THE UNITED REPUBLIC OF TANZANIA

MINISTRY OF EDUCATION, SCIENCE AND TECHNOLOGY

NATIONAL EXAMINATIONS COUNCIL OF TANZANIA

GUIDELINES ON ASSESSMENT PROCEDURES FOR SECONDARY SCHOOLS AND PROFESSIONAL LEVELS

Prepared by
The National Examinations Council of Tanzania
P.O. Box 2624
Dar es Salaam

November, 2021

THE UNITED REPUBLIC OF TANZANIA
MINISTRY OF EDUCATION, SCIENCE AND TECHNOLOGY
THE NATIONAL EXAMINATIONS COUNCIL OF TANZANIA

GUIDELINES ON ASSESSMENT PROCEDURES FOR SECONDARY SCHOOLS AND PROFESSIONAL LEVELS

Prepared by
The National Examinations Council of Tanzania
P.O. Box 2624
Dar es Salaam

November, 2021

Prepared by:
The National Examinations Council of Tanzania,
P.O. Box 2624,
Dar es Salaam, Tanzania.

TABLE OF CONTENTS

FOREWORD

The National Examination Council of Tanzania (NECTA) is a Government institution under the Ministry of Education, Science and Technology. It was established by Act No. 21 of 1973 that was then amended in 2019 by the National Examinations Council of Tanzania Act (CAP 107 R.E. 2019). It is responsible for the administration of examinations and assessment at national level. It is also responsible for the collection of school Continuous Assessment (CA) records which contribute 30 percent to Form Four, Form Six and Professional the final examination. In this regard, NECTA is pleased to issue Guidelines on Assessment Procedures at Secondary and Professional Levels to empower teachers and tutors to prepare valid, fair, adequate and reliable assessment tools and procedures to enhance learners learning experiences.

The guidelines provide essential techniques for preparing assessment tools and procedures that accommodate all types of learners, regardless of their ability and needs. They also provide ways of administering and recording test results for various purposes such as regulation of teaching and collection of CA records. Besides, they show the best ways to monitor and evaluate the quality of prepared assessment tools and procedures.

The National Examinations Council of Tanzania (NECTA) believes that these guidelines will help teachers and tutors to prepare valid and reliable assessment tools geared towards improving learning. The guidelines will also help School Quality Assurers in monitoring and evaluating teachers' competences in assessing learners' achievement.

Dr. Charles E. Msonde
EXECUTIVE SECRETARY

1.0 INTRODUCTION

The guidelines cover the meaning of, and types of assessment. They focus on the preparation of various assessment procedures such as test items, for normal learners and those with special needs. They also describe procedures for setting and moderating tests items. They further describe administering and marking tests as well as recording test results. The guidelines also inform about monitoring and evaluation of the assessment activities.

It is NECTA's expectation that the guidelines will help to enhance secondary school teachers' and college tutors' knowledge and skills on assessment by supplementing what they have acquired from universities or colleges.

The knowledge acquired from the guidelines will enable secondary school teachers and college tutors to prepare quality examinations, thereby generating quality Continuous Assessment (CA). CA includes tests, terminal and annual examinations. These contribute 30 percent of the final national examination results which NECTA compiles through the PReMS system.

Furthermore, using these guidelines, school quality assurers will have the capability to monitor and evaluate teaching and learning process by focusing on the kinds of assessment activities used by teachers.

2.0 ASSESSMENT

Assessment is a systematic, continuous process of monitoring various pieces of learning to evaluate learners' achievement and instructional effectiveness. It also refers to activities that are designed to measure learners' achievement from an instructional programme.

2.1 Purpose of Assessment

The teaching and learning process is not complete without assessment. Hence, assessment is carried out for various purposes, which include to:

(i) help teachers determine the effectiveness of their teaching techniques and learning materials;
(ii) determine the general trend in the development of the teaching and learning process;
(iii) identify problems that might hinder or prevent the achievement of set goals;
(iv) provide educational administrators with adequate information about teachers' effectiveness and school needs;
(v) identify learners' growth or lack of growth in acquiring desirable knowledge, skills, attitudes and societal values;
(vi) motivate learners to learn more as they discover their progress or challenges in a given task;
(vii) encourage learners to develop a sense of discipline and systematic study habits;
(viii) provide a base for determining the promotion of learners from one class to another; and
(ix) make objective and reliable decisions about educational planning.

2.2 Types of Assessment

2.2.1 Diagnostic Assessment

Diagnostic assessment is a form of pre-assessment that allows teachers/tutors to determine learners' individual strengths, weaknesses, knowledge and skills before instruction. It is primarily used to diagnose students' difficulties and to guide lesson and curriculum planning. Diagnostic assessment benefits both the instructor and the learners due to the following reasons:

(i) It allows the instructor to plan meaningful and efficient instruction. When the instructor knows exactly what learners know or do not know about a topic, they can easily focus lessons on the areas the learners still need to learn about rather than on what they already know.

(ii) It provides information to individualise instruction. It may show a teacher that a small group of learners needs additional instruction on a particular portion of the lesson. This makes it easy to provide remedial teaching to those learners so that they can fully engage with new content. Similarly, if the teacher discovers that a group of learners has already mastered a large portion of the unit of study, they can design activities that allow the group to go beyond the standard curriculum in that topic through independent or small group study.

2.2.2 Formative Assessment

Formative assessment refers to a range of formal assessments conducted during the learning process to modify teaching and learning activities for improving learners' future performance. The purpose is to find out whether the learners can do what they previously could not do after the learning experience. The formative assessment administered by teachers/tutors at secondary schools and college levels includes daily exercises, homework, weekly tests, monthly tests and terminal examinations. The formative assessment is used to:

(i) monitor learners' progress;
(ii) adjust instruction to maximise student achievement;
(iii) provide effective and timely feedback;
(iv) reveal learners who need remediation; and
(v) provide continuous assessment (CA) reports that contribute to the Form Four, Form Six and College students' final examination results.

2.2.3 Summative Assessment

Summative assessment refers to formal assessment conducted at the end of an educational activity to provide feedback that sums up the teaching and learning process at a particular level. It is judgemental. At secondary school and college levels, annual examinations serve as summative assessment. Summative assessment is used to:

(i) measure students' achievement at the end of a unit of study;
(ii) determine learning priorities for the learners;
(iii) evaluate group instruction or curriculum effectiveness; and
(iv) predict future success in other courses or standards.

2.3 Assessment Tools

The common assessment tools which will be reflected in these guidelines include quizzes, assignments, tests, laboratory activities and examinations. Throughout these guidelines, one type of assessment tools, namely tests, will be used to represent others.

2.4 Factors Affecting the Quality of the Test

2.4.1 Validity of the Test

Validity is the extent to which a test accurately measures what it is supposed to measure. In educational testing, "validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests." Generally, validity determines the meaningfulness and usefulness of assessment results drawn from accumulated evidence. The following are the main factors affecting the quality of a test:

(a) Factors originating from the test itself

Each test contains items and a close inquiry of the test items, which indicates whether the test measures the intended subject matter. Certain issues within the test item can limit the item from functioning as desired. These issues are related to the item's structure, clarity and comprehensiveness (i.e. completeness); proper use of the task words; inappropriate level of difficulty; arrangement and organization. These factors are summarized in the following subsections as follows:

(i) Length of the test

A test usually represents a sample of many questions. If the test is too short to become a representative one, validity is affected. Homogeneous lengthening of a test increases validity.

(ii) Unclear directions

If directions for responding to the test, whether it is permissible to guess and how to record the answers, are not clear to the learner, validity is reduced.

(iii) Too difficult reading vocabulary

Reading vocabularies in the test item are words or phrases that briefly tell an examinee what to do. Complicated vocabulary items in the test limit easy communication of the item task to the examinee.

(iv) Sentence structures

Too difficult sentences may fail to measure the aspects of learner performance, lowering the test validity.

(v) Inappropriate level of difficulty of test items

When test items have an inappropriate level of difficulty, they will affect the validity of the results. For example, in criterion-referenced tests, failure to match the difficulty specified by the learning outcome will lower the validity.

(vi) Poorly-constructed test items

Test items that provide unintentional clues to the answer will tend to measure learners' alertness in detecting the correct answer as well as the learners' performance, which ultimately affect its validity.

(vii) Ambiguity

Ambiguous statements in the test items lead to misinterpretation, different interpretation and confusion. Sometimes, they may confuse highly performing learners more than the low performing ones, resulting in the discrimination of items being in a negative direction. Consequently, the validity of the test is lowered.

(viii) Inappropriate test items for the outcomes being measured

Sometimes examiners are over ambitious. They wish to set items which measure complex levels of Higher Order Thinking Skills (HOS). However, the nature of some topics and the competences to be measured may not allow them to measure

HOS but Lower Order Thinking Skills (LOS). For example, if one wants to measure knowledge of procedures, terminologies and facts, they should set items from topics that will measure LOS. Hence, forcing the items to measure HOS may fail to meet the intended goals and lower its validity.

(ix) Improper arrangement of items

Items in the test are generally arranged in order of difficulty, with the easiest items first. If difficult items appear early in the test, they will make examinees spend too much time on these and fail to reach other items which they could answer easily. Such an improper arrangement may also influence validity by having a negative effect on learner motivation.

(x) Identifiable pattern of answers

When learners identify the systematic pattern of the correct answers (e.g.T,T,F,F or A,B,C,D), they can cleverly guess the answers to the rest of the items and this will affect the validity of the test.

(b) Functioning content and teaching procedures

The content words of the test and the teaching procedures can affect the validity of the test since the functioning content of the test items is not determined only by examining the form and content of the test. The teacher has to teach fully how to solve a particular problem before including it in the test.

A test on complex learning outcomes appears to be valid if the items function as intended. If the learners have previously experienced the solution of the problem included in the test, then

such a test is not a valid instrument for measuring complex mental processes. The items thus affect validity.

(c) Test administration and scoring

Test administration and scoring procedures may also affect the validity of the results. For instance, in teacher-made tests several factors might lower validity. These include insufficient time to complete the test, unfair help to individual learners, cheating during the examination and unreliable scoring of essay answers.
Similarly, in standardised tests, failure to follow standard directions and time limits, having irregularities for the learners and errors in scoring would lower validity. Whether it is a teacher made test or a standardised test, adverse physical and psychological conditions during testing may affect the validity.

(d) Learners' emotions and nature of responses

Certain personal factors influence the learner's response to the test situation and invalidate test interpretation. Emotionally disturbed, unmotivated, and nervous learners may not respond to the items as expected, and this may ultimately affect validity.

The response set also influences the test results. The test taking habit affects the learner's score. If the same test is administered several times, its validity will be reduced.

Therefore, test items should be set in such a way that they do not create anxiety among students in responding to the items.

(e) Nature of the group and the criterion

Validity is always specific to a particular group. For example, an arithmetic test based on word

problems may help in measuring reasoning ability in a group of slow learners. In contrast, a combination of a simple recall of information and computational skills is appropriate for measuring reasoning ability in a more advanced group.

Certain factors influence test measures. Examples are age, sex, ability level, educational backgrounds, interest in responding to the questions and cultural backgrounds. Therefore, the nature of the validation group should be considered in the test.

2.4.2 Reliability of the Test

Reliability refers to how dependably or consistently a test measures a characteristic. If a person takes the test again, will he or she get a similar test score, or a much different score? A test that yields similar scores for a person who repeats the test is said to measure a characteristic reliably. The factors that affect the reliability of the test items can be categorised into intrinsic and extrinsic ones. These are described in more detail below.

(a) Intrinsic factors

Intrinsic factors are those which lie within the test itself. The principal intrinsic factors that affect reliability are as follows:

(i) Length of the test

Reliability has a definite relation with the length of the test. The more the number of items the test contains, the greater will be its reliability, and vice-versa. Logically, the more sample of items one takes in a given area of knowledge, skill and the like, the more reliable the test will be.

However, it is difficult to ensure the maximum length of the test to achieve an appropriate value of reliability. The length of the test in such a case, should not lead the learner into fatigue. Nevertheless, examiners should prefer longer tests to shorter tests since the latter are less reliable.

