

# Extensiveness of Tutors Adherence to the Objectives of Test Construction on Enhancing Quality Assessment Procedures in Teachers' Colleges in Kilimanjaro Region

Severa Tesha<sup>1</sup>; Fortunatus Mbua<sup>2</sup>; Catherine Muteti<sup>3</sup>

<sup>1</sup>Masters Student at Mwenge Catholic University, Tanzania

<sup>2</sup>Lecturer, Department of Education, Psychology and Curriculum Studies, Mwenge Catholic University, Tanzania

<sup>3</sup>Lecturer, Department of Education, Psychology and Curriculum Studies, Mwenge Catholic University, Tanzania

Publication Date: 2025/09/10

**Abstract:** This study investigated the extent to which tutors' adherence to the objectives of test construction enhances quality assessment procedures in teachers' colleges in the Kilimanjaro region. The study employed a convergent research design under a mixed methods approach, guided by the Value-Added Theory. The target population included seven (7) teachers' colleges in Kilimanjaro Region, comprising seven (7) college principals, seven (7) academic tutors, and 150 tutors. Through census sampling and total population sampling, the study selected a sample of six (6) teachers' colleges, six (6) college principals, six (6) academic tutors, and 120 tutors. Quantitative data were collected using questionnaires, while qualitative data were collected through interview guides and document analysis guide. The validity of data collection instruments was established by consult research experts at Mwenge Catholic University (MWECAU), and a pilot study was conducted in one Teachers' Colleges to test the tools. Reliability for Likert-type items from questionnaires was ensured through Cronbach's Alpha, which yielded a coefficient of 0.879 for tutors, indicating high reliability. The trustworthiness of qualitative data collection instruments was ensured through triangulation. Quantitative data were analyzed using descriptive statistics through SPSS version 27, while qualitative data were subjected to thematic analysis. Ethical considerations were strictly observed throughout the study; informed consent was obtained from all participants, and confidentiality and anonymity were maintained. Findings indicated that tutors in teachers' colleges across the Kilimanjaro region aligned test items with curriculum objectives, covered various learning domains, prepared valid and reliable assessments, and incorporated student feedback to enhance assessment quality to a high extent. The study concludes that tutors adhere to test construction objectives to a high extent, thereby promoting fairness, reliability, and competency-based assessment practices in teacher education. Summarize. The study recommended that tutors should be continuously provided with regular training on test construction, item writing, and assessment standards aligned with national guidelines to ensure consistent and effective assessment practices.

**Keywords:** *Objectives of Test Construction, Quality, Assessment Procedures, Teacher Training.*

**How to Cite:** Severa Tesha; Fortunatus Mbua; Catherine Muteti (2025) Extensiveness of Tutors Adherence to the Objectives of Test Construction on Enhancing Quality Assessment Procedures in Teachers' Colleges in Kilimanjaro Region. *International Journal of Innovative Science and Research Technology*, 10(8), 2835-2848. <https://doi.org/10.38124/ijisrt/25aug992>

## I. INTRODUCTION

The adherence to objectives during test construction is globally acknowledged as a crucial element in promoting quality assessment procedures within teacher education programs. In the current dynamic educational environment, ensuring that assessment practices are objective, valid, and reliable is fundamental to producing competent and professional teachers (Ketonen et al., 2023). Globally, there

is an increasing emphasis on enhancing assessment systems that are aligned with learning objectives, especially for student-teachers undergoing professional training (Hoppe et al., 2024). This aligns with international frameworks such as UNESCO's Agenda 2030, specifically Sustainable Development Goal (SDG) 4, which advocates for inclusive and equitable quality education. In this context, aligning test items with instructional objectives is seen as a strategic approach to ensure that assessments not only measure what is

intended but also contribute meaningfully to teaching and learning outcomes.

Across the world, education systems have intensified efforts to enhance quality of assessment procedures by emphasizing strict adherence to test construction objectives. Countries such as the United States have made formative assessments a central element of instructional planning, ensuring that assessments are aligned with learning objectives and offer continuous, objective-driven feedback to support differentiated teaching (Cañadas, 2021). In the United Kingdom and other European countries, assessment reforms have prioritized the balance between formative and summative approaches, with clear emphasis on structuring test items around defined instructional goals to ensure fairness and reliability (Włodarczyk et al., 2021; Nada & Legutko, 2022). Similarly, in Asia, countries like Singapore and South Korea have implemented rigorous assessment frameworks that emphasize cognitive and non-cognitive skill evaluation, underpinned by the alignment of assessment tools with curriculum objectives (Ferretti et al., 2021). These global trends underscore the critical role of educators, including tutors in teachers' colleges, in adhering extensively to test construction objectives to ensure that assessment procedures contribute meaningfully to the development of competent professionals.

In Sub-Saharan Africa, efforts to improve assessment practices increasingly focus on enhancing tutors' adherence to test construction objectives. Ghana and Nigeria have introduced national frameworks to guide assessment, yet studies reveal ongoing gaps in tutors' competencies, especially in aligning test items with learning objectives and conducting item analysis (Ankomah, 2020; Salihu, 2019). In East Africa, countries like Kenya and Uganda have integrated continuous assessment into their education systems, though questions remain about tutors' capacity to consistently apply structured test construction principles (Atuhurra & Kaffenberger, 2020). In Tanzania, NECTA's 2021 guidelines emphasize fairness, inclusivity, and reliability in assessments, with a strong focus on objective-driven test development. Despite these efforts, the extent to which tutors in teachers' colleges effectively adhere to these objectives in practice remains uncertain, raising concerns about the overall quality of assessment procedures and their role in developing competent teacher graduates (Hidayah & Syahrani, 2022).

Assessment in teachers' colleges takes two primary forms: summative and formative. Summative assessments are typically administered at the end of a term or program to evaluate the overall learning achievements of student-teachers and are often used for certification (Nutov et al., 2021; Kunuba, 2022). Formative assessments, on the other hand, provide continuous feedback aimed at improving instructional delivery and learner outcomes. When properly implemented, formative assessments can foster motivation and improve academic performance among student-teachers (Thomas & Brown, 2021). Both types of assessment must be grounded in clearly defined objectives to ensure they accurately reflect the intended learning outcomes.

In Tanzania, teachers' colleges are responsible for producing professional educators equipped with the competencies needed for classroom practice. These competencies are expected to be nurtured through systematic training and high-quality assessments administered by tutors (URT, 1999; Keller-Schneider et al., 2020). However, there are growing concerns about the effectiveness of current assessment practices in bridging the gap between theoretical knowledge and practical teaching skills. Graduates are reported to lack the professional competencies needed in real classroom settings (Yoong et al., 2018), raising questions about the credibility of internal assessment procedures used in teachers' colleges.

To address such concerns, NECTA's 2021 guidelines emphasize the need for adherence to key principles of test construction, including setting clear learning objectives. These practices are designed to ensure alignment between curriculum goals and assessment items, promote objectivity, and accommodate learners with special needs. Despite this guideline, empirical studies report that tutors struggle to apply these principles effectively due to limited training, time constraints, and lack of collaboration (Mkimbili & Kitta, 2021). Although some tutors understand the theoretical underpinnings of formative assessment, many lack the practical competencies to prepare a well-structured objective or engage in item moderation (Kitta et al., 2021).

A key aspect of enhancing quality assessment procedures in teachers' colleges is tutors' adherence to clearly defined test construction objectives. These objectives are essential for guiding the development of assessments that are valid, reliable, and aligned with intended learning outcomes (Brookhart, 2024). When tutors design test items based on these objectives, they help ensure that assessments accurately measure what students are expected to learn and can do (Brookhart, 2011). However, in the Tanzanian context, many tutors face challenges in applying these principles effectively due to limited training in assessment literacy and insufficient understanding of objective-based test design (Rubeba & Kitula, 2024; Rubeba & William, 2023). As a result, test items often fail to reflect curriculum demands, leading to inconsistencies in measuring student-teachers' competencies. These issues raise critical concerns about the extent to which tutors in teachers' colleges adhere to test construction objectives, which is fundamental to improving the quality of assessment procedures and the effectiveness of teacher preparation programs.

