 85% of the respondents strongly agreed that they prepared tests with the intention that both low-achievers and high-achievers could adequately comprehend each question. The findings further indicated that very few of the respondents from Lindi (10.1%) and Mtwara (18.9%) subjected their proposed test items to

item analysis, and, whilst somewhat higher, less than one-third of the respondents from Kilimanjaro (29.7%) and Arusha (32.5%) reportedly applied the practice. Regarding keeping a resource bank of questions which could be referred to when constructing tests, the vast majority of the respondents (90%) indicated having maintained a resource bank of questions. In total, 75% of the respondents from the Lindi and Mtwara regions set tests that took into consideration the time availability for testing, whereas the figure was 66.9% from Kilimanjaro and 71.2% from Arusha regions.

On the topic of ensuring that test items appropriately measured the predetermined course objectives, a vast majority of the respondents, 90% from Arusha, 89.3% from Mtwara, 88.2% from Kilimanjaro, and 87.8% from Lindi, stated that they ensured that the test items they developed measured the predetermined course objectives. In addition, more than 50% of the respondents strongly agreed that they prepared marking guides whilst constructing tests. In terms of the respondents appropriately considering the age of learners during test-item preparation, the responses ranged from 70.6% in Kilimanjaro to 78.3% in Lindi. Also, the majority of respondents in both Lindi (78.3%) and also in Kilimanjaro (84.3%) strongly agreed that they avoided the issue of gender stereotyping when constructing test items.

When an independent sample *t*-test was performed for the purpose of determining if any statistical differences were found between the respondents knowledge of procedures for the construction of quality test items, no statistically significant differences were found between the low-performing regions ($M = 94.4$, $SD = 11.5$) and the high-performing regions ($M = 99.4$; $SD = 11.9$); $t\,(243) = 3.34$, $p = .01$ in terms of the teachers' awareness of test construction procedures. One-way Anova was also conducted so as to determine the effect of the respondents' teaching experience on their procedural knowledge of test construction. The results showed that teaching experience did not affect the teachers' awareness of appropriate test construction procedures ($f\,(3, 241) = 1.307$, $p = .27$).

Despite awareness of the various skills at play, it is evident that the majority of teachers in the current study had limited understanding of how to prepare and use a "Table of Specifications" (ToS), which is considered a necessary tool to ensure the construction of valid, quality classroom tests. The ToS, which is also referred to as a "test blueprint," helps teachers to align objectives, instruction, and assessment (Airasian, 2004; Miller et al., 2009; Notar et al., 2004). Therefore, teachers need to be aware of this tool because it can be used for a variety of assessment purposes, but is most commonly associated with test construction (Fives & DiDonato-Barnes, 2013). It is argued that when constructing a test, teachers need to be concerned with measuring an adequate sample of the class content according to the cognitive level that the course content/materials were taught (Fives & DiDonato-Barnes, 2013).

Another aspect that teachers showed little awareness of was item analysis, which is a process that examines student responses to individual test items and multiple-choice configurations in order to assess the quality of those items (Airasian, 2004; Miller et al., 2009), and also of the test as a whole. Item analysis is deemed important because the quality of a test depends upon each constituent item of a test (Sharma, 2000). Item analysis is especially valuable in improving items which are intended to be reapplied again in later tests, but may also be used to eliminate ambiguous or misleading items in a single test administration. In addition, item analysis is valuable for increasing instructors' skills in test construction, and for the identification of specific areas of course content which deserve

greater emphasis or where increased clarity is needed. Item analysis helps to identify items which are deemed likely "too difficult" or "too easy," for the intended target base, or which cannot differentiate between students who have adequately learned the content vs. those who have not, or questions that have distracters which are not plausible (Airasian, 2004; Miller et al., 2009). Therefore, once identified, teachers can remove them from the pool of items or change the items or modify the instruction accordingly.

Additionally, the preparation of a marking guide whilst constructing a test was not found to be highly regarded by the participant teachers in the current study, despite their numerous reported merits; with the reason given as because of a shortage of available time (Masayile et al., 2017). It may be argued further that a marking guide is an essential tool to help ensure teacher accountability for every mark, whilst enabling students to self-evaluate themselves while enhancing standardized grading (Masayile et al., 2017). Generally, Galle (2019) argues that the quality of tests given by a teacher, whether for formative or summative purposes, can be closely associated to its ability to provide the kind of information needed regarding students' learning outcomes. Thus, is it important for teachers to possess the relevant skills and knowledge for the preparation of quality tests and examinations, so as to ensure quality and reliable learning outcomes are met in full.