(ii) Homogeneity of items

Homogeneity of items has two aspects: item reliability and the homogeneity of traits measured from one item to another. If the items measure different functions and the inter-correlations of items are ‘zero’ or close to it, then the reliability is ‘zero’ or very low. In contrast, if the items measure similar functions and the inter-correlations of items are far greater than zero, then the reliability is very high.

(iii) Difficulty level of items

A test item level of difficulty and clarity of expression also affect the reliability of test scores. If the test items are too easy or too difficult for group members, they will produce unreliable scores since both tests have a restricted spread of scores.

(iv) Discriminative value

When items can discriminate well between higher and lower achievers, the item total-correlation is high; the reliability is also likely to be high, and vice-versa.

(v) Test instructions

Clear and concise instructions increase reliability. Conversely, complicated and ambiguous instructions cause difficulties for

learners in understanding the questions and the nature of the response expected from the tested learners, ultimately leading to low reliability.

(b) Extrinsic factors

Extrinsic factors are those which originate outside the test itself. The following are extrinsic factors which influence reliability:

(i) Group variability

When the group of learners being tested is homogeneous in ability, the reliability of the test scores is likely to be lowered. In contrast, if the group of learners being tested is heterogeneous in ability, the reliability of the test scores is likely to be increased.

(ii) Guessing and chance errors

Guessing in the test increases error variance and, as such, reduces reliability. For example, in two-alternative response options, there is a 50 percent chance of answering the items correctly by guessing.

(iii) Environmental conditions

The testing environment should be uniform. Arrangement should be well planned such that light, sound, and other comforts should be equal to all tested learners; otherwise, it will affect the reliability of the test scores.

(iv) Momentary fluctuations

Momentary fluctuations may raise or lower the reliability of the test scores. Broken pencil, a pen that stops working, sudden unusual sound from outside the classroom, anxiety over uncompleted homework, mistake in giving the answer and not knowing how to change it are factors that which may affect the reliability of test score.

2.5 Competence-Based Assessment and the 21st Century Skills

2.5.1 Competence-based assessment

Competence-based assessment focuses on what an individual can do in his/her environment using the skills acquired in education. That is the application of knowledge and skills acquired in performing day-to-day activities and addressing environmental challenges. This type of assessment is conducted in contrast to content-based assessment, which focuses on what learners are expected to know.

Recognising the importance of the competence-based assessment, the National Examinations Council of Tanzania has shifted its emphasis from content-based assessment to competence--based assessment. The main purpose is to develop skilled personnel to apply their knowledge, skills and attitudes to different sectors, thereby fitting well in the new era of advanced technology. This change of direction was influenced by the 2014 Educational Policy which guides the country through achieving middle level economy and industrial development by 2025, hence the need for skilled workforce.

Competence-based assessment focuses on the understanding of concepts and the acquisition of skills and competences, as it emphasises on higher order thinking skills. Ultimately, competence-based assessment will produce people who can think critically, analyse and clarify issues, solve problems in their environment, and give opinions and suggestions based on evidence. Table 1 shows sample questions for competence-based assessment versus content-based assessment.

Table 1: Sample Competence-Based vs Content-Based Ordinary Level Assessment Questions

S/N	Competency-Based Assessment	Content-Based Assessment
1.	Between Dar es Salaam and Dodoma regions which one is the capital city of Tanzania? (Civics).	The capital city of Tanzania is ___.
2.	How does an animal cell differ from a plant cell? Give four key differences. (Biology).	State four key features of an animal cell.
3.	In five points, argue against the idea that "female genital mutilation is important in today's world." (Civics).	Explain the effects of female genital mutilation on today's world
4.	It has been reported that a lathe machine in a school workshop has a problem of producing tapered work. Assuming you are an expert in machine tool maintenance. What procedures would you use to identify and fix the problem? (Mechanical Engineering).	Describe the procedures for fixing a machine which produces tapered work.
5.	Chambua muundo, matumizi ya lugha na wahusika katika matini ifuatayo ya kifasihi/ubeti wa shairi lifuatalo:…… (Kiswahili).	Eleza muundo, matumizi ya lugha na wahusika katika kazi za fasihi.

S/N	Competency-Based Assessment	Content-Based Assessment
6.	Suppose you were a site technician and you are required to supervise the curing activity for concrete beams, slab and columns. Describe three methods that would help you to accomplish the work. (Civil Engineering).	Explain the curing process of concrete beams, slabs and columns.
7.	A milling machine that was bought two years ago has now started to make noise in the gearbox. Which maintenance process can be recommended for this machine? (Mechanical Engineering) A. Restoring maintenance B. Predictive maintenance C. Inspection maintenance D. Preventive maintenance E. Retaining maintenance	The type of maintenance which is done before a machine fails is …… A. restoring maintenance. B. predictive maintenance. C. inspection maintenance. D. preventive maintenance. E. retaining maintenance.
8.	Assume you have developed a friendship with a person of the opposite sex whom you think is the right candidate to get married to for about a year now. How can you describe such a relationship? (Civics) A. Premature marriage B. Wedding C. Honeymoon D. Courtship E. Marital relationship	Before marriage, men and women develop friendships. This period of friendship is called A. courtship. B. infidelity. C. early marriage. D. initiation period. E. kitchen party.
9.	The power consumption in a house whose circuit is connected to the 240 V main supply is as follows: Electric kettle 750 W, Bulb 50 W, Television 300 W and Heater 1500 W. Being a	Which fuse rate is suitable to use in a circuit connected to the 240 V main supplies with power consumption of 750 W, 50

S/N	Competency-Based Assessment	Content-Based Assessment
	technical advisor, which fuse rate will you suggest to use for proper current flow in the circuit? (Physics) A. 10 A B. 12 A C. 8 A D. 5 A E. 3 A	W, 300 W and 1500 W? A. 10 A B. 12 A C. 8 A D. 5 A E. 3 A
10.	Amani and Asha bought Coca-cola and Pepsi drinks for a farewell party. Amani spent 9,950 Tshs. to buy 12 bottles of Coca-cola and 5 bottles of Pepsi. Asha spent 8,150 Tshs. to buy 9 bottles of Coca-cola and 5 bottles of Pepsi drinks. Formulate a system of linear equations and apply the matrix method to find the price of one bottle of each type of the drinks. (Basic Mathematics).	Using the matrix method solve: $\begin{cases} 12x + 5y = 9950 \\ 9x + 5y = 8150 \end{cases}$
11.	Form Three students were asked by their Geography teacher to conduct a survey around their school. Explain eight pre-survey activities they needed to consider. (Geography).	Explain eight pre-survey activities.
12.	A team of experts from the district visited a recently constructed health facility, and they were less satisfied with the quality of the floor. Examine four factors to consider in effectively mending the floor finishing. (Building Construction).	Examine four factors to consider when selecting the type of floor finishing.

Table 2: Sample of Competency-Based VS Content-Based Assessment for Advanced Level Assessment Questions

S/N	Competence-Based Questions	Content-Based Questions
1.	Show how literacy techniques used by playwrights in two plays of your choice carry the intended messages. (English Language)	Mention the literacy techniques used by playwrights in two plays of your choice.
2.	A Form Five Biology student used a light microscope and observed a round large organelle located at the centre of a plant cell. (Biology) (i) Identify the organelle. (ii) Give five points to justify the role of the organelle in plants.	State five roles of a vacuole in plant.
3.	"The boiling point of sea water is slightly higher than that of fresh water". Briefly explain this fact and suggest appropriate term for this scenario. (Chemistry).	Give a reason responsible for boiling point elevation of a solution.
4.	Tathmini upungufu uliojitokeza wakati wa kusanifisha Lugha ya Kiswahili Tanzania kabla ya uhuru. Toa hoja sita. (Kiswahili).	Taja matatizo sita yaliyojitokeza wakati wa kusanifisha Lugha ya Kiswahili Tanzania kabla ya uhuru.
5.	Mwandishi wa fasihi ni chachu ya mabadiliko chanya katika jamii ili kuwa na maisha bora. Thibitisha kauli hii kwa kutumia riwaya mbili ulizosoma. Toa hoja nne kwa kila kitabu. (Kiswahili).	"Mwandishi wa fasihi ni muhimu katika jamii ili kuwa na maisha bora." Eleza kauli hiyo kwa kutumia riwaya mbili ulizosoma. Toa hoja nne kwa kila kitabu.

S/N	Competence-Based Questions	Content-Based Questions
6.	Asha experienced a change in temperature from hot to cold while travelling from Dar es Salaam to Mbeya. She did not understand the causes of the change. Analyze eight causes of the situation. (Geography).	Analyze eight factors which affect the temperature.
7.	Why the Kilimanjaro mountain does not decrease in height despite the continuous erosion taking place in the area? (Geography).	Describe the isostacy theory.

Table 3: Sample of Competency-Based VS Content-Based Teacher's Education Assessment Questions

S/N	Competency-Based Questions	Content-Based Questions
1.	In five points, justify the need for a teacher to prepare a scheme of work in the implementation of the subject's syllabus. (Curriculum and Teaching).	State the importance of the scheme of work.
2.	Mr Peter is a newly appointed headmaster of Mtakuja Primary school. He has discovered that the teachers are highly demotivated in performing their duties. How would you help him to address the situation? Give five points. (Psychology).	State five ways of motivating teachers.
3.	With examples, show how Benjamin Bloom's hierarchy of intellectual functions assists teachers in setting learning objectives. (Research Measurement and Evaluation).	What are the five functions of Benjamin Bloom's hierarchy of intellectual functions in setting learning objectives?

4.	Assess the usefulness of any five principles of learning according to Classical Conditioning Theory. (Psychology).	State five uses of the principles of learning according to Classical Conditioning Theory.
5.	Mr Ally is planning to design educational media and technology for teaching learners with intellectual impairment. Explain five factors that he should consider. (Educational Media and Psychology).	State five factors to consider in designing educational media and technology for teaching learners with intellectual impairment.
6.	Promoting cooperative learning is among the principles of teaching and learning Biology. How would you ensure that this principle is adhered? Give six points. (Biology Teaching Methods)	Explain six strategies that can be used to adhere to the principle of promoting cooperative learning among learners during teaching and learning Biology.
7.	How does the concept of philosophy differ from other pseudophilosophies? (Foundation of Education)	Define the following terms: (a) Philosophy (b) Pseudophilosophies

2.5.2 The 21st Century Skills

The 21st Century Skills are twelve (12) abilities that today's learners need to succeed in their careers during the information age. The twelve skills are critical thinking, creativity, collaboration, communication, information literacy, media literacy, technology literacy, flexibility, leadership, initiative, productivity and social skills.

These skills are intended to help learners keep up with the lightning-pace of today's modern markets. Each skill is unique in how it helps learners, but all the skills have one quality in common.

2.5.3 The link between the competence-based assessment and the 21st century skills

The education system is expected to equip learners with the 21st century skills which will contribute to their future careers. The National Examinations Council of Tanzania has been working hard to assess whether learners have acquired the 21st century skills. Though NECTA conducts only formative and summative assessments, it ensures that other types of assessments, such as diagnostic assessments, are well conducted to meet the requirements of competency-based assessment and the challenges of the 21st century.

2.6 The Taxonomy of Educational Objectives

The Taxonomy of Educational Objectives which is popularly known as Bloom's Taxonomy is a useful guide for developing a comprehensive list of instructional objectives. It attempts to identify and classify all possible educational learning outcomes.

2.6.1 The importance of Bloom's taxonomy in assessment

(a) It helps instructors to design performance tasks, construct questions by considering learners and provide feedback on learners' work.

(b) It helps to determine the level of learning outcome one intends to achieve.

(c) It assists curriculum developers and teachers in ensuring that all aspects of developing learners, especially at higher levels of thinking and operation, are included in the learning process.

(d) It provides an important framework for instructors to focus on higher order thinking.

2.6.2 Major Domains of the Taxonomy of Educational Objectives

The Taxonomy of Educational Objectives are classified into three major domains:

- Cognitive domain (Bloom et al., 1956)
- Affective domain (Krathwohl, 1964)
- Psychomotor domain (Simpson, 1972)

2.6.2.1 Cognitive domain

The cognitive domain involves knowledge and development of intellectual skills. There are six major levels of cognitive processes, starting from the simplest to the most complex (See Appendix 1 and 2). These are as follows:

- Level 1 – Remembering
- Level 2 – Understanding
- Level 3 – Applying
- Level 4 – Analysing
- Level 5 – Evaluating
- Level 6 - Creating

The key verbs and sample questions for each level of Bloom's taxonomy are shown in Table 4.