In Kilimanjaro region, despite the nationwide dissemination of the 2021 NECTA guidelines, concerns remain about the fidelity of their implementation in teachers' colleges. This is evidenced by notable discrepancies between internal assessment results and NECTA examination outcomes. As shown in Table 1, internal exams consistently record higher distinction and credit rates than NECTA results, raising questions about the validity of the internal assessments.

Table 1 Teachers College Examination Results in Kilimanjaro Region 2019-2023

YEAR	INTERNAL EXAMINATION					NECTA EXAMINATION				
	DISTITION	CREDITS	PASS	SUPP	FAILED	DISTITION	CREDITS	PASS	SUPP	FAILED
2019	7	97	17	0	0	0	62	59	0	0
2020	16	173	5	0	0	2	152	40	0	0
2021	12	101	8	0	0	0	62	59	0	0
2022	14	113	8	1	0	0	96	40	0	0
2023	17	112	2	1	0	2	109	21	0	1

Source: Principals Teachers College (2023)

Table 1 highlights a disparity between internal examination results and those administered by NECTA in teachers' colleges within the Kilimanjaro region. Internal assessments consistently report higher numbers of distinctions and credits; for example, in 2023, internal distinctions reached 17, while NECTA recorded only 2. This inconsistent gap raises important concerns regarding the validity and reliability of internal assessments and suggests possible weaknesses in tutors' adherence to NECTA's test construction objectives. Such discrepancies prompted the researcher to conduct this study which investigated the extent to which tutors adhere to the objectives of test construction on enhancing quality assessment procedures in teachers' colleges in the Kilimanjaro region.

## II. STATEMENT OF THE PROBLEM

The inconsistency in assessment procedures within teachers' colleges has been closely linked to tutors' inadequate adherence to the objectives of test construction, particularly in developing and implementing effective continuous assessments. Scholars such as Amani et al. (2021) and Mkimbili and Kitta (2020) have reported that tutors struggle with applying essential assessment practices due to limited skills and insufficient understanding of guidelines. This shortfall has resulted in poorly constructed internal assessments that do not reflect national standards. Notably, student-teachers often attain high scores in internal examinations, averaging a 93.86% pass rate, while their performance in NECTA examinations averages only 66.80% (Saimon & Mtenzi, 2022). Such discrepancies raise concerns over the validity and reliability of internal assessments and suggest that test construction does not follow structured procedures such as setting objectives (Frank & Dickson, 2021).

To address these challenges, NECTA introduced the 2021 assessment guidelines, aiming to standardize and improve the quality of continuous assessment in teachers' colleges. The guidelines emphasize systematic procedures, including the preparation of test items based on the objectives of test construction. However, studies reveal that tutors either lack the capacity or fail to implement these procedures effectively, due to inadequate training, limited oversight, and time constraints (William et al., 2020; Kitta et al., 2021). This situation has contributed to the production of graduates who

lack the competencies required for effective classroom practice (Nsubuga et al., 2020; Hey et al., 2020). While previous studies have examined general issues related to assessment in higher education (Baker, 2020; James, 2021), few have focused specifically on tutors' adherence to the full cycle of test construction as outlined in NECTA's 2021 guidelines. This study, therefore, investigated the extent to which tutors' adherence to the objectives of test construction enhances quality assessment procedures in teachers' colleges in the Kilimanjaro region.

## III. RESEARCH QUESTION

➤ *This Study is Guided by the Following Research Question;*

To what extent do tutors adhere to the objectives of test construction on enhancing quality assessment procedures in teachers' colleges in Kilimanjaro region?

## IV. SIGNIFICANCE OF THE STUDY

This study on the extent of tutors' adherence to the objectives of test construction in enhancing quality assessment procedures in teachers' colleges in Kilimanjaro Region will be significant to various stakeholders. Tutors will benefit by gaining practical insights into aligning assessments with learning objectives through the proper use of objectives, which will be essential for developing valid and reliable assessment tools. College principals will be informed on how effectively tutors implement the 2021 NECTA assessment guidelines, enabling them to identify areas needing support and improve assessment supervision and training. Student-teachers will gain from fairer and more objective assessments that accurately measure their competencies and better prepare them for professional practice. For policymakers, the findings will provide evidence on the effectiveness of test construction practices in teacher education, which will inform the ETP 2014 version 2023. Academically, the study will add to the existing literature and support the implementation of standards-based assessment practices in Tanzania's teacher education system.

## V. THEORETICAL FRAMEWORK

The Value-Added Theory (VAT) by Neil Smelser (1962) provides a relevant framework for understanding the implementation of the 2021 NECTA guidelines as a structured response to systemic weaknesses in test construction and assessment practices in teachers' colleges within the Kilimanjaro Region. According to VAT, social systems undergo strain when structural inconsistencies such as disjointed assessment procedures, lack of adherence to test construction objectives, and unreliable scoring create instability (Smelser, 1962). In such cases, formalized interventions emerge to restore equilibrium. The 2021 NECTA guidelines can be seen as a corrective mechanism designed to address these tensions by standardizing assessment processes across colleges, thereby reinforcing consistency, fairness, and alignment with national education standards (Mupupa, 2017; Saffer, 2018). VAT, therefore, explains the guidelines as a collective and institutional response aimed at improving the overall quality and reliability of assessments.

### ➤ *Strengths of Value-Added Theory*

VAT is particularly useful in educational research for explaining how formal reforms arise in response to institutional strain. In the context of this study, the theory justifies the 2021 NECTA guidelines as necessary interventions aimed at mitigating fragmentation and enhancing quality assurance in assessments (Saffer, 2018). It also offers a conceptual lens to monitor and evaluate whether such reforms are functioning effectively to reduce systemic weaknesses (Mupupa, 2017). By focusing on structured, goal-driven change, VAT provides a basis for understanding how organized responses can stabilize and improve test construction practices within teacher education systems.

### ➤ *Weaknesses of Value-Added Theory*

A key limitation of VAT is its assumption that all institutions respond uniformly to systemic strain, which is often not the case in practice. In Tanzania, teacher colleges differ in administrative support, resource availability, and tutor competence levels, factors that can result in varied responses to the same reform (Ngussa & Waiswa, 2021). VAT does not sufficiently account for these contextual variables, making it less effective in explaining why some tutors or institutions adhere to NECTA guidelines while others do not. This theory also tends to overlook micro-level resistance or informal practices that may undermine the intended outcomes of reforms (Saffer, 2018).

### ➤ *Application of Value-Added Theory*

In this study, VAT will guide the analysis of how the 2021 NECTA guidelines serve as a systemic response to persistent issues in test construction, such as poor alignment with learning objectives, lack of standardized test blueprints, and inconsistent scoring. The theory will help examine whether these guidelines have reduced systemic strain by enhancing tutors' adherence to clear test construction practices. Furthermore, VAT will assist in identifying areas where implementation is lacking and highlight the need for further interventions, including professional development,

monitoring systems, or localized strategies to ensure sustained improvement in assessment quality (Ngussa & Waiswa, 2021).

## VI. EMPIRICAL REVIEW

### ➤ *The Extent Tutors Adhere to the Objectives of Test Construction According to the Examination Guidelines*

The objectives of test construction focus on creating assessments that are valid, reliable, fair, and aligned with learning outcomes. Adherence to these objectives ensures that tests effectively measure student understanding and support educational goals. This study examines how well tutors align their test construction practices with these objectives.