4. CONCLUSION AND RECOMMENDATIONS

Although the current study found that the majority of respondents seemed to be aware of a good number of procedures for constructing classroom tests, they did not, however, possess the required skills and competences for preparing a Table of Specifications or for conducting item-based analysis. These results could be partly attributed to the participant teachers having received inadequate training on the topic of assessment and evaluation in the course of executing their profession. However, it may be said that these findings have key implications for teacher education in current-day Tanzania. The results underline that classroom-based assessment may require additional focus during initial teacher-training programs, and specifically in the preparation of courses on assessment and evaluation.

It was also noted that the participant teachers preferred essay-type questions than matching items, short answers, or multiple-choice items in the tests they prepared for their students. This finding may imply a simple lack of knowledge and/or familiarity with how best to prepare test items. The current study concludes that valid classroom tests require proper planning, whilst adhering to relevant principles and procedures for the construction and administration of valid academic tests. Therefore, in order to ensure the delivery of quality education, teachers need to possess the relevant skills to be able to prepare and manage classroom-based tests, and thereby to provide effective and useful feedback on student learning. Based on these findings, the current study recommends strengthening both initial teacher education in view of the competence-based assessment in secondary schools, and also emphasizes regular inservice teacher professional development in the area of assessment and evaluation.

LIMITATIONS AND FUTURE RESEARCH

Despite the current study contributing to the existing knowledge on teachers' skills for quality test construction, and also the status of professional development amongst secondary school teachers in Tanzania, one major limitation of the study should be taken into consideration. The study adopted a survey design that employed the use of questionnaires to enable researchers to collect quantitative data on the phenomenon within

a relatively short period of time. The choice of such a research design, however, may be said to limit eliciting any deeper understanding of the matter under investigation. The use of mixed-methods research could have provided greater insight into the preference of one type of test item over another, and how the procedures for test construction are applied in the actual classroom setting. Therefore, it is recommended that a qualitative inquiry should be applied in order to measure the extent to which secondary teachers prepare and administer competence-based classroom testing, and how the feedback of such an assessment is communicated back to the students, to the school's management, and also the students' parents so as to improve the students' learning and the general quality of education.

DECLARATIONS

Author Contributions Jaquiline Amani developed the background and literature review of the study, as well as reviewed the research instruments, collected and analyzed data, and reported on the findings. Septimi Kitta developed the background and literature review, and also developed the research instruments, collected and entered data into SPSS, analyzed it, and then reported on the findings. He also reviewed the final version of the manuscript. Orestes Silverius Kapinga was responsible for developing the background, literature review, and methodology, and then for writing the final report and reviewing the finalized version.Christina Mbilinyi worked on the final version of the research proposal, collected and entered data into SPSS, and then analyzed it. She also reviewed the first draft of the manuscript. All of the authors have read and approved the published final version of the article.

Conflicts of Interest The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Ethical Approval Permission to collect data was granted by the University of Dar es Salaam.

Funding This research project was funded by Mkwawa University College of Education, a Constituent College of the University of Dar es Salaam.

Data Availability Statement Data for this study is available upon reasonable request.

Acknowledgments Data collection and the development of this manuscript were supported by research grants from Mkwawa University College of Education. The study sponsors had no role in the study design, data collection, or the analysis or interpretation of the collected data, the writing of the manuscript, or the decision to submit the article for publication. The content is solely the responsibility of the authors and does not necessarily represent the official views of the study sponsors. The authors would like to thank the management, staff, and students of the participant secondary schools in Lindi, Mtwara, Kilimanjaro, and Arusha where the data of the study was collected.

REFERENCES

Airasian, P. W. (2004). *Classroom assessment: Concepts and applications* (3rd ed.). McGraw Hill.

Ashfrod, S. J., & De Stobbeleir, K. E. M. (2013). Feedback, goal setting, and task performance revisited. In E. A. Locke & G. P. Latham (Eds.), *New developments in goal setting and task performance* (pp. 51-64). Routledge.

Black, P. J., & Wiliam, D. (2009). Developing a theory of formative assessment. *Educational Assessment Evaluation and Accountability, 21*(1), 5-31. https://doi.org/10.1007/s11092-008-9068-5

Byabato, S., & Kisamo, K. (2014). Implementation of School Based Continuous Assessment (CA) in Tanzania Ordinary Secondary Schools and its Implications on the Quality of Education. *Developing Country Studies*, *4*(6), 55-62. https://www.iiste.org/Journals/index.php/DCS/article/view/11781

Chidolue, M. (1999). *Introduction to tests and measurements* [Unpublished mimeograph]. Awka, Nnamdi Azikiwe University, Nigeria.