Table 4: Verbs and Sample Questions Used in Different Levels of Bloom's Taxonomy

S/N	Bloom's Level	The Behaviour Assessed	Key Verbs	Sample Questions
1.	**Level I: Remembering**	The learner's ability to recall or remember information	define, give, outline, describe, list, name, give, choose, identify, label, find, show, spell, who, what, when, where	(a) Define the term Chemistry. (b) What are the four reasons for the outbreak of cholera? (Biology) (c) Jairus complains of

S/N	Bloom's Level	The Behaviour Assessed	Key Verbs	Sample Questions
			and which	having burning sensation around the chest. (Biology) (i) What digestive disorder is he suffering from? (ii) Give five measures he should take to treat the disorder. (d) Name two breast milk substitutes. (Biology)
2.	**Level II: Understanding**	The learner's ability to organize, translate, interpret and state main ideas	explain, rephrase, classify, demonstrate, interpret, describe, illustrate, relate, translate, summarise, how and why	(a) Write a number that follows in the following sequence 1, 3, 5, 7, ___. (Basic Mathematics) (b) Summarise the following text in 20 words. (English Language). (c) You are invited by a Fema Club of a certain secondary school as a health practitioner. Explain how you will educate the members on four different ways of HIV/AIDS transmission. (Civics). (d) Classify human being to Class level. (Biology). (e) Why are chemical symbols useful in chemistry? Give three reasons. (Chemistry).

S/N	Bloom's Level	The Behaviour Assessed	Key Verbs	Sample Questions
3.	**Level III: Applying**	The learner's ability to use the information learned to solve problems in new situations	develop, apply, build, construct, make use and solve	(a) How would you use the principle of floatation to teach children how to swim? (Physics). (b) One angle of the right angled triangle is 40˚. Find the value of the second and the third angle. (Basic Mathematics). (c) Eleza kwa ufupi mazingira manne ambayo kielezi huweza kujipambanua. Kisha, tunga sentensi moja kwa kila aina ya mazingira. (Kiswahili). (d) Three airplanes arrived at Kilimanjaro International Airport (KIA) at the intervals of 30 minutes, 40 minutes and 55 minutes. If all three airplanes arrived at KIA at 2:00 p.m. on Saturday, when and at what time would they arrive together again? (Geography) (e) Solve the following equations using the given formulae. (Basic Mathematics).
4.	**Level IV: Analysing**	The learner's ability to differentiate various parts of the whole or to break	analyse, categorise, classify, compare, contrast/differentiate and dissect	(a) Chambua sentensi ifuatayo kwa kubainisha aina za maneno zilizotumika kuunda sentensi hii. (Kiswahili). (b) Classify rocks based on their mode of

S/N	Bloom's Level	The Behaviour Assessed	Key Verbs	Sample Questions
		information into parts		formation. (Geography) (c) Compare the climate of Kenya to that of Ghana. (Geography). (d) Analyse the important procedures for removing stains from clothes. (Textile and Dressmaking).
5.	**Level V: Evaluating**	The learner's ability to make judgment about information or validate ideas/ quality of work based on set criteria	defend, avaluate assess, judge, justify, give opinion, support and criticize	(a) With six reasons, defend the opinion that Mt Kilimanjaro is the National asset. (Geography). (b) Assess the importance of vegetation in the lives of wild animals. (Geography). (c) "Betrayal is one of the main sources of conflicts in many societies." Support this statement by referring to two novels that you have read. Provide three points from each novel. (English Language).
6.	**Level VI: Creating**	The learner's ability to create a new product, generate a new idea or create a different thought or process	compose, construct, create, design, develop, formulate, plan and propose	(a) Compose an essay on the effects of COVID 19. (The number of words should be between 200-300). (Civics) (b) Propose six ways that Tanzania can use to eradicate neo-colonialism. (History). (c) Rachel is three years older than her brother John. Three years to come, the product of their ages will be 130

S/N	Bloom's Level	The Behaviour Assessed	Key Verbs	Sample Questions
				years. (Basic Mathematics). (i) Formulate a quadratic equation representing this information. (ii) Use the formula to find the present ages of Rachel and John. Take *X* as the present age of Rachel.

2.6.2.2 Affective *domain*

The affective domain involves feelings, attitudes and emotions. It includes the ways in which people emotionally deal with external and internal phenomenon such as values, enthusiasms and motivations. This domain is categorized in five levels: receiving, responding, valuing, organizing and characterizing. All these levels are arranged from simple to complex, as shown in Table 5.

Table 5: Verbs and Sample Scenarios for Different Levels of Krathwohl's Taxonomy

S/N	Bloom's Level	The Behaviour Assessed	Key Verbs	Sample Scenario of occurrence
1.	Receiving	The learner's ability to attend to stimuli (ie. Ability to pay attention and being aware of the existence of certain ideas, materials or phenomena).	Ask, acknowledge, attend (to), follow, listen, meet, observe, receive	It takes place through listening attentively to someone, watching a movie and listening to a lecture

2.	Responding	The learner's ability to actively participate in the learning process by not only being aware of the stimulus but also reacting or responding to it in some way	agree, allow answer, ask assist, attempt, choose communicate comply, conform cooperate demonstrate describe, discuss display, exhibit follow, give help, identify	It takes place when having a conversation, participating in a group discussion, giving a presentation, complying with procedures or following directions.
3.	Valuing	The learner's the ability to see the worth of something, attach significance and express it	adopt, aid, care (for), complete, complement, contribute, delay, encourage, endorse, enforce, evaluate, expedite, foster, guide, initiate, interact, join, justify, maintain, monitor, praise, preserve, propose, query, react, respect, seek, share, study, subscribe, suggest, support, thank, uphold.	It takes place from a point a person's desires for a team to improve its skills to the point of taking responsibility for the overall improvement of the team.
4.	Organization	The learner's ability to put together different values, information and ideas then relating them to already held beliefs to bring it into an internally consistent philosophy. Or the ability to prioritise one value over another and create a	anticipate, collaborate, confer, consider, consult, coordinate, design, direct, establish, facilitate, follow through, investigate, judge, lead, manage, modify, organize, oversee, plan, qualify, recommend, revise, simplify, specify, submit, synthesize, test, vary, weigh,	It takes place when comparing, relating and assessing values to create a unique value system. For example, it can be done through spending more time studying then playing sports, recognising the need for balance between

		unique value system.		studying and outdoor activities or prioritizing time effectively to meet goals.
5.	Characterization	The learner's ability to act consistently in accordance with the set of values or philosophy about life. It involves internalizing values that guide behavior.	act, administer advance advocate, aid challenge, change commit (to) counsel, criticize debate, defend disagree, dispute empathize, enhance excuse, forgive influence motivate negotiate, object, persevere, persist, praise profess, promote, promulgate question, reject, resolve seek, serve strive, solve tolerate volunteer (for).	It takes place when a person internalises values and lets them control or guide his/ her behaviour. This can be realized through spending time with friends or fellow learners.

2.6.2.3 Psychomotor domain

The psychomotor domain is demonstrated by physical skills such as movements, coordination and manipulation. It includes behaviour demonstrated by all coordination activities, which involve arms, hands, fingers and feet.

The domain includes utilising motor skills and coordinating them. Its sub domains include perception, set, guided response, mechanism, complex overt response, adaptation, and origination.

Table 6: Verbs and Sample Scenarios for Different Levels of Simpson's Taxonomy

S/N	Bloom's Level	The Behaviour Assessed	Key verbs	Sample Task or Scenario of occurrence
1.	Perception	The learner's ability to use various senses to obtain clues	choose, describe, detect, differentiate, distinguish, identify, isolate, relate, select	Observe the correct technique for preparing oxygen gas in the laboratory.
2.	Set	The learner's readiness to take action	begin, display, explain, move, proceed, react, show, state, volunteer	Describe the steps involved in testing for simple sugars in the laboratory.
3.	Guided responses	The learner's knowledge of the steps required to perform a task	copy, trace, follow, react, reproduce, respond.	Demonstrate safely the procedures for making an alloy material and correctly on multiple occasions.
4.	Mechanism	The learner's ability to perform tasks habitually	assemble, calibrate, construct, dismantle, display, fasten, fix, grind, heat, manipulate, measure, mend, mix, organize, sketch.	It can be vivid to the learner when showing that they can solve examination questions after they have confidently answered some past questions previously. i.e. Asking a skilled shot putter to perform a thorough shot put
5.	Complex overt responses	The learner's skillful performance of motor acts	assemble, build, calibrate, construct, dismantle, display, fasten, fix, grind, heat, manipulate,	It can be used to assess the student's typing speed when using a computer. Sample question:

S/N	Bloom's Level	The Behaviour Assessed	Key verbs	Sample Task or Scenario of occurrence
			measure, mends, mix, organize, sketch.	How quick and correctly can you perform a thorough shot put?
6.	Adaptability	The learner's skillful performance of motor acts and modification of movement in problematic or new situations	adapt, alter, change, rearrange, reorganise, revise, vary	It takes place when a student who has learnt various underlying theories can develop a working model using everyday materials. Sample question: How best can the procedures for producing electricity from fossil fuels be altered?
7.	Origination	The learner's ability to create new movement patterns for problematic or new situations or create new tasks that incorporate the learned ones	arrange, build, combine, compose, construct, create, design, initiate, make, originate.	Develop a skill pattern that can enable the one to sing and dance at the same time according to a particular tone

It is important to note that most of the classroom assessment techniques including teacher-made tests, terminal and annual examinations, practical, assignments, exercises, projects and quizzes cover all the three domains that is cognitive, affective and psychomotor. This should be done seriously since most of the national examinations focus on cognitive domain.

2.7 Assessment Procedures

When doing assessment, teachers and tutors should follow appropriate procedures. By doing so, they will be able to prepare valid and reliable assessment tools. These can also help them assess relevant learning outcomes, as stipulated in the syllabus and or course outlines.

There are several assessment techniques and tools as explained earlier on. In this section, special attention is paid to testing, which is the most popular assessment technique for getting information about students' achievement.

In conducting assessment by using tests, teachers should follow these procedures:

(a) Preparing a table of specifications;
(b) Preparing test items based on the table of specifications
(c) Moderating or reviewing test items
(d) Administering the test.
(e) Marking or scoring the test.

2.7.1 Preparing a table of spacifications

The table of specifications (ToS) is also known as a test blueprint or test grid. This is a two-way chart which relates the instructional objectives/competences and the content. It aims to balance what is to be examined and what was taught. Practically, ToS is developed before the test is written.

2.7.2 The importance of the table of specifications

The table of specifications is important in developing quality test items in the following ways:

(i) Acts as a guide for test construction;
(ii) Ensures validity and reliability of the test;
(iii) Ensures that there is a match between what is taught and what is tested;
(iv) Ensures that the examination measures learners' learning across various cognitive levels;
(v) Identifies the weight for each topic in the test;
(vi) Ensures the presence of representative samples of questions in the test as per topics' relative weight; and
(vii) Ensures that the test is balanced and fair to all learners.

2.6.3 Procedures for constructing the table of specifications

Teachers can use various approaches to construct ToS. However, the following is a common and simple approach:

(i) Outline the topics to be covered and the instruction time for each topic, as shown in Table 7.

Table 7: Topics to be Covered and Instructional Time

Topics	Instructional Time for the Topic
Our nation	1
Promotion of life skills	3
Human rights	2
Responsible citizenship	1
Work	2
Family life	1
Total Instructional Time	**10**

(ii) Calculate the percentage weight of the topic using the following formula:

$$\text{Weight of the topic} = \frac{\text{Time spent in the topic}}{\text{Total instructional time}} \times 100$$

Fill the data into the table, as shown in Table 8.