Azarias et al. (2023) investigated the influence of a college-based assessment mode on the adherence to test construction objectives among teacher college students in Ghana. Using a quasi-experimental pre-test-post-test non-equivalent control group design, the study found that students exposed to the school-based assessment approach demonstrated greater alignment with test construction objectives, leading to improved academic performance compared to those subjected to the summative assessment strategy. The study recommended establishing assessment teams in teacher colleges, consisting of subject teachers, assessment experts, and school administrators, to ensure effective planning, resource allocation, and adherence to test construction objectives. The study by Azarias et al. was conducted in Ghana, where the educational policies, examination guidelines, and challenges faced by educators may differ from those in Tanzania. Tanzania's educational context, including unique examination guidelines, regional disparities, and resource constraints, necessitates a localized study to provide context-specific insights and recommendations.

Ajayi and Idusogie (2021) analyzed item writing flaws (IWFs) to evaluate whether examination questions adhered to the objectives of examination construction at the Federal College of Education, Asaba, Nigeria. Using a descriptive cross-sectional design, they examined the prevalence of IWFs in 57 objective question papers from centrally conducted end-of-semester examinations over three academic sessions (2016/2017, 2017/2018, and the first semester of 2018/2019). Their study revealed that, while lecturers utilized short-answer, multiple-choice, and alternative-response question formats, a significant proportion of objective examination questions failed to align with the intended objectives of assessment. The results highlighted frequent violations of standard item-writing rules, with the most common item-writing flaws (IWFs) involving irrelevant difficulty. The study by Ajayi and Idusogie (2021) analyzed item writing flaws (IWFs) in examination questions to evaluate adherence to test construction objectives at a Nigerian teachers' college. The study was limited to item analysis with little information about the broader practices and understanding of tutors in adhering to test construction objectives. This study focused on assessing tutors' adherence to test construction objectives to ensure high-quality assessments that meet examination standards.



In Kenya, Ochieng (2021) conducted a study on the relationship between college-based tests and the objectives of test construction in the Kenya Certificate of Teachers Education within Nairobi City County. The study employed a survey design, targeting student-teachers in teachers' colleges in Nairobi. Data were collected through structured questionnaires to capture perceptions regarding the alignment of tests with test construction objectives. For hypothesis testing, the study utilized the t-test to analyze the relationship between college-based tests and their adherence to test construction objectives. The findings revealed a negative association, indicating that the tests were not effectively aligned with the intended objectives, thus compromising their validity and reliability in assessing teacher competencies. The study by Ochieng (2021) examined the alignment of college-based tests with test construction objectives in Kenya. The study centered on the perceptions of student-teachers rather than directly assessing tutors' adherence to test construction objectives. Tutors play a crucial role in designing and implementing assessments, making their practices pivotal in ensuring adherence to examination guidelines. The current study focuses on tutors in Tanzania, investigating their knowledge, practices, and alignment with test construction objectives to address gaps in assessment quality at the source.

Amani et al. (2021) carried out a study about Teachers' Knowledge of Procedures for Constructing Quality Classroom Tests in Tanzania. This study involved a convenient sample of 246 tutors in the teachers' college from Lindi, Mtwara, Kilimanjaro, and Arusha regions in Tanzania. The research utilized a quantitative research approach, employing semi-structured questionnaires as the primary data collection tool. The findings revealed that a significant number of participating tutors lacked the necessary competencies for creating high-quality classroom tests, particularly in areas such as using a Table of Specification and conducting test-item analysis. The study also discovered that over 70% of the tutors had not received any in-service training on assessment and testing. Furthermore, the tutors reported a lack of professional support in preparing matching items, short answers, and multiple-choice test items. The study by Amani et al. (2021) explored tutors' knowledge of test construction in Tanzania, identifying gaps in areas like Tables of Specification and test-item analysis. The current study extends on the extent to which tutors in teacher training colleges align the objectives of test construction practices with national examination guidelines.

In Tanzania by Chacha and Onyango (2022) to investigate the adherence to test procedures in teachers' colleges in Mwanza region. The primary objective of this study was to explore the challenges faced by tutors in implementing the Competency-Based Curriculum (CBC), with a specific focus on the influence of teachers' competence. The study utilized a convergent parallel design within a mixed research approach. The findings revealed that many teachers' colleges lacked sufficient training on test construction and competent chemistry tutors, which significantly hindered the effective implementation of the competence-based curriculum in the subject. The study by Chacha and Onyango (2022) examined adherence to test

procedures in Tanzanian teacher colleges, focusing on challenges related to implementing the Competency-Based Curriculum (CBC) in Mwanza, particularly in chemistry. The study primarily explored the influence of teachers' competence on CBC implementation, with little information on addressing adherence to broader test construction guidelines. The current study extends their research by investigating how tutors across various disciplines in Tanzanian teacher training colleges adhere to national examination guidelines.

## VII. SUMMARY OF LITERATURE REVIEW AND RESEARCH GAPS

The reviewed studies from global, African, and Tanzanian contexts highlight various assessment practices but offer limited focus on tutors' adherence to test construction objectives. Methodological weaknesses, such as inappropriate designs and limited tools, raise concerns about the validity of their findings. Most studies emphasize student outcomes, item flaws, or subject-specific issues without thoroughly assessing how tutors follow key objectives like aligning tests with learning outcomes, setting clear goals, and ensuring fairness. In the Tanzanian context, few studies evaluate tutors' practices in relation to the 2021 NECTA Guidelines. Therefore, this study addresses both methodological and conceptual gaps by assessing the extent to which tutors in Kilimanjaro teachers' colleges adhere to test construction objectives, aiming to enhance assessment quality in line with national standards.

## VIII. DESIGN AND METHODOLOGY

This study employed a convergent research design under a mixed methods approach, allowing for the simultaneous collection and analysis of both quantitative and qualitative data sets to achieve a comprehensive understanding of the extent to which tutors adhere to test construction objectives. Quantitative and qualitative data sets were gathered concurrently, analyzed separately, and then compared to identify consistencies or contradictions (Creswell & Creswell, 2018). The target population included seven (7) teachers' colleges in Kilimanjaro Region, comprising seven (7) principals, seven (7) academic tutors, and 150 tutors. Through census sampling, total population sampling and the study selected a sample of six (6) teachers' colleges, six (6) principals, six (6) academic tutors, and 120 tutors (URT, 2023). Quantitative data were collected using structured questionnaires, while qualitative data were gathered through interview guides and document analysis guide. The validity of the instruments was ensured by research experts at Mwenye Catholic University (MWECAU), and a pilot study was conducted in one (1) teachers' college to test the tools. Reliability of the questionnaire was established using Cronbach's Alpha, which yielded a coefficient of 0.879 for tutors, indicating high reliability. The trustworthiness of qualitative data was ensured through triangulation. Quantitative data were analyzed using descriptive statistics through SPSS version 27, while qualitative data were subjected to thematic analysis involving familiarization with the data, coding, generating

themes, reviewing themes, defining and naming themes, and final report writing. Ethical considerations were strictly observed throughout the study; informed consent was obtained from all participants, confidentiality and anonymity were maintained, and the study received approval from relevant educational and research authorities to ensure compliance with institutional and professional ethical standards.

## IX. FINDINGS IN RESPECT TO THE THEMES FROM THE RESEARCH QUESTION

This study aimed to address the research question regarding the extent to which tutors adhere to test construction objectives in enhancing quality assessment practices in teachers' colleges within the Kilimanjaro region. Data were collected from tutors, academic officers, and college principals. The findings were organized and discussed based on the main theme derived from the research question.

### ➤ *The Extent Tutors Adhere to Test Construction Objectives in Enhancing Quality Assessment Practices in Teachers' Colleges within the Kilimanjaro Region*

This study specifically sought to examine the extent to which tutors in teachers' colleges across the Kilimanjaro region comply with test construction objectives to support quality assessment practices. A five-point Likert scale was used for analysis, where 1 = Very Low Extent (VLE), 2 = Low Extent (LE), 3 = Moderate (M), 4 = High Extent (HE), and 5 = Very High Extent (VHE). Frequency (F) and percentage (P) values were also employed in data presentation. As per the interpretation guidelines by Hashim et al. (2022), mean scores were categorized as follows: 1.00–1.89 signified a very low extent, 1.90–2.49 low extent, 2.50–3.49 moderate extent, 3.50–4.29 high extent, and 4.30–5.00 very high extent. Furthermore, percentage responses were interpreted following the scale by Taherdoost (2019): ≤20% denoted an extremely small minority; 21–49% a minority; 50–59% moderate; 60–69% majority; 70–89% very high majority; 90–99% extremely majority; and 100% represented an overwhelming majority. A summary of tutors' responses is presented in Table 1.