Chonjo, P. N., Osaki, K. M., Possi, M. A., & Mrutu, M. (1996). *Improving science education in secondary schools: A situational analysis of science teaching in selected secondary schools in Tanzania Mainland*. Tanzania Ministry of Education and Culture/GTZ.

Deluca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher Candidates' learning. *Assessment in Education Principles Policy and Practice Policy & Practice, 17*(4), 419-438. http://dx.doi.org/10.1080/0969594X.2010.516643

Fives, H., & DiDonato-Barnes, N. (2013). Classroom Test Construction: The Power of a Table of Specifications. *Practical Assessment, Research and Evaluation*, *18*, Article 3. https://doi.org/10.7275/cztt-7109

Hamafyelto, R. S., Hamman-Tukur, A. H., & 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. https://doi.org/10.5923/j.edu.20150505.01

Irvine, M. (2010). *The Quality of Teaching and Learning in Tanzanian Preschool and Pre-primary Classrooms (for Children Aged 3-5 and 5-6): A Rapid Appraisal Baseline Study*. Tanzania Ministry of Education and Vocational Training.

Isaacs, T., Zara, C., Herbert, G., Coombs, S. J., & Smith, C. (2013). *Key Concepts in Educational Assessment*. Sage.

James, M., & Lewis, J. (2012). Assessment in harmony with our understanding of learning: Problems and possibilities. In J. Gardner (Ed.), *Assessment and learning* (pp. 187-205). Sage. http://dx.doi.org/10.4135/9781446250808.n12

Jones, C. A. (2005). *Assessment for Learning.* Learning and Skills Development Agency.

Koksal, D. (2004). Assessing teachers' testing skills in ELT and enhancing their professional development through distance learning on the net. *Turkish Online Journal of Distance Education, 5*(1). https://dergipark.org.tr/tr/pub/tojde/issue/16931/176755

Masayile, Q., Nyoni, T., Nyoni, M., Bonga, W. G., & Musisinyani, B. J. (2017). An Evaluation of the use of Marking Schemes in the Teaching and Learning of Commerce at Ordinary level: A Case Study of Five (5) Secondary Schools in Chiefs Sayi and Jiri areas in Gokwe South District, Midlands, Zimbabwe. *Journal of Economics and Finance (DR-JJEF), 2*(6). https://storage.googleapis.com/wzukusers/user-28924863/documents/59550ebf8f131NAa6Zzw/DRJ-JEF20170602.pdf

Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools, 12,* 101-113. http://dx.doi.org/10.1177/1365480209105575.

Miller, M. D., Linn, R. L., & Gronlund, N. E. (2009). *Measurement and Assessment in Education*. Pearson.

Mkumbo, K. (2013). Factors associated with teachers' motivation and commitment to teach in Tanzania. *Journal of Educational Sciences and Psychology, LXV 3*(1), 58-71. http://jesp.upg-ploiesti.ro/index.php?option=com_phocadownload&view=file&id=386:factors-associated-with-teachers-motivation-and-commitment-to-teach-in-tanzania&Itemid=16

Moon, T. R. (2005). The role of assessment in differentiation. *Theory into Practice, 44*(3), 226-233. https://doi.org/10.1207/s15430421tip4403_7

Mosha, H. J. (2011, June). *The state of quality of education in Tanzania: A candid reflection*. [Conference presentation]. Symposium on Education, Dar es Salaam.

Mtema, N. (2019, January 25). Form 4 Exams Pass Rate Up. *Daily News*. https://www.dailynews.co.tz/news/2019-01-255c4aa8a83735a.aspx

Notar, C. E., Zuelke, D. C., Wilson, J. D., & Yunker, B. D. (2004). The table of specifications: Insuring accountability in teacher made tests. *Journal of Instructional Psychology, 31,* 115-129.