Table 8: Topics, Time Spent for Teaching and Percentage Weight of each Topic

Topics	Instructional Time for the Topic	Percentage Weight of the Topic
Our nation	1	10
Promotion of life skills	3	30
Human rights	2	20
Responsible citizenship	1	10
Work	2	20
Family life	1	10
Total	**10**	**100**

(iii) Determine the number of questions to be included in the test. For example, if the total number of questions in a test is 20, calculate the number of questions per specific topic using the following formula:

$$\text{Number of questions per specific topic} = \frac{\%\text{age weight of the topic}}{100} \times \text{total number of questions in the paper}$$

E.g., The number of questions for the topic of Promotion of life skills = $\frac{30}{100} \times 20 = 6$ questions, as shown in Table 9.

Table 9: Topics, Instructional Time and Number of Questions per Topic

Topics	Instructional Time in the Topic	Percentage Weight of the Topic	Number of Items per Topic
Our nation	1	10	2
Promotion of life skills	3	30	6
Human rights	2	20	4
Responsible citizenship	1	10	2
Work	2	20	4
Family life	1	10	2
Total	**10**	**100**	**20**

(iv) Decide on the type of test items you want to use in your test/examination e.g. Multiple choices, matching items or essays, based on the coverage of the topics in the syllabus.

(v) Distribute the questions on each topic based on the competences/objectives intended to be tested, as shown in Table 10.

Table 10: Sample of a Table of Specifications for the Civics Subject

Topics	Skills to be Tested (Objectives)						Number of Questions per Topic	Percentage Weight per Topic
	Remembering	Understanding	Applying	Analysing	Evaluating	Creating		
Our nation		1		1			2	10
Promotion of life skills	1	2	1		1	1	6	30
Human rights			2			2	4	20
Responsible citizenship			1	1			2	10

Topics	Skills to be Tested (Objectives)						Number of Questions per Topic	Percentage Weight per Topic
	Remembering	Understanding	Applying	Analysing	Evaluating	Creating		
Work		1		1	2		4	20
Family life	1			1			2	10
Total Number of Questions per Skill	**2**	**4**	**4**	**4**	**3**	**3**	**20**	
Percentage Weight per Skill	**10**	**20**	**20**	**20**	**15**	**15**		**100**
Total number of items							**20**	
Total percentage weight								**100**

2.8 Preparing Test Items

The construction of good test items is a science and an art as it involves planning, preparing, administering, scoring, statistically analysing and reporting the results. It is a science because it involves a systematic procedure to design and construct the test item. This includes the proper use of language by considering the right choice of words to convey the intended meaning. It is an art because it involves the beauty or aesthetic aspect of the test item, including formatting, item arrangement, printing style, and the general presentation of the examination. However, the level of difficulty of the item will be determined by the task included in the item. Thus, the skills it requires are similar to those found in effective teaching, which include the following:

(a) A thorough grasp of the subject matter;
(b) A psychological understanding of learners;
(c) Persistence (perseverance, determination, diligence, pushiness);
(d) A sound judgement;
(e) A touch of creativity.

(f) A clear conception of the desired learning outcomes; and
(g) Ability to use different test formats.

The commonly used types of test items in assessment include the following:

(a) Multiple choice items
(b) Matching items
(c) True/False items
(d) Short answer items
(e) Essay type questions

2.8.1 Multiple Choice Items

Multiple-choice items are a good measuring instrument in assessment since they can cover a broader area and can measure various skills. A multiple choice item consists of a stem, which presents a problem/situation, and several alternatives (options/choices), which provide possible solutions to the problem. The stem may be a question or an incomplete statement. The alternatives include the correct answer and several plausible wrong answers, called distracters. Examples (i) and (ii) are samples of multiple choice items.

(i) In a light experiment, results showed that less light was transmitted and the image was distorted. Which type of material was used? (Physics)
   A. A translucent material
   B. An opaque material
   C. A luminous material
   D. A tranaparent material
   E. A non-luminous material

(ii) The type of material which allows a small part of light to pass through it is called (Physics)

A. a translucent material.
B. an opaque material.
C. a luminous material.
D. a transparent material.
E. a non-luminous material.

Although stated differently, both stems pose the same problem, though, the incomplete statement is briefer than the question. However, the question form is more clear because it is the normal way we ask questions in the day life situation.

Forms of Multiple Choice Items

(a) One correct answer

This is the question that has only one answer to be chosen.

For example:

What arithmetic process should be used in deciding how many oranges each child should get when distributing twelve oranges equally among three children?

A Addition
B Subtraction
C Multiplication
D Division*

(b) Incomplete statement

In this question, the stem is presented as an incomplete sentence whereby the student has to complete the sentence with the correct statement.

For example:

Under the electron microscope, you observe a cell with three different types of large organelles, each bounded by two membranes. The cell is most likely to be from.

A a plant.
B an animal.
C a fungus.
D a bacterium.
E a virus.

(c) Direct question

The direct question is looking for either one correct or one false statement. Each of the five responses is related to a common concept and is either true or false.

For example:

Which statement about chemical bonds is

CORRECT?

A A covalent bond forms between a sodium ion and a chloride ion.
B A hydrophobic interaction links an oxygen atom to the hydrogen atoms in a water molecule.
C A covalent bond links an iron atom to the protein haemoglobin.
D An ionic bond binds complementary base pairs together in a double-stranded DNA molecule.
E A hydrogen bond forms between water molecules.

(d) Clustered statements

A question with more than one sentence is often needed to present the information required to answer the question or solve a problem.

For example:

A microscopic observation demonstrates that the mitotic phase is one hour long. How long does the post-synthesis ‘gap’ take in this case?

A 1 hour
B 6 hours
C 9 hours
D 11 hours
E 23 hours

(e) Combined response

This form identifies which of the available statements (i-v or 1 to 5) are correct and then select the response (a, b, c, d, or e) that contains these statements.

For example:
Which of the following capacitors are suitable for use in a high frequency circuit?

(1) Air capacitor
(2) Mica capacitor
(3) Ceramic capacitor
(4) Electrolytic capacitor

A (1), (2) and (3) *
B (1), (2) and (4)
C (1), (3) and (4)
D (2), (3) and (4)
E (2), (3), and (4)

(f) Best answer

In these questions, all distracters are correct, but only one is the most correct or best answer.

For example:

Which of the following substances is the poorest conductor of heat?

A Air*
B Glass
C Brick
D Water

Uses of Multiple-Choice Items

(i) They are used to measure various learning outcomes from simple to complex.

Example 1: *Remembering*

- Identify the name of the source of historical information characterized by the narration of past events. (History)
  A Oral traditions*
  B Archives
  C Museum
  D Archaeology
  E Linguistics

Example 2: *Applying*

- What temperature would be experienced by the tourists in Manyara at 1500 m above sea level, if the temperature of Zanzibar, at sea level is 32°C? (Geography)

  A 23°C*
  B 9°C
  C 19°C
  D 0.6°C
  E 17°C

Example 3: *Analysing*

- Which characteristic of animals in the amphibian group
  differentiates them from birds? (Biology)
  A. Having warm blood
  B. Short bones
  C. Laying eggs
  D. Living in water*
  E. Living in land

(ii) They are used to test the broad subject matter content and different skills at a time.

Strengths of Multiple-Choice Items

(i) Flexibility in measuring all levels of cognitive skills.
(ii) Allow a wide coverage of content and objectives.
(iii) Provide highly reliable test scores.
(iv) Reduce guessing as compared to true-false items.

Limitations of Multiple-Choice Items

(i) Difficult to construct.
(ii) Time consuming to construct.
(iii) Cannot measure some problem-solving skills in Mathematics and Science.
(iv) Cannot measure the ability to organise and present ideas as compared to essay type questions.

Guidelines for Constructing Multiple-Choice Items

The following are the guidelines for constructing good multiple choice items:

(i) Each item should be designed to measure an important learning outcome.

When constructing an item, focus on the functioning content of the item and avoid including irrelevant materials or less significant content to increase item difficulty. The purpose of each item is to test the extent to which the intended learning outcomes have been achieved. Thus, multiple choice items can measure any skills in Bloom's taxonomy, depending on the learning outcomes that are measured.

(ii) The stem should state the problem clearly.

When constructing multiple choice items, fully develop the problem or question in the stem.

The task set forth in the stem of the item should be clear for the student to understand it without reading the alternatives. The clarity and

completeness of a multiple-choice stem is determined by whether it could be answered with or without the choices. Consider the following examples:

A poorly constructed item

Non-metallic minerals (Chemistry)

A Coal, gold and copper.
B Natural gas, silver and diamond.
C Oil, diamond and coal.
D Copper, silver and oil.
E Gold, natural gas and diamond.

The stem is not well constructed because does not present a definite problem.

A well constructed item

Suppose you were given several types of minerals and you were asked to select three non-metallic minerals. What would be your selection from the following alternatives?

A Coal, gold and copper
B Natural gas, silver and diamond
C Oil, diamond and coal
D Copper, silver and oil
E Gold, natural gas and diamond

The problem is clearly stated in the stem i.e. to select the non-metallic minerals.

(iii) The stem of the item should be stated in simple and clear language.

The problem in the stem for a multiple-choice item should be stated as precisely as possible and should be free from unnecessary complex wording. Accordingly, the candidate will be able to answer the question correctly. Poorly-stated item stems hinder a knowledgeable student from

responding correctly since it can create ambiguity. The following examples are illustrative.

A poorly constructed item

A Mathematics professor wants to determine whether or not his teaching of the unit on probability has had a significant effect on his learners. He decides to analyse their scores from the test they took before the instruction and their scores from another examination taken after the instruction. Which of the following tests is appropriate to use in this situation? (Basic Mathematics)

A Dependent samples

B Heterogeneous samples

C Homogeneous samples

D Independent samples

E Uniform samples

The stem is too long and unclear.

A well-constructed item

Which statistic for t-test should be used when analysing learners' pretest and post-test scores to determine if teaching has had a significant effect?

A Dependent samples

B Heterogeneous samples

C Homogeneous samples

D Independent samples

E Uniform samples

The stem is short and clear.

(iv) Observe conciseness in the stem of the item.

Conciseness and clarity of expression are important goals to strive for in test construction. Avoid repeating the same words in each of the alternatives. By moving all of the common content to the stem, you can clarify the problem and reduce the time the student needs to read the alternatives. The following item have been improved by considering this rule.

A poorly constructed item

To which direction do monsoon winds move during the cooler months? (Geography)

A They move towards the northeast
B They move towards the northwest
C They move towards the southeast
D They move towards the southwest *
E They move eastward

There is a repetition of the same material in each of the alternatives.

A well-constructed question

To which direction do monsoon winds move during the cooler months?

A Northeast
B Northwest
C Southeast
D Southwest *
E Eastward

The repeated materials have been removed from the alternatives.

(v) The intended answer should be the only one which is correct

When setting a multiple-choice item, make sure that there is only one correct answer, and it must be unquestionably correct. All forms of ambiguity

must be avoided. It is also necessary to check each of the distracters in the item to make sure that none of them could be defended as the correct answer. This will not only improve the quality of the item but also prevent a disruption in discussing test results.

A poorly constructed item

How long does an *annual plant generally live? (Geography)*

A It dies after the first year.*
B It lives for many years.
C It lives for more than one year.
D It needs to be replanted each year.*
E It lives for more than two years.

There are two correct answers: A and D.

A well-constructed item

How long does an *annual plant generally live?*
A Only one year.*
B Many years.
C One and half years.
D Three years
E Two years

There is only one answer, which is A.

(vi) Distracters should be plausible and attractive.

Distracters in a multiple-choice item should be appealing to the learners who lack the knowledge called for by the item that they prefer one distracter to the correct answer. The art of constructing good multiple-choice items depends heavily on the development of effective distracters. You can do several things to increase

the plausibility and attractiveness of distracters, such as:

(a) Using common misconceptions or learner errors as distracters.
(b) Stating the alternatives in the language of the student.
(c) Using 'good-sounding' words, such as accurate and important, in the distracters and in the correct answer.
(d) Making distracters similar to the correct answer in both length and wording complexity.
(e) Making the alternatives homogeneous, but in doing so, be aware of fine discriminations that are educationally insignificant.

A poorly-constructed item

Which substance is the poorest conductor of heat? (Physics)

A Air*
B Glass
C Brick
D Iron
E Wet wood

This question is not well constructed because alternative D "iron" is a good conductor of heat. Hence, it is not plausible as it is a clue to the incorrect answer.