Table 1 Tutors' Responses on the Extent Tutors Adhere to Test Construction Objectives in Enhancing Quality Assessment Practices in Teachers' Colleges within the Kilimanjaro Region (n=115)

Statements	VLE		LE		ME		HE		VHE		MEAN
	f	%	f	%	f	%	f	%	f	%	
I ensure that test items align with curriculum objectives.	6	5.2	9	7.8	36	31.3	25	21.7	39	33.9	3.71
I construct test items that accurately measure students' learning outcomes.	3	2.6	9	7.8	26	22.6	42	36.5	35	30.4	3.84
I create assessments that cover various cognitive levels based on subject knowledge.	7	6.1	13	11.3	25	21.7	37	32.2	33	28.7	3.66
I create assessments that cover various aspects of affective domain based on subject knowledge.	2	1.7	17	14.8	22	19.1	41	35.7	33	28.7	3.74
I create assessments that cover various aspects of the psychomotor domain based on subject knowledge.	0	0.0	24	20.9	34	29.6	32	27.8	25	21.7	3.50
I develop valid and reliable test items following the examination guidelines.	8	7.0	3	2.6	29	25.2	36	31.3	39	33.9	3.82
I incorporate student feedback in improving test construction based on training programs	3	2.6	10	8.7	28	24.3	46	40.0	28	24.3	3.74
Test items prepared reflect the competences to be tested as stipulated in the syllabus	4	3.5	7	6.1	19	16.5	41	35.7	44	38.3	3.99
I prepare test items which cover the content taught	1	0.9	7	6.1	20	17.4	42	36.5	45	39.1	4.06
I avoid ambiguous statements in the test items which could lead to misinterpretation.	3	2.6	5	4.3	24	20.9	39	33.9	44	38.3	4.00
<b>GRAND MEAN</b>											<b>3.80</b>

Source: Field Data (2025) 1=Very Low Extent (VLE), 2= Low Extent (LE), 3=Moderate (M), 4=High Extent (HE), 5=Very High Extent (VHE), F=frequencies, P=Percentages

Data in Table 1 shows that a moderate (55.6%) of tutors rated a high and very high extent that they ensure test items align with curriculum objectives. In contrast, an extremely

minority (13.0%) of tutors rated this practice at a very low and low extent. Additionally, a minority proportion (31.3%) indicated a moderate extent. The mean score is 3.71,

indicating that majority of tutors recognize the importance of aligning test items with curriculum objectives to a high extent. This implies that most tutors demonstrate commendable commitment to curriculum alignment, a key element of quality assessment practices that contributes to the overall improvement of educational assessment. However, the existence of tutors who adhere only moderately or to a low extent signals the need for strengthened capacity-building efforts and regular follow-up to ensure more consistent and widespread application of this critical practice across teachers' colleges in the Kilimanjaro region. The researcher, during a face-to-face interview with the Academic Tutor 1 (AT<sub>1</sub>) had this to share:

Ensuring test items align with curriculum objectives is central to quality assessment. In my role, I often review drafts of test papers and notice some tutors include content that is either beyond or not clearly reflected in the syllabus. I always emphasize the need to begin with the curriculum and make sure every question links directly to what was planned and taught. It's not just about testing knowledge, it's about assessing the right knowledge. Some tutors follow this well, but others need regular reminders and support (AT<sub>1</sub> Personal communication, 15 April 2025).

- *Academic Tutor 2 (AT<sub>2</sub>) had this to Say:*

Alignment to curriculum objectives is non-negotiable, but from experience, not every tutor gives it the weight it deserves. I've observed cases where assessment items are set based on personal interest rather than curriculum demands. During moderation meetings, I always stress the importance of using curriculum objectives as a reference point. We also conduct internal workshops to support tutors in interpreting these objectives accurately. There's progress, but it's gradual. Some still struggle with bridging the gap between teaching content and assessment items (AT<sub>2</sub> Personal communication, 17 April 2025).

Information from both AT<sub>1</sub> and AT<sub>2</sub> illustrates that aligning test items with curriculum objectives is recognized as a core element of quality assessment. AT<sub>1</sub> highlights the importance of linking every test question directly to what is planned and taught in the syllabus, while noting varying levels of compliance among tutors. AT<sub>2</sub> reinforces this view, sharing experiences of misalignment and the efforts taken through internal support and training to ensure curriculum objectives are properly interpreted and applied. Together, their insights reflect a shared commitment to curriculum-based assessment, although consistent adherence remains an ongoing challenge requiring structured oversight and mentoring. The findings of the study are consistent with those by Azarias et al. (2023), who reported that aligning assessment items with curriculum objectives improves the effectiveness of learning outcomes and enhances academic performance. The study revealed that tutors ensured test items aligned with curriculum objectives to a high extent, indicating that curriculum-based assessment is central to quality assurance practices in teachers' colleges within the Kilimanjaro region.

Data in Table 1 shows that a majority (66.9%) of tutors rated to a high and very high extent that they construct test items that accurately measure students' learning outcomes. In contrast, an extremely minority (10.4%) of tutors rated this practice at a very low and low extent, while a minority (22.6%) indicated a moderate extent. The mean score of 3.84 suggests that the majority of tutors recognize the importance of constructing test items that accurately reflect students' learning outcomes to a high extent. This implies that tutors are strongly committed to developing valid assessment tools aligned with expected learning outcomes, which is a fundamental aspect of quality assessment practices. However, the presence of tutors who reported only moderate or low adherence indicates gaps in engagement or understanding regarding the development of outcome-based test items, underscoring the need for targeted support and professional development in this area. The researcher, during a face-to-face interview with Academic Tutor 4 (AT<sub>4</sub>) had this to share:

Measuring students' learning outcomes accurately is fundamental for fair assessment. As an academic tutor, I ensure that test items are designed to reflect the specific skills and knowledge that students are expected to demonstrate. However, sometimes tutors focus too much on recalling facts rather than applying or analyzing concepts, which doesn't fully capture student learning. We regularly conduct moderation meetings to adjust and improve the test items to better measure different levels of understanding (AT<sub>4</sub> Personal communication, 28 April 2025).

- *Academic Tutor 6 (AT<sub>6</sub>) had this to Say:*

Accurate measurement of learning outcomes depends largely on how well tutors understand the intended competencies. I've noticed that while some tutors are diligent in aligning questions with learning outcomes, others struggle with creating items that truly test comprehension and application. In my supervision role, I provide guidance and sample questions that reflect higher-order thinking skills. It's a process of continuous improvement, but I see positive changes as tutors gain more experience and training (AT<sub>6</sub> Personal communication, 12 May 2025).

Information from both AT<sub>4</sub> and AT<sub>6</sub> emphasizes the importance of constructing test items that accurately assess student learning outcomes. AT<sub>4</sub> points out the tendency of some tutors to focus on lower cognitive levels and the corrective role of moderation meetings. Meanwhile, AT<sub>6</sub> highlights variation in tutors' understanding of competencies and the supportive measures taken to improve question quality. Their reflections underscore a shared responsibility to enhance assessment accuracy through ongoing training and supervision within teachers' colleges in the Kilimanjaro region. The findings of the study are consistent with those by Ajayi and Idusogie (2021), who reported that poor adherence to item-writing principles resulted in assessment flaws such as irrelevant difficulty and reduced test validity. In this study, tutors demonstrated high commitment to constructing test items that accurately measure students' learning outcomes, showing that awareness and application of assessment objectives are increasingly being internalized in practice.