Oduro, G. K. T., Dachi, H., & Fertig, M. (2008, September, 8-12). *Educational leadership and quality education in disadvantaged communities in Ghana and Tanzania* [Conference presentation]. Commonwealth Council for Educational Administration and Management Conference, Durban, South Africa. https://www.edqual.org/publications/presentations/cceam08.pdf/at_download/file.pdf

Ololube, N. P. (2008). Evaluation competencies of professional and non professional teachers in Nigeria. *Studies in Educational Evaluation*, *34*(1), 44-51. https://doi.org/10.1016/j.stueduc.2008.01.004

Open Africa. (n.d.). *Ufaulu Kimkoa kwa Mitihani wa Kidato cha Nne 2012 [Regional Passing for 2012 Form Four Examinations].* [Dataset]. https://africaopendata.org/dataset/ufaulu-wa-mitihani-ya-kidato-cha-nne-kimkoa/resource/007dbf62-29a3-4589-9372-e8f38f14b484?inner_span=True

Perrigrino, J. W. (2014). Assessment as a positive influence of the 21st century teaching and learning: A system approach to progress. *Psicología Educativa, 20*(2), 65-77. https://doi.org/10.1016/j.pse.2014.11.002

Popham, W. J. (2009). Assessment Literacy for Teachers: Faddish or Fundamental? *Theory into Practice*, *48*(1), 4-11. https://doi.org/10.1080/00405840802577536

Rahman, S., & Majumder, A. A. (2014). Is it assessment of learning or assessment for learning? *South East Asia Journal of Public Health, 4*(1), 72-74. https://doi.org/10.3329/seajph.v4i1.21847

Schafer, W. D. (1991). Essential assessment skills in professional education of teachers. *Educational Measurement: Issues and Practice, 10*(1), 3-6. https://doi.org/10.1111/j.1745-3992.1991.tb00170.x

Sharma, S. R. (2000). *Modern teaching strategies.* Omsons.

Tanzania Institute of Education. (2013). *Maboresho na Mabadiliko ya Mitaala toka Mwaka 1961 hadi 2010* [Curriculum Improvements and Changes from 1961 to 2010].

Umar, A. M. (2018). The impact of assessment for learning on students' achievement in English for specific purposes: A case study of Pre-medical students at Khartoum University. *English Language Teaching, 11*(2), 15-25. https://doi.org/10.5539/elt.v11n2p15

UnistoreTz. (2018, January 30). *Necta Form Four Results Performance Report: Top 10 Students, Best Performed Schools Students and Poor Performed Schools*. http://www.unistoretz.com/magazine/necta-form-four-results-performance-report-top-10-studentsbest-performed-schools-students-and-poor-performed-schools/

United Republic of Tanzania, Ministry of Education and Culture. (1995). *Education and* training *policy*.

United Republic of Tanzania, Ministry of Education, Science and Technology. (2017). Basic Education Statistics in Tanzania.

United Republic of Tanzania, Ministry of Education, Science and Technology. (2019). Basic Education Statistics in Tanzania.

United Republic of Tanzania, Office of the President. (2017). *Pre-Primary, Primary, Secondary, Adult and Non-Formal Education Statistics Regional Data*. Regional Administration and Local Government.

Van den Berg, M., Bosker, R. J., & Suhre, C. J. M. (2018). Testing the effectiveness of classroom formative assessment in Dutch primary mathematic education. *School Effectiveness and School Improvement, 29*(3), 339-361. https://doi.org/10.1080/09243453.2017.1406376

Vlachou, M. A. (2018). Classroom assessment practices in middle school science lessons: A study among Greek science teachers. *Cogent Education*, *5*(1), Article 1455633. https://doi.org/10.1080/2331186X.2018.1455633

Volante, L., & Fazio, X. E. (2007). Exploring teachers' candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education, 30*(3), 749-770. https://journals.sfu.ca/cje/index.php/cje-rce/article/view/2973

Wolf, A. (1995). *Assessing Assessment: Competence based assessment*. Open University Press.

Zhang, Z., & Burry Stock, J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education, 16*(4), 323-342. https://doi.org/10.1207/S15324818AME1604_4

ABOUT THE CONTRIBUTORS

Jaquiline Amani Mkwawa University College of Education, Tanzania.
E-mail: jacque.amani@gmail.com
ORCID ID: https://orcid.org/0000-0001-7025-6771

Septimi Kitta School of Education, University of Dar es Salaam, Tanzania.
E-mail: sekitta@yahoo.com
ORCID ID: https://orcid.org/0000-0002-4320-819X

Orestes Silverius Kapinga Mkwawa University College of Education, Tanzania.
E-mail: orestes.katopu@gmail.com
ORCID ID: https://orcid.org/0000-0003-1698-5433

Christina Mbilinyi Mkwawa University College of Education, Tanzania (email:)
E-mail: mbilinyichristina@gmail.com
ORCID ID: https://orcid.org/0000-0002-2839-3403