A well-constructed question

Which substance is the poorest conductor of heat?

A Air*
B Glass
C Brick
D Water
E Wet wood

This question is well constructed because it increases the plausibility of the distracters and calls for a more significant type of discrimination.

(vii) Options should be grammatically parallel and consistent with the stem.

Examinees are quick to take advantage of extraneous clues such as inconsistent stems and options, which easily attract them.

The correct answer is usually carefully phrased so that it is grammatically consistent with the stem. Unless care is taken to check them against the wording in the stem and in the correct answer, they may be inconsistent in tense, article, or grammatical form. This could provide a clue to the correct answer or make some of the distracters ineffective.

A general step that can be taken to prevent grammatical inconsistency is to avoid using articles ‘a’ or ‘an’ at the end of the stem of the item, as shown in the following example:

A poorly-constructed question

A piece of land that is surrounded by sea water is called an (Geography)

A gulf.
B bay.
C island.
D lagoon.
E port.

There is no grammatical agreement between the stem and the alternatives A, B, D and E because of the use of article "an" at the end of the stem. Hence, a student can easily guess the correct answer "C" from the grammatical point of view. This is because "an" goes with a vowel and not a consonant (with exception of special words that are pronounced as vowels, e.g. hour, honest, etc.)

A well-constructed question

A piece of land that is surrounded by sea water is called

A a gulf.
B a bay.
C an island.
D a lagoon.
E a port.

There is grammatical agreement between the stem and the alternatives.

(viii) The correct responses should not be consistently shorter or longer than the distracters.

If the examiner consistently writes correct options which are different in length from the distracters,

examinees will quickly learn to select correct answers based on these peculiarities.

(ix) Each item should be independent of the other items in the test.

Occasionally, information given in the stem of one item will help the learners to answer another item. This can be rectified easily by carefully reviewing the items before they are assembled into a test.

A different type of problem occurs when the correct answer to an item depends on knowing the correct answer to the item preceding it. The student who fails to answer the first item, of course, will not have a basis for answering the second. Such chains of interlocking items should be avoided. Each item should be independently scored.

(x) Avoid using the alternative 'all of the above,' and use 'none of the above' with extreme caution.

When test makers are having difficulty in locating a sufficient number of distracters, they frequently resort to the use of 'all of the above' or 'none of the above' as the final option. These special alternatives are seldom used appropriately and almost always render the item less effective than it would be without them.

The inclusion of 'all of the above' as an option makes it possible to answer the item based on partial information. Since learners are to select only one answer, they can detect 'all of the above' as the correct choice simply by noting that two of the alternatives are correct. They can also detect it as a wrong answer by recognising that at least one of the alternatives is incorrect. Another difficulty with this option is that some learners,

recognising that the first choice is correct, will select it without reading the remaining alternatives.

For example;
Which group of organisms is Mammalia?
(Biology)

A Cat, cow and goat
B Elephant, lion and raven
C Giraffe, raven and lion
D Mouse, raven and lion
E All of the above

It is easy for examinees to eliminate the alternative 'All of the above' after recognising that one of the organisms such as raven is not a Mammalia.

(xi) The position of the correct answer should be randomly varied.

The correct answer should appear in each alternative position about the same number of times, but its placement should not follow a pattern that may be apparent to the person taking the test. Learners who detect that the correct answer never appears in the same position more than twice in a row, or that A is the correct answer on every fourth item, are likely to obtain a higher score than their knowledge would warrant. Such clues can be avoided by random placement of the correct answer.

(xii) The normal rules of grammar should be followed.

If the stem is in question form, begin each alternative with a capital letter and end with any appropriate punctuation mark. Omit the full stop with numerical answers to avoid confusion with decimal points. When the stem is an incomplete

statement, start each alternative with a lower case letter and end with a full stop.

2.8.2 Matching Items

Matching items are presented in two columns. One column consists of stems or problems to be answered, and another column consists of responses from which the answers are to be chosen. Traditionally, the column of stems/premises is placed on the left and the column of responses is placed on the right, as seen in Example 1.

Example 1

Match the major functions of the brain in **Column A** with the brain structures in **Column B** by writing the letter of the correct response beside the corresponding item number in the answer booklet provided (Biology).

Column A		**Column B**	
(i)	Helps to initiate and control rapid movement of the arms and legs.	A.	Basal ganglia
(ii)	Serves as a relay station connecting different parts of the brain.	B.	Cerebellum
(iii)	Involved in the regulation of basic drives and emotions.	C.	Corpus callosum
(iv)	Helps to control slow, deliberate movements of the arms and legs.	D.	Hypothalamus
(v)	Connects the right and left cerebral hemispheres.	E.	Limbic system
		F.	Medulla
		G.	Thalamus

Column A	(i)	(ii)	(iii)	(iv)	(v)
Column B					

This is a well-constructed item because the content tested comes from the homogeneous context i.e. all of the premises specify functions served by brain structures and all the responses are brain structures.

Example 2

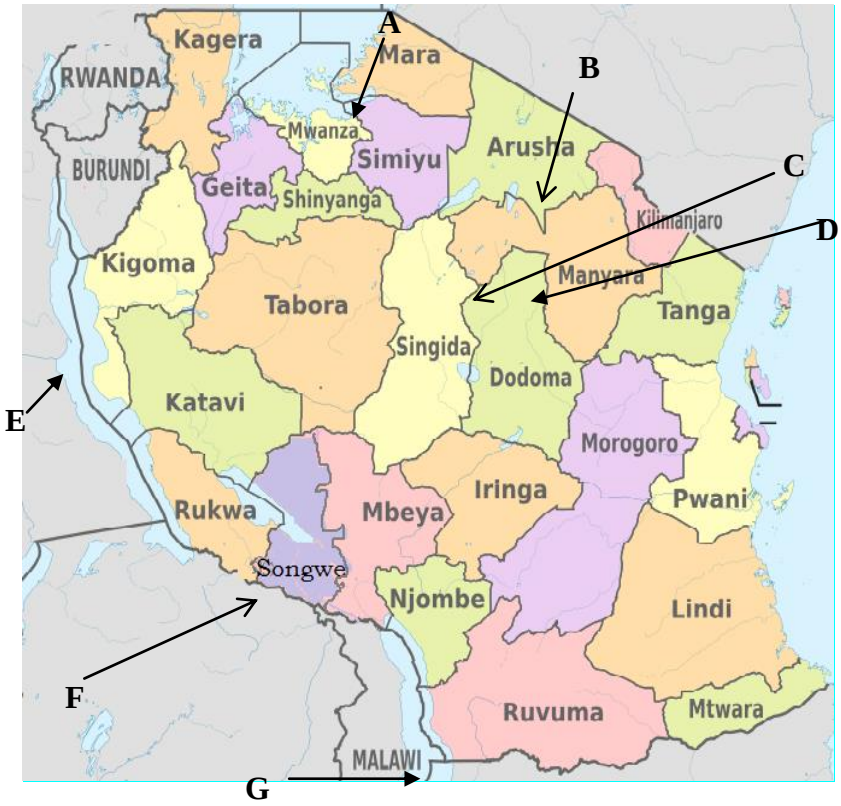
Match the descriptions in **List A** with the responses in **List B** by writing the letter of the correct response beside the corresponding item number in the answer booklet(s) provided. (English Language)

List A	List B
(i) The one who specialises in making furniture. (ii) The type of poem which contains 14 lines. (iii) How do you call the father of your father? (iv) They are trusted to act for another in any business transaction. (v) It is a literary genre characterised by many incidents and characters.	A. Agent B. Carpenter C. Undertaker D. Great Grandfather E. Novel F. Sonnet G. Agent

This is a poorly constructed item because there is no homogeneity of items. The same question includes the descriptions of people, poetry, family relationship and novel, which is contrary to one of the principles, that is the inclusion of homogeneous materials in a single matching exercise. Besides, in List B, responses A and G are the same. Another principle that has been violated is that the responses are not organised in order.

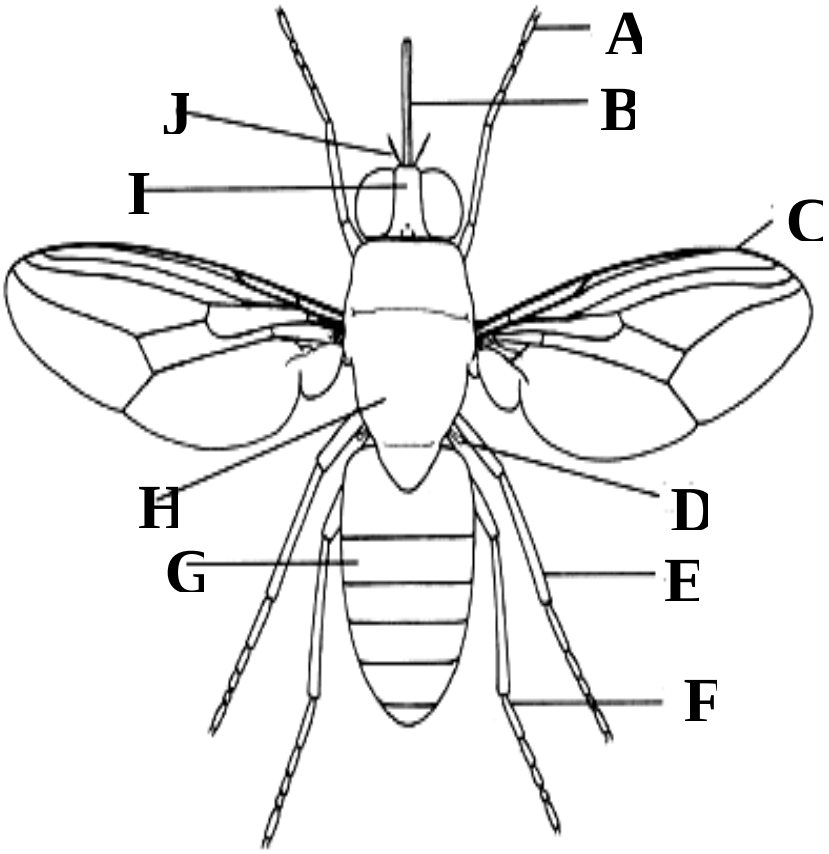
Other Forms of Matching Items

Example I



Directions: The following are the names of the lakes found in Tanzania. In the space provided after each lake, write the letter that represents it on the given map:

1. Eyasi: _____
2. Manyara : _____
3. Natron: _____
4. Nyasa: _____
5. Rukwa: _____
6. Tanganyika: _____
7. Victoria: _____



Directions: The following are parts of a housefly.Write the letter that represents each part on the given diagram in the space provided after each:

1. Abdomen: ___
2. Antenna: ___
3. Foreleg: ___
4. Halter: ___
5. Head: ___
6. Hind leg: ___
7. Mid leg: ___
8. Proboscis: ___
9. Thorax: ___
10. Wing: ___

Uses of Matching Items

Matching items are used for:

(i) measuring the relationship or association between two things;

(ii) relating terms and definitions, objects or pictures and labels, symbols and proper names, causes and effects, scenarios and responses, as well as principles and scenarios to which they apply;

(iii) connecting words, sentences or phrases in one column to corresponding words, sentences or phrases in a second column;

(iv) relating pictorial materials and words; and

(v) identifying positions on maps.

Strengths of Matching Items

(i) Provide objective measurement of learners' achievement or ability.
(ii) Provide highly reliable test scores.
(iii) Allow coverage of more content.
(iv) Provide scoring efficiency and accuracy.

Limitations of Matching Items

(i) Cannot measure learning objectives beyond simple recall of information.
(ii) Difficult to construct since they base on homogeneous material.

Guidelines for teachers in Constructing Matching Items

The following are the guidelines for examiners in constructing matching items:

(i) Limit matching items to homogeneous material.