Moreover, the findings also support Smelser's Value-Added Theory (1962), which holds that perceived deficiencies in social or institutional systems generate a need for structured corrective responses. In this case, the construction of test items that accurately measure learning outcomes addresses the strain caused by traditional assessments that may overemphasize rote learning or neglect application and critical thinking. Tutors' deliberate focus on outcome-based assessment represents a mobilized, value-adding effort that not only ensures fairness and validity but also stabilizes and strengthens the assessment process in teachers' colleges.

Data in Table 1 shows that a majority (60.9%) of tutors rated to a high and very high extent that they create assessments that cover various cognitive levels based on subject knowledge. In contrast, an extremely minority (17.4%) of tutors rated this practice at a very low and low extent. Additionally, a minority proportion (21.7%) indicated a moderate extent. The mean score is 3.66, indicating that the majority of tutors recognize the importance of incorporating multiple cognitive levels, such as knowledge, comprehension, application, and analysis, when constructing assessments, to a high extent. This implies that tutors are making deliberate efforts to design assessments that go beyond rote recall and instead promote critical thinking and deeper understanding. However, the presence of tutors who apply this principle only moderately or to a low extent reflects limited awareness or inadequate training in cognitive-level differentiation. The researcher, during a face-to-face interview with Principal 3 (P<sub>3</sub>) had this to share:

As the principal, I emphasize the importance of assessments that target different cognitive levels. It's not enough for tutors to ask simple recall questions; they must challenge students to analyse, evaluate, and apply their knowledge. We encourage tutors to design tests that reflect Bloom's taxonomy, ensuring a range of cognitive demands. However, not all tutors are yet fully skilled in doing this, so we provide ongoing training and supervision to improve their assessment design capabilities (P<sub>3</sub> Personal communication, 24 April 2025).

- *Principal 6 (P<sub>6</sub>) had this to Speak:*

From my experience, assessments that cover various cognitive levels are crucial for developing critical thinking skills in students. I regularly review test papers and provide feedback to tutors on the cognitive demands of their questions. While many tutors have improved in this area, some still focus predominantly on lower-order thinking skills, which limits students' intellectual growth. Addressing this requires consistent capacity building and monitoring to ensure assessments are comprehensive and rigorous (P<sub>6</sub> Personal communication, 12 May 2025).

Information from both P<sub>3</sub> and P<sub>6</sub> highlights the degree to which tutors develop assessments addressing a range of cognitive levels. Principal 3 emphasizes the deliberate effort to encourage tutors to design questions that move beyond simple recall to higher-order thinking skills such as analysis, evaluation, and application. However, P<sub>3</sub> acknowledges that

some tutors still require support to fully implement this. Principal 6 points out that while progress has been made, several tutors continue to focus mainly on lower cognitive levels, limiting the breadth and depth of student learning assessments. Their reflections collectively indicate that creating assessments covering diverse cognitive levels remains a key focus area, with leadership actively promoting capacity building. The findings of the study are in agreement with Amani et al. (2021), who identified gaps in tutors' competencies regarding valid and reliable test construction. This study found that while a majority of tutors develop test items following examination guidelines, some still demonstrate only moderate or low adherence.

Data in Table 1 shows that a majority (64.4%) of tutors rated to a high and very high extent that they create assessments covering various aspects of the affective domain based on subject knowledge. In contrast, an extremely minority (16.5%) of tutors rated this practice at a very low and low extent. Additionally, an extremely minority proportion (19.1%) indicated a moderate extent. The mean score is 3.74, indicating that the majority of tutors adhere to this practice to a high extent. This suggests that tutors recognize the value of assessing attitudes, values, and emotional responses alongside cognitive skills, which is essential for nurturing holistic development in teacher trainees. However, the presence of tutors who reported only moderate or low adherence indicates limited understanding or confidence in designing assessments that effectively target affective learning outcomes. The researcher, during a face-to-face interview with Principal 1 (P<sub>1</sub>) had this to share:

In our college, we recognize that assessing the affective domain, such as attitudes, values, and feelings, is essential for holistic teacher education. I always encourage tutors to integrate these aspects into their assessments to ensure students develop not only knowledge but also the right dispositions. However, I've noticed that many tutors are more comfortable focusing on cognitive skills and tend to overlook affective outcomes. We are working on sensitizing staff about the importance of this domain and its role in producing well-rounded teachers (P<sub>1</sub> Personal communication, 15 April 2025).

- *Academic Tutor 3 (AT<sub>3</sub>) had this to Say:*

Covering the affective domain in assessments is often challenging because it's less tangible than cognitive skills. Tutors sometimes struggle to design questions or activities that accurately measure attitudes or values. In my supervisory role, I guide tutors on using reflective journals, peer assessments, and observational checklists as tools for assessing affective outcomes. While progress is slow, there is growing awareness of its significance, and more tutors are beginning to incorporate these methods into their assessments (AT<sub>3</sub> Personal communication, 22 April 2025).

Information from both P<sub>1</sub> and AT<sub>3</sub> highlights the growing attention given to incorporating the affective domain into assessment practices. Principal 1 points out the institutional recognition of its importance, but also the tendency for tutors to prioritize cognitive skills. Academic



Tutor 3 explains the practical challenges tutors face and the strategies used to support them in effectively assessing attitudes and values. Together, their insights reveal a shared commitment to promoting holistic assessment, though the integration of affective domain assessments remains a developing area within teachers' colleges in the Kilimanjaro region. The findings of the study are consistent with those by Chacha and Onyango (2022), who reported that many tutors lacked competence in constructing tests aligned with the affective domain, which reflects Competency-Based Curriculum (CBC). In this study, a large proportion of tutors reported that they ensure test items reflect the competencies stipulated in the syllabus, implying growing awareness of competency-based assessment despite some lingering inconsistencies.

Data in Table 1 shows that a minority (49.5%) of tutors rated to a high and very high extent that they create assessments covering various aspects of the psychomotor domain based on subject knowledge. In contrast, an extremely small minority (20.9%) rated this practice at a low extent, while a moderate proportion (29.6%) indicated a moderate extent. The mean score of 3.50, which falls within the high extent range, suggests that tutors generally recognize the importance of assessing practical and physical skills linked to learning. This indicates a commitment to incorporating psychomotor aspects into assessments, which is essential for evaluating skill-based competencies, particularly in subjects requiring demonstrations or hands-on activities. However, the lower percentage of tutors showing high adherence compared to other domains, along with a notable proportion reporting moderate or low engagement, points to possible gaps in awareness or confidence in effectively applying psychomotor assessment principles.

- *The Researcher, During a Face-to-Face Interview with Academic Tutor 5 (AT<sub>5</sub>), had this to Share:*

Assessing the psychomotor domain is crucial, especially in teacher training, where practical skills are just as important as theoretical knowledge. I've observed that some tutors integrate hands-on activities well, like microteaching demonstrations or instructional material production. However, others still rely heavily on written tests, even for subjects that demand practical demonstration. I always remind them that psychomotor assessment requires active student involvement and must be aligned with the competencies outlined in the syllabus (AT<sub>5</sub> Personal communication, 6 May 2025).

- *Academic Tutor 2 (AT<sub>2</sub>) had this to Speak:*

In my experience, the psychomotor domain is often the most neglected in assessment planning. Tutors sometimes lack confidence or creativity in designing tasks that evaluate physical or practical abilities. I support them by showing how lesson planning, classroom setup, and teaching demonstrations can be used as effective assessment tools. There's a noticeable improvement, but we are not yet where we should be. Continuous exposure to best practices is helping some tutors incorporate psychomotor tasks more meaningfully (AT<sub>2</sub> Personal communication, 17 April 2025).

Information from both AT<sub>5</sub> and AT<sub>2</sub> reflects the varying degrees of tutor engagement in assessing the psychomotor domain. AT<sub>5</sub> emphasizes the importance of hands-on performance and notes inconsistency in tutor practices, while AT<sub>2</sub> highlights the challenge tutors face in designing appropriate psychomotor tasks due to limited experience or confidence. Both responses underline that although progress is evident, the consistent integration of practical skill assessment remains an area requiring ongoing attention in teachers' colleges. The findings of the study are inconsistent with those of Ochieng (2021), who reported a disconnect between college-based assessments and test construction objectives in Kenyan colleges. In contrast, this study found that a majority of tutors prepare test items that directly reflect the content taught and avoid ambiguous statements, suggesting stronger efforts in upholding fairness, clarity, and content relevance in the Tanzanian context.

Furthermore, the findings also resonate with Smelser's Value-Added Theory (1962), which emphasizes that social systems adapt to stress through rationalized, value-adding responses. The inclusion of psychomotor skills in assessments addresses historical neglect of hands-on, skill-based learning outcomes and identified strain in teacher training. Tutors' efforts to incorporate performance-based tasks and practical demonstrations into test design represent an institutional adaptation aimed at producing more competent graduates. This adaptation supports the stabilization of teacher education by ensuring that future teachers are evaluated on practical as well as theoretical competencies, in line with VAT's principles of systemic correction and equilibrium.

Data in Table 1 shows that the majority (65.2%) of tutors rated to a high and very high extent that they develop valid and reliable test items in accordance with examination guidelines. In contrast, an extremely minority (9.6%) rated this practice at a very low and low extent, while a minority proportion (25.2%) indicated a moderate extent. The mean score is 3.82, which reflects that the majority of tutors implement this practice to a high extent. This implies that tutors demonstrate a solid commitment to ensuring that their test items are both valid, accurately measuring the intended learning outcomes and reliable, producing consistent results. Such practices are essential for upholding the credibility and effectiveness of assessments. Nevertheless, the presence of tutors who demonstrate only moderate or low adherence indicates variability in the application of valid and reliable test construction practices. This suggests that while most tutors uphold examination standards strongly, some lack full understanding or consistency in developing test items that meet these criteria. The researcher, during a face-to-face interview with Principal 2 (P<sub>2</sub>), had this to say:

In my capacity as principal, I constantly emphasize that test items must not only be aligned with the curriculum but also meet the standards of validity and reliability as outlined in national examination guidelines. While a good number of tutors adhere to these expectations, inconsistencies still arise. At times, we find test items that are either too general or fail to capture the intended learning outcomes. For this reason,

we've institutionalized a review process where assessments are checked before approval. This helps us maintain a high level of accuracy and fairness in student evaluation (P2 Personal communication, 17 April 2025).

- *Principal 4 (P4) Also had this to Contribute:*

What I've observed over the years is that tutors often understand what to teach, but translating that into reliable assessment tools is another challenge. Validity and reliability require thoughtful construction, not just picking any question from a textbook or past paper. Some tutors are very precise in this, but others need support in framing questions that actually measure what students have been taught and ensure consistent scoring. That's why we encourage collaborative item review sessions; it helps tutors sharpen their skills and standardize their approach to test development (P4 Personal communication, 30 April 2025).

Information from Principal 2 and Principal 4 indicates that tutors are generally aware of the need to develop test items that are valid and reliable, as required by examination guidelines. Principal 2 emphasizes structured review processes to ensure items meet expected standards, while Principal 4 notes that while some tutors construct appropriate items, others struggle with consistency and alignment. These responses suggest that the development of valid and reliable test items is practiced but not uniformly across all tutors. Moreover, the researcher during document analysis revealed that assessment tools such as internal examinations, portfolios, and tests used in teachers' colleges showed deliberate efforts to reflect examination guidelines regarding validity and reliability. Review of Teachers' College Assessment Reports confirmed that many institutions have put in place structured mechanisms for pre-assessment moderation and review to ensure the quality of test items. However, some test moderation minutes revealed inconsistencies in item construction, including vague instructions, lack of mark schemes, or test items not addressing higher-order thinking skills. The findings of the study are consistent with those by Winstone et al. (2023), who emphasized that active student involvement in feedback processes enhances assessment quality and fairness. Their research highlighted that when students engage with and make sense of feedback, educators are better able to refine test items, reducing ambiguity and improving alignment with learning targets.

Data in Table 1 shows that the majority (64.3%) of tutors rated to a high and very high extent that they incorporate student feedback to improve test construction based on training programs. In contrast, an extremely minority (11.3%) proportion of tutors rated this practice at a very low and low extent, while a minority proportion (24.3%) indicated a moderate extent. The mean score is 3.74, indicating that the majority of tutors apply this practice to a high extent. This suggests that most tutors value and utilize student feedback as a constructive tool to refine their assessment methods, contributing to more relevant, clear, and learner-centered test items. The presence of tutors who apply this practice only moderately or at low levels indicates variability in how student feedback is utilized in test

construction. This variation suggests that not all tutors consistently recognize or prioritize the role of student feedback in enhancing assessment quality, which affects the overall effectiveness of test improvement efforts in teachers' colleges within the Kilimanjaro region. The researcher, during a face-to-face interview with Academic Tutor 3 (AT3), had this to share:

Incorporating student feedback has become an important aspect of improving assessment practices. From what I've observed, some tutors use reflections from students, either through class discussions or course evaluations to adjust how they frame or structure test questions. It may not be formal in all cases, but the feedback helps identify areas where students face challenges, such as unclear wording or misalignment with taught content. I always encourage tutors to take this seriously, especially following training programs that emphasize learner-centered assessment (AT3 Personal communication, 22 April 2025).

- *Principal 2 (P2) had this to Speak:*

The idea of using student feedback to improve test construction is well recognized in our college. Tutors who have undergone assessment training are more likely to consider student input, especially in areas such as clarity of questions and relevance to what was taught. However, this practice is still emerging, and not every tutor applies it consistently. During staff meetings, we highlight the value of student perspectives not to undermine professional judgment, but to refine it. It's part of building a responsive and reflective assessment culture (P2 Personal communication, 17 April 2025).

The responses from AT3 and P2 reflect that the incorporation of student feedback into test construction is recognized but not yet consistently practiced. AT3 notes that tutors informally use student reflections to identify weaknesses in test clarity and alignment with content, especially after attending training programs. Similarly, P2 acknowledges that tutors who have undergone training are more likely to integrate student input, though this remains inconsistent across staff. These views suggest that while the practice of using student feedback exists, it is still developing and not uniformly embedded in assessment practices among tutors in teachers' colleges. The findings of the study align with Christopher (2024), who reviewed competency-based curricula in East Africa and noted that assessments must map directly to syllabus-stated competencies to maintain content validity. Ensuring that test items reflect prescribed competencies supports consistent measurement of essential skills.

Data in Table 1 show that a very high majority (74.0%) of tutors rated to a high and very high extent that the test items they prepared reflect the competences to be tested as stipulated in the syllabus. In contrast, an extremely minority (9.6%) rated this practice at a very low and low extent, while a minority proportion (16.5%) indicated a moderate extent. The mean score is 3.99, which falls within the high extent category. This indicates that the majority of tutors demonstrate strong adherence to aligning test items with the

expected competences outlined in the syllabus, thereby supporting the validity and relevance of assessments. Such alignment ensures that assessments accurately target the intended learning outcomes, which is central to quality assurance in education. The existence of a small group of tutors who apply this practice inconsistently indicates variations in how well tutors align test items with syllabus requirements. This inconsistency suggests that not all tutors fully adhere to syllabus guidelines, which affect the uniformity and quality of assessments across teachers' colleges in the Kilimanjaro region. The researcher, during a face-to-face interview with the Academic

- *Tutor 1 (AT1) had this to Share:*

Ensuring that test items reflect the required competencies is something I take seriously in my work. Every topic we teach comes with specific competencies outlined in the syllabus, and it is our duty to ensure that assessments target those. I always refer to the syllabus when preparing questions, and I encourage others to do the same. In some cases, tutors may unknowingly focus too much on content recall instead of competency-based outcomes, but regular internal moderation helps us catch that early (AT1 Personal communication, 15 April 2025).