Possibly the most important guideline to remember when constructing matching items is to ensure that the lists contain homogeneous content. That is, the lists should be from a common theme. What should be avoided is including heterogeneous material in your lists. Consider the following example:

On the line to the left of each compound in Column I, write the letter of the compound's formula presented in Column II. Use each formula only once

Column I	Column II
(i) ____Water (ii) ____Table salt (iii) ____Ammonia (iv) ____Sulfuric Acid (v) ____Alkali (vi) ____Marble	A. $CaCO_3$ B. H2O C. H2SO4 D. HCl E. NaCl F. NaOH G. NH3

List I	(i)	(ii)	(iii)	(iv)	(v)
List II					

In the given example, all items in Column I are names of compounds, and items in column II are formulae of compounds in column I.

(ii) Include more responses than premises.

By including more responses than premises, you reduce the chance that an uninformed candidate/student can narrow down options and successfully match items by guessing.

Indicate whether responses may be used once, more than once or not at all (if needed at all). By adding this statement to your directions, you reduce the impact of guessing.

(iii) Limit the number of items.

For several reasons keeping the list of items fairly brief is desirable. It is easier for the test setter to ensure that the lists are homogenous when they are brief. A student taking the test can easily read and respond to a shorter list of items. Although there is no consensus on the number of items to include in a matching list, a maximum of five is reasonable to ensure homogeneity.

(iv) Place all items on the same page.

Keep the directions and all items of a single matching exercise on one page. It greatly reduces efficiency in responding if the learner must turn the page looking for responses. The learner is also more likely to transpose a letter or number if they have to look back and forth across two pages, leading to errors in measuring what the student has learned.

2.6.1 True/False Items

A true/false question consists of a statement that requires a True or False response. Other variations of the true/false format are Yes or No, Correct or Incorrect, Agree or Disagree and Fact or Opinion. Effective true/false questions are factual -based, rather than opinion-oriented. These are designed to quickly and efficiently test learners' knowledge about a particular idea or concept. The following example is worth considering:

For each of the items (i) – (v), write **True** for a true statement or **False** for a false statement in the space provided. (Geography).

(i) The hydrological cycle is a continuous circulation of water from the earth's surface to the atmosphere……
(ii) Wind is air in motion from a low pressure to high-pressure area……..
(iii) A rain gauge is an instrument used to measure the amount of rainfall……
(iv) The side of the mountain facing the direction of the wind is known as the leeward side……..
(v) Bush fallowing is an agricultural system in which land is cultivated and left for some years to recover its fertility……..

Uses of True/False Items

True/False items are used to
(i) test a wide content;
(ii) get facts; and
(iii) reflect on the materials learned.

Strengths of true/false items

The following are the strengths of true/false items:
(i) Easy to create
(ii) Enable coverage of a wide area of knowledge.

(iii) Can be constructed out of a few words, which make them less dependent on the learners' ability to read carefully
(iv) Take a short time to be answered: On average, a learner can answer 3 to 4 true/false items per minute
(v) Easy to grade.

Limitations of true/false items

(i) Learners have 50% chance of guessing answers if the items are not well constructed.
(ii) It is difficult to measure accurately a learner's actual understanding of the content using the results of a test composed of true/false questions.
(iii) They can be too easy.

Guidelines for Constructing True/ False Items

The following are important things to adhere to when constructing true/false items:

(i) Concentrate on one key idea or concept. Every true/false item should focus on one specific topic. The primary reason for this is that true or false items are limited. They do not call for a short answer response, nor are there multiple answers to choose from. The student is simply being asked to declare whether or not the statement or assertion is based on truth.

(ii) Keep both true and false statements similar in length.

(iii) Include a good mix of true and false answers.

(iv) Avoid using an abundance of words in your true and false question to make it more challenging.

(v) Do not use negative statements when constructing true/false items. Many test writers may add the word **NOT** to a true statement simply to trick learners.

(vi) For a statement to be true, each part must be true. One detail can make an entire statement false.

(vii) True/false questions may use words called "absolutes" or "qualifiers". Absolute words imply there are no exceptions to the facts stated in the question. Examples of such words are *never*, *none*, *always*, *all*, *every* and *only*. These words tend to make a statement false (but not always). Qualifiers such as *some*, *few*, *often*, *many* and *frequently* limit meaning, thus allowing exceptions and possibilities that can make a question true (but not always). See the following examples.

Example 1:
Everyone should exercise daily. (Biology)

Answer: False: Due to the word everyone. There are many people who do not exercise daily. Besides, others may be unable or prohibited from exercising.

Example 2:
All types of cars have some type of engine. (Mechanical Engineering)

Answer: True. Even though the absolute term "all" could tend to make this question false, the qualifier "some" makes the question more general and allows for possibilities ("some type of engine": does not have to be the familiar gasoline- driven engine).

Example 3:
Children sometimes have emotional problems following their parents' divorce. (Civics)

Answer: True. Because the qualifier sometimes limits the scope of this statement, allowing for exceptions.

(viii) Avoid using double negatives. Double negatives (two negative words in a sentence) tend to indicate a positive relationship in Standard English usage, just as in algebra multiplying two negative numbers equals a positive number. See the following example.

Aspirin is not an illegal drug. (Civics)

Hint: Cancel the negatives to turn the question into a positive statement; then, select your answer. (Aspirin is not an illegal drug = Aspirin is a legal drug.) **Answer: True**

2.6.2 Short Answer Items

Categories of short answer items

(i) Completion/Fill in the Blanks/Insert Items
(ii) Short answer items: structured and unstructured
(iii) Jumbled items: Jumbled sentences

Short answer and the completion items are both supply-type test items that can be answered by a word, phrase, number, symbol or a simple sentence. They are essentially the same, differing in the method of presenting the problem. The short-answer item uses a direct question, whereas the completion item consists of an incomplete statement. In contrast, jumbled items assess learners' ability to process sentence information in units larger than a single word. They assess organization skills in reading comprehension. They also assess the ability to organize facts, concepts, principles and procedures; and to produce a coherent and cohesive response.

Examples 1: Short Answer and Completion Items

(i) What is the moisture content of sample L? (Chemistry)

(ii) The moisture content of sample L is……(Chemistry)

(iii) In which year did Tanganyika gain her independence? (History)

(iv) Write symbols that are used to represent water. (Chemistry)

(v) Briefly describe the main character in the novel titled Passes Like a Shadow by B.M. Mapalala. (English Language)

(vi) Given that the local time at Greenwich Meridian is 12:00 noon, what will be the local time at Mikindani which is 100S 400E? (Geography)

Example 2: Jumbled Items

- Re-arrange the following sentences into a logical sequence to make a meaningful paragraph by writing the corresponding letter in the answer booklet provided.

  A It was so unfortunate that the man died in Amina's house because of drunkenness.

  B Amina was struggling to support him to get into the house and I went to help her.

  C One day, I saw her coming back home with a man.

  D There was a certain girl living near our house; she was known as Amina Matanuzi.

  E The man was so drunk that he could not walk on his own, and he looked so terrible.

Uses of Short Answer Items

Short-answer items are suitable for measuring a wide variety of learning outcomes such as:

(i) Knowledge of terminologies

E.g. What is the name of the line that divides the circle into two equal halves? (Geography).

(ii) Knowledge of specific facts
E.g. How Theory X differs from Theory Y?

(iii) Knowledge of principles
E.g. If the temperature of a gas is held constant while the pressure applied to it is increased, what will happen to its volume? (Chemistry).

(iv) Knowledge of method or procedure
E.g. The process of judging, analyzing and comparing the educational achievement is called…… (Research Measurement and Evaluation).

(v) Simple interpretation of data
E.g. What does the correlation coefficient of 0.85 between two variables imply? (Research Measurement and Evaluation).

(vi) Ability to solve problems/numerical problems. E.g.
- Solve the quadratic equation $x^2+9x+18=130$, where *x* is a positive integer. (Basic Mathematics)
- Badili sentensi zifuatazo katika hali ya kutendwa. (Kiswahili)

(vii) Skill in manipulating procedures/mathematical symbols. E.g.
- Make ***a*** the subject of the formula *F* = *ma* (Basic Mathematics)
- Describe the two initial steps to follow in a successful preparing an effective teaching resource. (Educational Media and Technology).

(viii) Ability to complete and balance chemical equations. E.g.
Balance the following chemical equation: (Chemistry)
H_2O = H_2 + O_2

Strengths of Short Answer Items

(i) They are easy to construct partly because of the relatively simple learning outcomes they usually measure.

(ii) Answers are supplied by the learner, not pre-selected as in objective questions, reducing chances of guessing.

(iii) They are precise and specific as to the scope and length of answers.

Limitations of Short Answer Items

(i) They are unsuitable for measuring complex learning outcomes (If not well phrased).

(ii) They are difficult in scoring, especially when the question is not well phrased.

Forms of Short Answer items

(i) Selective Recall
Example: Sera ya kidini ya Akbar ni ipi hasa?

(ii) Evaluative Recall
Example: Given that the local time at Greenwich Meridian is 12:00 noon, what will be the local time at Mikindani, which is 100S 400E?

(iii) Comparison of two things on a single designated basis
Example: Compare the uses of a stale cheque to those of an uncredited cheque.

(iv) Comparison of two things in general.
Example: Compare the activities in the Early Stone Age with those of the Late Stone Age.

(v) Decision for or against
Example: “The atmosphere is multiple layered in one layer”. Do you agree with this statement? Support your answer using four points.

(vi) Causes or effects
Example: Account for the need to observe the laboratory rules.

(vii) Explanation of the use or exact meaning of a phrase in a passage or a sentence
Example: As an expert in Book Keeping which errors would you encounter when preparing a trial balance agreement?

(viii) Classification
Example: Soma kwa umakini orodha ya maneno iliyopo kwenye sanduku la maneno na kisha yapange katika makundi yafuatayo:
Vivumishi
Viwakilishi
Vitenzi
Vihisishi

(ix) Application of rules or principles in given situations
Example: Which principle would you apply when teaching kids how to swim?

(x) Summarise a unit of the text or an article
Example: Shorten the following passage to not more than 40 words.

(xi) Analysis
Example: The relative density of a substance is 2.5. What is its density?

(xii) Relationship items
Example: What is the relationship between the Uhuru torch and the national flag in terms of their functions?

(xiii) Illustration of your own examples of principles in science, language, etc.

Example: Using two examples, state the importance of Tanzanian national emblems/symbols.

(xiv) Factual/definitional short-answer questions: Describe or define things.
Example: Using two reasons briefly explain the effects of using charcoal on the environment.

(xv) Interpretive short-answer questions: Ability to apply your knowledge
Example: With two reasons explain why there is no work done on the books when carried horizontally?

(xvi) Graphing questions
Graphing questions typically require an answer in the form of an graph.
Example: Draw a diagram of a digestion process showing the relationship between the type of food and the responsible enzyme.

(xvii) Calculation questions
Calculation questions require the candidate to calculate or compute a numerical answer.
Example: "In a sales promotion, the price of a shirt costing Tshs. 15,000/- is reduced by 15%. What is the new price of the shirt?

Guidelines for Examiners in constructing Short-Answer/Completion /Jumbled Items

(i) Word the item so that the required answer is brief and specific as seen in the following examples.

A poorly constructed item

An animal that eats flesh of other animals is ..... (Biology)

The question is too general, so there is no specific answer.

A well constructed item

In which class does the animal that eats flesh of other animals belong?

The question is specific, it requires classification.

(ii) Do not use statements directly from text books as a basis for short-answer items.

A poorly constructed item

Chlorine is a …….(Chemistry)

A well constructed item

In which group of elements does chlorine belong?

(iii) A direct question is generally more desirable than an incomplete statement.

(iv) If the item is to be expressed in numerical units, indicate the units of measurement needed. See the following examples.

A poorly constructed item

What is the diameter of the earth? (Geography)
Here the students give multitude of answers e.g. diameter in mm, cm, dm. km, etc.

A well constructed item

What is the diameter of the earth in kilometres?
The answer here is specific, i.e. diameter in km.

(v) When completion items are used, blanks for answers should be placed at the end of the item, and they should be equal in length.

(vi) When you use completion items, do not include too many blanks.

Example:

In the year _____, Mr._____ discovered_____, which led to a _____ which was _____.

Including too many blanks in a question, my distort the meaning of the statement and even confuse students.

(vii) Jumbled sentences/statements must observe coherence and cohesion of ideas, concepts or facts.

2.6.3 Essay Questions

Essay questions require candidates to express themselves. They are suitable for assessing learning at higher cognitive levels that cannot be measured by other means. They allow integration and application of different ideas and problem solving skills. Essay items are easy and quick to construct. Some subjects such as Mathematics, Physics and Engineering mostly, use structured questions in which the candidate is given a diagram or a problem to do manipulations and calculations. Thus, structured questions enable the candidate to think critically and solve problems if well-constructed.

Types of Essay Questions

(i) Extended response questions

Extended response questions allow examinees to select any factual information that they think is pertinent, to organise the answer in accordance with their best judgment and to integrate and evaluate ideas as they deem appropriate.

Examples of extended response questions:

- Kwa kutumia vyombo vya habari vilivyopo nchini hivi sasa, onesha mchango wa kila kimoja katika kukuza na kueneza lugha ya Kiswahili. (Kiswahili)

- Explain how to handle chemicals having the warning signs of corrosive, flammable, harmful, explosive and toxic. (Chemistry)

- Explain how child spacing influences the health of the family. (Biology)

(ii) Restricted response questions

- Limit the content and response.
- Limitations on the form of response are generally indicated in the question.

Examples of restricted response questions:

- Form Three students were asked by their Geography teacher to do a survey around the school compound. In doing this task, explain eight pre-survey activities they need to consider. (Geography)
- Tuberculosis (TB) is among the dangerous and contagious diseases in the country. Educate people about four symptoms and five preventive measures to minimise the spread of the disease. (Biology)

(iii) Structured Questions

Example 1:

Amani and Asha bought Coca-cola and Pepsi drinks for a farewell party. Amani spent 9,950 Tshs to buy 12 bottles of Coca-cola and 5 bottles of Pepsi drinks. Asha spent 8,150 Tshs to buy 9

bottles of Coca-cola and 5 bottles of Pepsi. Formulate a system of linear equations and hence apply the matrix method to find the price of one bottle of each type of drinks. (Basic Mathematics)

Example 2:

Figure 1 shows the profile of a radio wave. Study it carefully and determine its wavelength. (Physics)

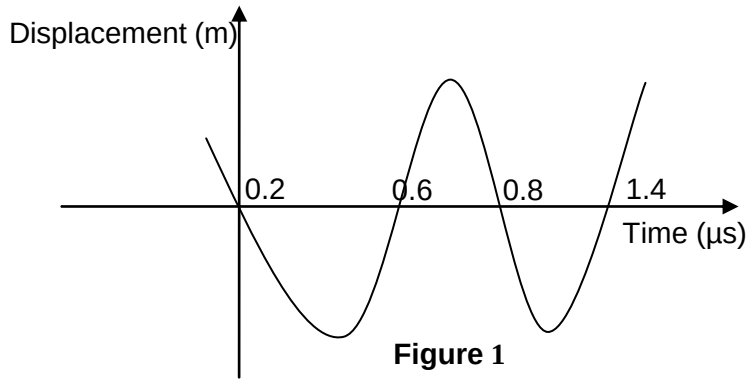


Figure 1

Forms of Essay Questions

(i) Quotation and Discuss questions

These are among the most common types of essay questions. They allow the candidate to argue for or against the statement.

Example: To what extent do you agree with or differ from this statement?

(ii) Double-barrelled questions

Questions containing several issues that need to be answered separately.

Example: Parents today complain of decline in moral conduct among school graduates and argue for change. To what extent do you agree?

(iii) A direct question

Example:

- "Metaphysical poetry values intellect above emotion." To what extent have you found this to be the case? Limit your answer to five (5) points.
- Do you consider the love poetry of these days to be persuasive? Justify using five points.

(iv) A general question which specifies a text, author or genre for you to write about but which does not indicate any particular issue to look at

Example: Write an essay about the impact of HIV Aids on Tanzania. (Limit you response to six social impacts.)

Uses of Essay Questions

Essay questions are used to measure complex learning outcomes or higher order thinking skills that cannot be measured by other means.

Strengths of Essay Questions

The following are the strengths of essay questions:

(i) They are an effective way to measure higher levels of cognitive objectives.
(ii) They are unique in measuring learners' ability to select content, organise, integrate and present it logically.
(iii) They present a more realistic task to examinees.
(iv) Since examinees must present their answers in their own writing, essay questions help to improve writing skills.
(v) They are easy to construct.
(vi) They are quick to construct.

Limitations of Essay Questions

(i) Grading is often subjective and inconsistent, coloured by the student's prior performance,

neatness and handwriting, spelling and grammar, as well as time of the day where the actual test falls in

(ii) Limit sampling of the content.

(iii) Require time for learners to think, organise, write and revise.

(iv) Time consuming to mark.

(v) Favour learners with good writing and verbal skills as opposed to those who have difficulties in learning (visual and kinaesthetic).

(vi) Favour learners who are quick to develop argument, as opposed to those who are slow to do so.

Guidelines for Examiners in constructing Essay Questions

The guidelines for examiners in constructing essay questions are as follows:

(i) The question should be formulated in such a way that it will call forth the behaviour specified in the learning outcomes.

In constructing an essay question, the examiner should know the learning outcome they intend to measure. Hence, the question formed will be focused on the specific behaviour.

For example, if the learning outcome requires the student to use relevant information on biological science and related fields for self-study and lifelong-learning, they will be asked, "'Human beings possess endoskeleton while arthropods possess exoskeleton.' Describe four roles of the skeleton." Therefore, every constructed essay question should respond to a specific learning outcome.

(ii) Each question should be phrased so that the examinee's task is clearly indicated.

One way to clarify the question is to make it as specific as possible. Learners tend to interpret the questions differently and give varying responses if the question contains ambiguous phrasing.

An example of a clear phased question:
Give five reasons that justify that not all fungi are harmful to human beings. (Biology)
In this question, the examinee's task is to justify the given fact.
An example of an ambiguous question:

Give a detailed explanation on Majimaji War. (History)
This question does not indicate specifically what the examiner is looking for. It is too broad. Hence, there are possibilities of having varying answers. Some students may provide factors for which led to Majimaji War others may give effects of Majimaji War and so forth.

(iii) The Approximate time limit for each question should be indicated.

When examinees are given time limit, they can balance their time when responding to different questions. It also helps them to determine the extent to which they respond to the question.

(iv) If possible, optional questions should not be used due to the following reasons:

- When examinees answer different questions, they are actually taking different tests.
- Learners may not study all the subject matter if they know they will have a choice among the questions**.**

(v) The use of essay questions should be restricted to those learning outcomes that cannot be satisfactorily measured by objective items.

The use of essay questions is inevitable as complex learning outcomes such as organisation, integration and expression of ideas are required.

(vi) The number of points required for each question should be specified in order to:

- avoid extra points which will not be marked,
- maintain uniformity in number of points required, and
- save time to both the candidate responding to the question and the examiner since there will be no marking of extra points.

(vii) The number of marks to be earned for the correct response to each question should be indicated.

If time is running short, examinees may have to choose which questions to answer. They will want to work on the questions that are worth the most marks.

2.9 Preparation of a Marking Scheme

A marking scheme is a set of model answer prepared when constructing test items. It indicates **all key points,** the total marks for each item, and their subsequent sections/parts. The marking scheme helps to increase the reliability of test results. In preparing the marking scheme, the examiner/teacher should consider the following:

(a) The answers for each question/item should be prepared during the preparation of test items to avoid forgetting the required responses to a particular question/item.

(b) The answers to the objective questions should be exactly the ones needed or rather required.

(c) Objective questions should have an even distribution of marks i.e. one or $\frac{1}{2}$ mark for each correct response. (Consider the format of a particular paper.)

(d) Essay questions should be assigned some points and descriptions to the various subparts of the response. For example, points for the introduction, body and conclusion.

(e) For examination questions demanding drawings, the marking scheme should indicate the marks for the correct drawing, labelling and caption.

(f) Distribute the marks according to the weight of the task given in each question.

The marking scheme should be neat, and all answers should be very clear/correct in relation to the nature of the question asked. This will reduce bias in marking.

3.0 MODERATION, ADMINISTRATION, SCORING AND RECORDING OF TEST RESULTS

3.1 Moderation of the Test Items

Moderation refers to a range of quality assurance processes conducted by academic experts to ensure that the assessment tools are fair, valid, and reliable and that they are in line with the required standards. Members of the moderation panel at school, council, regional or national levels (depending on the purpose of the test) should be composed of qualified, experienced subject teachers in the subject to be moderated. The important tools required for the moderation process include items to be moderated, the marking scheme, the syllabus, some reference books/teaching and learning materials, dictionaries and the table of specifications.

When moderating the test items, the moderators should check the following criteria:

(a) Do the items reflect the competences to be tested as stipulated in the syllabus?
(b) Do the questions unambiguously communicate the examiner's intentions? i.e., could a student sitting the examination reasonably be expected to know what is required to be done?
(c) Is the marking scheme sufficiently detailed to allow efficient and consistent marking?
(d) Do the test items fairly cover the material examined?
(e) Are there any technical errors?
(f) Could the writing style be improved?
(g) Are the vocabulary and weighting of the test items appropriate to the level of learners?
(h) Can the test item be answered in the time allotted?

3.2 Administration of the Test

Test administration should follow examinations' regulations and conditions, which include the following:

(a) All students in a class must sit the same test paper under the same conditions.
(b) Calendars that indicate dates for midterm and end of term tests, as well as ward, council, regional or national examinations should be provided.
(c) All instructions given by the invigilators responsible for the administration of the test should be observed.
(d) Only authorised materials should be allowed into the test/examination room.
(e) Verbal or any other form of communication between students during the test/examination is strictly prohibited.
(f) Cheating in the test/examination room should be avoided.
(g) Students/candidates should not begin the test before authorisation.

3.3 Scoring the Test

Marking the test should be done in such a way fairness is observed to the greatest possible extent. The subject marking panel should ensure that each marker marks one question from the student assessment sheet. This will minimise the chances

of treating a student unfairly. Besides, checkers should be appointed to ensure that each question is marked according to the marking scheme, and the marks of each student are accurately transferred to the score sheet.

Where there is only one subject teacher, the school may contact a nearby inter-school subject panelist to be involved in the marking process. However, if there is no nearby school to form a subject panel, then the subject teacher may mark the' scripts. Thereafter, academic officers may ask teachers from other related subjects to check the marking against the detailed marking scheme prepared by the subject teacher.

However, if the test/examination is administered at council, regional or national level, a panel should be formed by the authority concerned based on the criteria set, which include being a teacher of the specific subject, academically qualified for the level marked and experienced in teaching.

3.4 Recording and Presentation of Learners' Tests/Candidates' examinations

The subject teacher or authority concerned with the administration of tests/examinations is required to prepare/record the following:

(a) Portfolio: This will include all activities given to learners as part of the assessment, marking scheme, worksheets and the score sheet.

(b) Continuous Assessment (CA): One copy will be kept in the student's academic file. Secondary school teachers and college tutors should submit their students' Continuous Assessment scores to the National Examinations Council of Tanzania (NECTA) through the PReMS System. CA involves Test, terminal and annual examinations as they will contribute to the final national examination results.

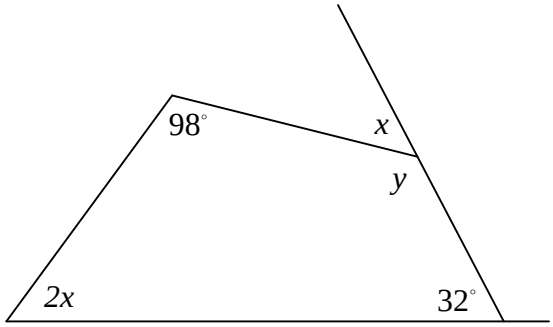
4.0 ASSESSMENT OF LEARNERS WITH SPECIAL NEEDS

Assessment of learners with special needs should consider the nature of their disabilities. Assessment for learners with physical

impairments depends on the type and nature of their disabilities. For example, learners who are unable to write using hands should not be given tests that require them to write, such as essay questions. Learners who are completely unable to write, do examinations orally, and their responses are recorded.