- *Academic Tutor 5 (AT5) had this to Speak:*

What I've noticed is that most tutors are aware of the competencies they need to assess, but translating those into well-structured test items can be challenging. The syllabus provides clear guidance, but some tutors struggle to differentiate between testing content and testing applied competence. I usually advise them to use action verbs from the syllabus when formulating questions. That way, the items remain directly linked to the expected outcomes. Over time, I've seen improvement, especially after workshops on competency-based education (AT5 Personal communication, 6 May 2025).

The responses from AT1 and AT5 show that tutors recognize the importance of aligning test items with the competencies specified in the syllabus. AT1 emphasizes personal adherence to syllabus-based competencies and the use of internal moderation to ensure alignment. AT5 supports this view but also notes that while awareness exists, some tutors still face difficulty in formulating items that truly reflect applied competence rather than surface-level content. These responses suggest that although the practice is generally understood and improving, varying levels of implementation still exist among tutors in teachers' colleges. Moreover, the researcher during document analysis revealed that assessment tools such as internal examinations, portfolios, and test used in teachers' colleges showed deliberate efforts to reflect syllabus-based competences. The findings of the study resonate with Arefiev et al. (2020), who argued that classroom assessments should be directly tied to teach content to ensure test fairness and authenticity. Their work highlighted that alignment between instruction and assessment enhances validity and student confidence.

Data in Table 1 shows that an extremely high majority (75.6%) of tutors rated to a high and very high extent that they

prepare test items which cover the content taught. In contrast, an extremely small minority (7.0%) rated this practice at a very low and low extent, while a minority proportion (17.4%) indicated a moderate extent. The mean score is 4.06, which lies within the very high extent range and is the highest among all items assessed. This indicates that the overwhelming majority of tutors are strongly committed to ensuring that their test items are directly derived from the content delivered in the classroom. This practice is essential in promoting fairness and validity in assessment, as it helps ensure students are evaluated on what they have actually been taught. Nonetheless, the small proportion of tutors not consistently applying this principle indicates some gaps in maintaining content relevance and coverage in assessments. This inconsistency and variability of student evaluations across teachers' colleges in the Kilimanjaro region.

- *The Researcher, During a Face-to-Face Interview with Academic Tutor 2 (AT2) had this to Share:*

From my experience, most tutors are keen on ensuring that the test items they prepare are directly linked to the content they have covered during instruction. This is important because testing beyond the scope of what was taught can be discouraging and unfair to learners. In my department, we routinely review test items to confirm that they reflect classroom instruction. However, there are occasional gaps when tutors rush through content or don't document their lesson coverage clearly, which can lead to misalignment between the test and what was actually taught (AT2 Personal communication, 17 April 2025).

- *Principal 6 (P6) had this to Speak:*

I make it a point to emphasize to all academic staff that test items must strictly reflect the content delivered in class. It's an academic and ethical responsibility. Students should be assessed on what they were actually taught not what is assumed they should know. While most tutors adhere to this principle, there are a few instances where discrepancies arise, often due to poor record-keeping or lack of adequate planning. That's why we have pre-test audits and post-test reviews to ensure assessment integrity is maintained (P6 Personal communication, 12 May 2025).

The responses from AT2 and P6 reveal that the practice of preparing test items that cover the content taught is generally well observed among tutors. AT2 emphasizes the link between classroom instruction and assessment, noting that review processes help maintain alignment, though occasional lapses may occur. P6 reinforces this by highlighting institutional mechanisms such as test audits to uphold content relevance. These responses suggest a shared commitment to ensuring that assessments fairly reflect the taught curriculum, though minor inconsistencies still emerge due to planning or documentation challenges. The findings of the study mirror the consensus found in Haladyna and Rodriguez (2020), summarized in updated testing standards, which noted that ambiguity in test items leads to misinterpretation and undermines measurement precision. Clear, coherent wording is critical to maintaining reliable and valid assessments.



Data in Table 1 shows that a very high majority (72.2%) of tutors rated to a high and very high extent that they avoid ambiguous statements in test items that could lead to misinterpretation. In contrast, an extremely minority (6.9%) proportion of tutors rated this practice at a very low and low extent, while a minority proportion (20.9%) indicated a moderate extent. The mean score is 4.01, which lies within the very high extent range, indicating that the majority of tutors are highly committed to ensuring clarity and precision in the construction of test items. This practice is essential for reducing confusion, improving test reliability, and promoting accurate measurement of student understanding. However, a portion of tutors reported only moderate or low levels of avoiding ambiguous language in test items, indicating inconsistencies in how clearly assessment questions are constructed. This reflects differences in tutors' attention to clarity and precision, which influence the reliability and fairness of assessments administered across teachers' colleges in the Kilimanjaro region.

• *The Researcher, During a Face-to-Face Interview with Academic Tutor 3 (AT3) had this to Share:*

Clarity in test construction is non-negotiable. In my view, ambiguous wording confuses students and compromises the fairness of assessment. I often tell tutors that if a question can be interpreted in more than one way, then it's not assessing what it's meant to. We do peer reviews before tests are finalized, and one of the key things we check is the precision of language. Tutors are becoming more aware of this, especially after workshops on item writing and assessment standards (AT3 Personal communication, 22 April 2025).

• *Academic Tutor 4 (AT4) had this to Speak:*

I've encountered cases where test items included vague terms or double meanings, and it really affects how students perform, not because they don't understand the content, but because the question isn't clear. I emphasize to tutors that technical accuracy and language simplicity matter. Most of us use mock testing or item analysis to identify such issues before final exams. Though not perfect, we're improving, and tutors are more careful now than they were a few years back (AT4 Personal communication, 28 April 2025).

The responses from AT3 and AT4 indicate that efforts to avoid ambiguity in test items are actively practiced and increasingly emphasized among tutors. AT3 highlights the role of peer reviews and training in enhancing clarity, while AT4 points to the practical impact of ambiguous language on student performance. Both responses suggest that while challenges still occur, there is growing awareness and commitment among tutors to ensure test items are clearly stated and free from misinterpretation. The findings are consistent with Kibicho and Mwangi (2021), who affirmed that tutors who followed prescribed guidelines on validity and reliability produced higher quality assessments, supporting standardized test construction protocols.

The findings also align with Neil Smelser's Value-Added Theory (1962), which holds that systemic clarity and order are restored through collective, rationalized actions that

address dysfunction. The commitment by tutors to avoid ambiguous statements in test items directly tackles the confusion and misinterpretation that can undermine assessment validity. This practice reduces learner frustration and supports accurate measurement of learning, thereby correcting communication failures in the assessment process. Such efforts reflect VAT's stage of institutional adaptation to restore equilibrium, ensuring that assessment systems function with precision and transparency.

Generally, the study found that tutors in teachers' colleges in the Kilimanjaro region largely adhere to test construction objectives, which enhances the quality of assessments. High adherence was observed in aligning test items with curriculum objectives, covering cognitive, affective, and psychomotor domains, and ensuring content and competency alignment. A grand mean score of 3.80 indicated strong tutor commitment to quality assessment practices to a high extent. Interviews confirmed that practices like peer reviews, student feedback, and moderation meetings support valid and reliable test construction. However, challenges remain, especially in assessing the psychomotor domain and applying higher-order thinking skills.