The test for learners with total blindness should be written using Braille Notation. Besides, if the test items contain maps or drawings, they should be transformed into wording. Moreover, tests for learners with low vision should use enlarged font size. Considering the nature of the disability, examiners should give all learners extra 10 minutes for each hour in all subjects, except in Mathematics, whose extra time is 20 minutes for each hour. The following are examples of modified questions to accommodate learners with total blindness.

Table 11: Examples of Modified Questions to Accommodate Learners with Visual Impairment

For Learners without Disability	For Learners with Visual Impairment
1. Find the values of *x* and y in the following figure. (Basic Mathematics)  	1. If $(180^\circ - x)$, $2x$, 32° and 98° are interior angles of a four sided polygon, find the value of *x.*
2. Draw a sketch map of Africa and locate the following by using roman numbers:(History) (i) The British East African colony where Indirect Rule Policy was more successful. (ii) The Portuguese Colony whose nationalist leader was assassinated by a parcel bomb	2. Fill in the blanks with the correct word/words given in brackets in item (i) – (v) by writing the correct answer beside the item number in the answer sheet(s) provided. (i) The British East African colony where Indirect

For Learners without Disability	For Learners with Visual Impairment
in Dar es Salaam. (iii) The country where the apartheid policy was applied. (iv) The German colony that resisted Direct Rule Policy in 1904 – 1907. (v) The British colony where Indirect Rule Policy was first experimented in Africa.	Rule Policy was more successful is called _____. (Tanganyika, Kenya, Nigeria, Uganda). (ii) Which Portuguese colony had its nationalist leader assassinated by a parcel bomb in Dar es Salaam? _____ (Nigeria, Namibia, Mozambique, South Africa). (iii) Where was the apartheid policy applied in Africa? _____ (Rwanda, Senegal, South Africa, Somalia). (iv) During which period did the Nama and Herero resistances against the German rule occur? _____ (1914 to 1918, 1939 to 1945, 1896 to 1897, 1904 to 1907) (v) The British colony where Indirect Rule Policy was first experimented in Africa is called _____ (Ghana, Uganda, Nigeria, South Africa).

For Learners without Disability	For Learners with Visual Impairment
3. Match the descriptions in **List A** with the corresponding political parties in **List B** by writing the correct letter of the corresponding response besides the item number in the answer booklet(s) provided. (History)	3. In items (i) – (v), fill in the blanks with the correct word/words given in the brackets by writing the correct answer beside the item number in the answer sheet(s) provided.

List A	List B
(i) A political party which struggled for the independence of Angola. (ii) A political party which was formed in 1963 in Zimbabwe. (iii) A political party which took active participation in the struggle for the independence of Ghana. (iv) A political party which was formed in 1952 to demand for the independence of Uganda. (v) A political party in Zanzibar that overthrew the oppressive Arab rule in 1964.	A AMNUT B ASP C FRELIMO D MPLA E UGCC F UNC G ZANU H ZPPP

(i) A political party which struggled for the independence of Angola is known as _____ (SWAPO, MPLA, ZANU, FRELIMO)

(ii) A political party that was formed in 1963 in Zimbabwe is called _____. (KANU, ZANU, ZPPP, ANC)

(iii) Which political party participated actively in the struggle for independence in Ghana? _____. (ZNP, ANC, UGCC, UNC)

(iv) The political party that was formed in 1952 to demand for the independence of Uganda is known as _____. (UGCC, ANC, UNC, UNDP)

(v) The political party in Zanzibar that overthrew the Arab rule in 1964 was called _____. (ZPPP, ZNP, ASP, UMMA PARTY)

5.0 MONITORING AND EVALUATION

The monitoring refers to the periodic tracking of school assessment by systematically analysing the tools of assessment. On the other hand evaluation refers to determination of the quality of assessment.

The National Examinations Council of Tanzania (NECTA), School Quality Assurers (SQA), Ward Education Officers, Heads of School and College Principals will be responsible for monitoring and evaluation of assessment to have quality Continuous Assessment (CA).

NECTA, being the body mandated to conduct formative and summative assessment at the national level in Tanzania, is responsible for making follow up on CA to ensure its standard. This is because it contributes to final examination results. Specifically, the role of NECTA in monitoring and evaluation is to analyse the quality of the assessment tools used to generate CA. The tools will include question papers, marking schemes and score sheets.

On the other hand, School Quality Assurers and Ward Education Officers are responsible for ensuring that teaching, learning and assessment processes adhere to the set standards. The processes involve, among others, the preparation of assessment tools, administration of tests and recording of test results. Heads of School and College Principals are mandated to coordinate the implementation of Curriculum, Education Policies and Directives given in collaboration of School/College Committees; hence, they will oversee all the teaching and learning activities, including assessment procedures.

For summative evaluation, NECTA will be analysing each item that has been done. The analysis will indicate the factors that enabled the students to respond to the items correctly and those that hindered them from responding correctly. The aim is to improve teaching, learning and assessment procedures at school level and college levels.

6.0 CONCLUSION

Assessment is an important task which, if well conducted, can inform and contribute to the effective teaching and learning processes. Hence, it helps to improve learners' performance. Although NECTA conducts assessment at national level, teachers and tutors are key players in the assessment process at school and college level. This is because they can largely assess not only the cognitive domain but also the affective and psychomotor domains. However, good assessment requires teachers and tutors with adequate assessment techniques, knowledge and skills. NECTA believes that these guidelines will enhance their competence in all aspects of assessment.

Appendix 1: A Summary of Cognitive Dimensions Assessed in the Final Examinations

Cognitive Levels Assessed	How to Measure	Sample of Action Verbs Used			
Remembering	Students' ability to recall, retrieve or recognise information, ideas, facts and principles in the approximate form in which they learned them	Define Recognise Recall Select Find Show Reproduce Give Mention Indicate	Identify Label List Memorise Name Outline Tell Point Out Locate Which	Duplicate Underline State Spell Repeat Record What Where When	Who
Understanding	Students' ability to translate, comprehend or interpret information, facts or ideas based on prior learning	Explain Summarise Paraphrase Describe Illustrate Approximate Clarify Complete Convert Extrapolate Fill in the Blanks	Convert Exemplify Restate Estimate Add Rewrite Insert Intrapolate Interpolate	Extend Rephrase Translate Indicate Show Specify Retell To What Extent Why	How
Applying	Students' ability to select, transfer, use data, ideas, facts and principles to complete a problem or task in a new situation with minimum direction	Use Compute Solve Apply Manipulate Demonstrate Derive Dramatise	Apply Change Calculate Illustrate Employ Assign Express Illustrate	Imitate Articulate Practice Instruct Simulate Make Use Of	Utilize
Analysing	Students' ability to distinguish, classify, and relate the assumptions, hypotheses, evidence, or structure of a statement or question	Analyse Categorise Compare Contrast Separate Break Down Select Group Order Distinguish Characterize Relate	Discriminate Associate classify Divide Factorise Sequence Tabulate Prioritize	Operate Relate Organise Arrange Inspect Correlate	Differentiat e
Evaluating	Students' ability to appraise, assess, or critique an idea or fact on a basis of specific	Judge Account For Infer Dispute Recommend Critique	Assess Conclude Argue Evaluate Defend Support	Reframe Predict Rate Verify Prioritise	

	standards and criteria	Justify Agree/Disagree Comment On Criticize Decide Deduce Diagnose	Appraise Disprove Grade Inspect Rule On Suggest		
Creating	Students' ability to compose, design, assemble, and combine ideas into a whole product, plan or idea that is new to him or her	Create Design Hypothesise Invent Develop Revise Prepare Discover Animate Enhance Improvise Improve Invent	Collect Combine Compose Construct Arrange Assemble Devise Formulate Join Modify Plot Draw	Plan Rewrite Set Up Synthesise Change Generate Redesign Sketch	Rearrange Reconstruct Relate Propose Integrate Express Compile Reorganise

Note

1. Interrogatives such as ***what*, *how*** and ***why*** can be used to assess all levels, depending on the qualifiers used in the stem of the question.
2. The action verbs such as ***investigate, discuss, debate on, determine*** and ***examine*** fall under the category of **complex action verbs** which, when used to set an examination item, do not show clearly the task of the item, unless many qualifiers are used. It is advised to use these action verbs sparingly. Otherwise, avoid them.

Appendix 2: A Summary of Affective Dimensions

Receiving	Responding	Valuing	Organization	Characterization
Attend to stimuli	**React to stimuli**	**Attach significance to ideas**	**Build value system**	**Internalize values**
Ask	Agree	Aid	Anticipate	that guide behavior
Acknowledge	Allow	Care (for)	Collaborate	Administer
Attend (to)	Answer	Complete	Confer	Advance
Follow	Ask	Complement	Consider	Advocate
Listen	Assist	Contribute	Consult	Aid
Meet	Attempt	Delay	Coordinate	Challenge
Observe	Choose	Encourage	Design	Change
Receive	Communicate	Endorse	Direct	Commit (to)
	Comply	Enforce	Establish	Counsel
	Conform	Evaluate	Facilitate	Criticize
	Cooperate	Expedite	Follow through	Debate
	Demonstrate	Foster	Investigate	Defend
	Describe	Guide	Judge	Disagree
	Discuss	Initiate	Lead	Dispute
	Display	Interact	Manage	Empathize
	Exhibit	Join	Modify	Enhance
	Follow	Justify	Organize	Excuse
	Give	Maintain	Oversee	Forgive
	Help	Monitor	Plan	Influence
	Identify	Praise	Qualify	Motivate
	Locate	Preserve	Recommend	Negotiate
	Notify	Propose	Revise	Object
	Obey	Query	Simplify	Persevere
	Offer	React	Specify	Persist
	Participate (in)	Respect	Submit	Praise
	Present	Seek	Synthesize	Profess
	Read	Share	Test	Promote
	Relay	Study	Vary	Promulgate
	Reply	Subscribe	Weigh	Question
	Report	Suggest		Reject
	Respond	Support		Resolve
	Select	Thank		Seek
	Try	Uphold		Serve
				Strive
				Solve
				Tolerate
				Volunteer (for)

1Waller, K. (n.d.). *Writing instructional objectives*. Retrieved from: http://www.cetla.howard.edu/teaching_resources/Curriculum_Design/docs/Learning%20Objectives.pdf

Appendix 3: A Summary of Psychomotor Dimensions

Perception	Set	Guided Response	Mechanism	Complex Overt Response	Adaptation	Origination
Use of senses to obtain clues	Readiness to take action	Knowledge of the steps required to perform a task	Perform tasks in habitual manner	Skillful performance of motor acts	Skillful performance of motor acts and modification of movement in problematic or new situation	Creating new movement patterns for problematic or new situation; creates new tasks that incorporate learned ones
Choose Describe Detect Differentiate Distinguish Identify Isolate, Relate Select Separate	Begin Display Explain Move Proceed React Respond Show Start Volunteer	Assemble Build Calibrate Construct Dismantle Display Dissect Fasten Fix Grind Heat Manipulate Measure Mend Mix Organize Sketch Work	Assemble Build Calibrate Construct Dismantle Display Dissect Fasten Fix Grind Heat Manipulate Measure Mend Mix Organize Sketch Work	Assemble Build Calibrate Construct Dismantle Display Dissect Fasten Fix Grind Heat Manipulate Measure Mend Mix Organize Sketch Work	Adapt Alter Change Rearrange Reorganize Revise Vary	Arrange Combine Compose Construct Design Originate

2 University of Central Florida. (2005). Program assessment handbook: Guidelines for planning and implementing quality enhancing efforts of program and student learning outcomes. (February 2008 ed.). Retrieved from: https://oeas.ucf.edu/doc/acad_assess_handbook.pdf

Appendix 4: **Extremely Vague Action Verbs or Phrases that Should be Avoided**

Words to Avoid where Necessary	Phrases to Avoid where Necessary
• Believe • Hear • Realise • Capacity • Intelligence • Recognise • Comprehend • Know • See • Feel • Conceptualise • Listen • Memorise • Self-Actualise • Think • Experience • Perceive • Understand	• Show your appreciation for • How acquainted are you with • How are you adjusted to • Show your awareness of • Comprehension of • Cognisant of • Enjoyment of • How conscious are you of • How familiar are you with • Indicate your interest in • Show your knowledge of • How knowledgeable are you about • Indicate your understanding of • Write short notes on