## X. CONCLUSION OF THE STUDY

Based on the findings, the study made the following conclusion: Tutors in teachers' colleges across the Kilimanjaro region contribute to a high extent in enhancing quality assessment practices by adhering to the objectives of test construction. The study found that tutors consistently align test items with curriculum objectives, cover cognitive, affective, and psychomotor domains, and ensure assessments reflect taught content and intended competencies. These efforts, supported by structured practices such as peer reviews, feedback integration, and moderation meetings, have contributed to improved validity, reliability, and fairness in internal assessments. However, interview data revealed that some tutors still struggle with constructing higher-order thinking questions and assessing practical competencies, reflecting inconsistencies in practice. Despite these challenges, the study underscores the vital role of tutors in maintaining assessment standards and enhancing the credibility of teacher education through the effective implementation of test construction principles in alignment with national examination guidelines.

## RECOMMENDATIONS FOR ACTION

➤ *Based on the Conclusion of the Study, the Following Recommendation was Made:*

Tutors in teachers' colleges should be provided with regular and targeted professional development on test construction aligned with national examination guidelines.



## RECOMMENDATIONS FOR FURTHER STUDIES

Future studies should examine the impact of professional training programs on tutors' adherence to test construction objectives and their influence on student-teacher performance in national assessments.

## REFERENCES

- [1]. Ajayi, M., & Idusogie, E. (2021). An analysis of item writing flaws in examination questions. *Nigerian Journal of Educational Measurement*, 18(2), 33–47.
- [2]. Amani, J., Msuya, O., & Mkumbo, K. (2021). Teachers' knowledge on procedures for constructing quality classroom tests in Tanzania. *African Journal of Teacher Education*, 10(1), 22–36.
- [3]. Ankomah, Y. A. (2020). Assessment practices among tutors in Ghanaian teacher education institutions. *Journal of Educational Measurement and Evaluation*, 14(1), 12–28.
- [4]. Atuhurra, J., & Kaffenberger, M. (2020). Measuring and improving learning in East Africa: Evidence from Kenya and Uganda. *RISE Programme Working Paper*, 20(035), 1–29.
- [5]. Azarias, K., Mensah, F., & Boateng, J. (2023). Influence of school-based assessment on adherence to test construction objectives. *Journal of Educational Assessment in Africa*, 7(2), 88–101.
- [6]. Baker, L. (2020). Evaluating assessment systems in higher education: Issues and perspectives. *International Review of Education*, 66(4), 521–539.
- [7]. Brookhart, S. M. (2011). *Educational assessment: Principles and practices*. Pearson Education.
- [8]. Cañadas, M. (2021). Formative assessment in American schools: Trends and practices. *International Journal of Assessment Tools in Education*, 8(3), 56–70.
- [9]. Chacha, J., & Onyango, P. (2022). Challenges facing implementation of competency-based curriculum in teacher colleges: A case of Mwanza region. *Tanzania Journal of Education and Pedagogy*, 9(2), 66–77.
- [10]. Ferretti, G., Lee, C., & Kim, S. (2021). Policy shifts and assessment reforms in East Asia: A comparative review. *Asian Education Review*, 13(1), 13–27.
- [11]. Frank, T., & Dickson, A. (2021). Challenges in aligning college-based assessments with national standards. *Journal of Education Measurement and Practice*, 12(2), 67–75.
- [12]. Hidayah, N., & Syahrani, S. (2022). Educational assessment reforms and their effectiveness in Sub-Saharan Africa. *Journal of African Education Development*, 5(1), 44–60.
- [13]. Hoppe, T., Seidel, T., Renkl, A., & Riess, W. (2024). Advancing preservice science teachers' skills to assess student thinking on-the-fly through practice-based learning. *Journal of Teacher Education*, 76, 42–56. <https://doi.org/10.1177/00224871241237497>
- [14]. Nsubuga, M., Howard, S., & Mollel, (2020). The impact of ineffective assessment on teacher graduate readiness in Tanzania. *East African Journal of Educational Research*, 8(3), 105–118.
- [15]. James, R. (2021). Innovations in assessment mechanisms of college teachers. *Journal of Higher Education Innovation*, 14(3), 121–135.
- [16]. Keller-Schneider, M., Zellweger, K., & Albisser, S. (2020). Professional competencies and self-efficacy beliefs among teacher education graduates. *European Journal of Teacher Education*, 43(2), 214–230.
- [17]. Ketonen, L., Körkkö, M., & Pöysä, S. (2023). Authentic assessment as a support for student teachers' reflection. *European Journal of Teacher Education*. <https://doi.org/10.1080/02619768.2023.2229004>
- [18]. Kitta, S., Mwandanji, M., & Mkimbili, S. (2021). Challenges facing implementation of NECTA assessment guidelines in Tanzania. *Tanzania Journal of Education and Assessment*, 5(1), 77–92.
- [19]. Kunuba, J. (2022). The role of national examinations in certification and placement in Tanzanian teacher colleges. *Journal of Education and Practice*, 13(9), 44–58.
- [20]. Mkimbili, S., & Kitta, S. (2020). Teachers' capacity in constructing competency-based classroom assessments. *Journal of Contemporary Issues in Education*, 15(1), 33–49.
- [21]. Mupupa, S. (2017). Application of Value-Added Theory in education reforms. *African Journal of Education Management*, 6(2), 14–27.
- [22]. Nada, C., & Legutko, J. (2022). Fairness and reliability in European assessment practices. *European Educational Research Journal*, 21(1), 63–79.
- [23]. Nitko, A. J., & Brookhart, S. M. (2024). *Educational assessment of students* (7th ed.). Pearson.
- [24]. Ochieng, D. (2021). Relationship between college-based tests and the objectives of test construction in Kenya. *Kenya Journal of Education and Curriculum Studies*, 4(1), 20–38.
- [25]. Rubeba, A. D., & Kitula, N. G. (2024). Enhancing test construction practices among tutors in Tanzania's teacher education: The role of in-service training. *Tanzania Journal of Teacher Education*, 11(1), 43–58.
- [26]. Rubeba, A. D., & William, G. (2023). Assessment literacy among Tanzanian tutors: Challenges and prospects. *Journal of Education and Assessment Studies*, 9(1), 56–70.
- [27]. Saffer, J. (2018). The relevance of Value-Added Theory in contemporary educational change. *Global Education Review*, 5(3), 99–112.
- [28]. Saimon, M., & Mtenzi, F. (2022). Facilitating accommodation of learners with disabilities in summative assessment of literacies in Tanzania through mobile application. *International Journal of Language, Literacy and Translation*, 5(1), 42–63.
- [29]. Salihu, M. (2019). Teachers' adherence to assessment frameworks in Nigeria. *Journal of African Educational Studies*, 7(2), 23–38.
- [30]. Thomas, G., & Brown, M. (2021). The impact of formative assessment on student motivation. *International Journal of Teaching and Learning*, 14(1), 11–26.
- [31]. URT. (1999). *Education and Training Policy*. Dar es Salaam: Ministry of Education and Culture.

- [32]. William, P., Nombo, S., & Mwandanji, M. (2020). Standardizing continuous assessment practices in Tanzania's colleges. *African Journal of Educational Assessment*, 8(2), 33–49.
- [33]. Włodarczyk, D., Basarabová, S., & Petrova, G. (2021). Summative and formative assessment: Dilemmas in a European context. *Assessment & Evaluation in Higher Education*, 46(4), 498–512.
- [34]. Yoong, D., Mhando, R., & Masatu, S. (2018). Bridging theory and practice in teacher education in Tanzania. *Tanzania Education Review*, 6(2), 73–85.
- [35]. Mupupa, M. E. (2017). *Policy implementation and systemic change in education: A study of exam reforms in southern Africa*. *International Journal of Educational Development*, 55, 45–53.
- [36]. Ngussa, B. M., & Waiswa, D. (2021). Enhancing assessment practices in higher learning: A reflection on challenges and reforms. *Journal of Educational Assessment in Africa*, 6(2), 85–98.
- [37]. Saffer, B. (2018). *Reform dynamics in educational systems: A value-added approach to institutional change*. *Journal of Education Policy*, 33(4), 503–519. <https://doi.org/10.1080/02680939.2017.1363724>
- [38]. Smelser, N. J. (1962). *Theory of collective behavior*. Free Press.