THE UNITED REPUBLIC OF TANZANIA

NATIONAL EXAMINATION COUNCIL OF TANZANIA

DIPLOMA IN SECONDARY EDUCATION EXAMINATION

762 EDUCATIONAL RESEARCH, MEASUREMENT AND EVALUATION

Time: 3 Hours. **ANSWER** **Year: 2017 a.m.**

Instructions

1. This paper consists of sections A and B with a total of **sixteen (16)** questions
2. Answer **all** questions in sections A and **four (4)** questions from section B.
3. Question **11** is **compulsory**.
4. Section A carries **forty (40)** marks and section B carries **sixty (60)** marks.
5. Cellular phones and unauthorized materials are **not allowed** in the examination room.
6. Non-Programmable calculators may be used.
7. Write your **Examination Number** on every page of your answer booklet(s).

SECTION A (40 Marks)
Answer **all** questions in this section

1. With suitable examples, distinguish between criterion-referenced measurement and norm-referenced measurement.

Criterion-referenced measurement evaluates a learner's performance against a specific set of objectives or standards. For example, a driving test requires a learner to demonstrate specific skills like parallel parking regardless of how others perform.

Norm-referenced measurement compares a learner's performance to that of a larger group, ranking students relative to one another. For example, a standardized test that places a student in the 70th percentile means they performed better than 70% of test takers.

2. Explain the following concepts:
a) Reliability
b) Validity

Reliability refers to the consistency of an assessment tool in producing similar results under similar conditions. A reliable test yields stable outcomes even when administered multiple times.

Validity is the extent to which an assessment measures what it is intended to measure. For example, a math test that accurately measures problem-solving ability rather than reading skills demonstrates validity.

3. (a) Define formative assessment
(b) Give three advantages of using formative assessment in teaching and learning.

Formative assessment is a continuous process of gathering feedback on student learning during instruction to guide teaching adjustments.

It helps identify learning gaps early, allowing teachers to address them before final evaluations.
It encourages active student engagement by involving them in their learning progress.
It supports differentiated instruction as teachers can adapt methods to meet individual student needs.

4. (a) State two types of observation used in educational research
(b) Explain three advantages of using observation in classroom studies.

Two types of observation are structured observation, where specific behaviors are recorded using a checklist, and unstructured observation, which is more open-ended and flexible.

Observation provides direct evidence of student behavior and engagement, offering richer insights than self-reports.
It allows researchers to capture learning processes as they occur naturally in the classroom.
It reduces reliance on memory since events are recorded as they happen.

5. (a) Briefly describe two measures of central tendency
(b) Give two reasons for using measures of central tendency in education.

The mean is the arithmetic average of all scores, providing a general indication of overall performance. The median is the middle value when scores are arranged in order, useful when data contains extreme values.

Measures of central tendency summarize large sets of data into a single representative value. They provide a basis for comparing different groups or classes.

6. Identify four purposes of administering standardized tests in schools.

They provide a common measure for comparing performance across different schools or regions.
They identify students' strengths and weaknesses to inform instruction.
They are used for selection into higher education or special programs.
They monitor the effectiveness of educational policies and curriculum.

7. Briefly explain four disadvantages of using oral examinations in assessing students' achievement.

They may be influenced by examiner bias, leading to unfair results.
Students may experience anxiety that affects their performance.
They are time-consuming to administer for large groups.
They may not comprehensively assess the full range of knowledge due to time limits.

8. Briefly explain four benefits of using performance-based assessment in education.

It assesses practical application of skills rather than rote memorization.
It promotes higher-order thinking as students solve real-world problems.
It provides opportunities for students to demonstrate creativity and innovation.
It offers more authentic evaluation aligned with workplace skills.

9. State and explain four principles of test construction.

Clarity: Test items must be clear and free from ambiguity.
Relevance: Items should directly relate to learning objectives.
Balance: Questions should cover a range of topics and cognitive levels.
Fairness: The test should not disadvantage any group of students.

10. Distinguish between:
(a) Reliability and validity
(b) Formative assessment and summative assessment

Reliability refers to consistency of results, while validity focuses on accuracy in measuring intended outcomes.

Formative assessment is conducted during learning to improve performance, while summative assessment evaluates learning at the end of an instructional period for grading purposes.

SECTION B (30 Marks)

Answer **two (2)** questions from this section

11. (a) Calculate the Pearson's correlation coefficient for the scores.

Math: 78, 65, 82, 55, 90, 70, 60, 85
Physics: 75, 60, 85, 58, 88, 72, 62, 80

Step 1: Calculate the mean for Math and Physics.

Math mean = (78 + 65 + 82 + 55 + 90 + 70 + 60 + 85) / 8 = 580 / 8 = 72.5
Physics mean = (75 + 60 + 85 + 58 + 88 + 72 + 62 + 80) / 8 = 580 / 8 = 72.5

Step 2: Calculate deviations, products, and squares.

X	Y	$X-\bar{X}$	$Y-\bar{Y}$	$(X-\bar{X})(Y-\bar{Y})$	$(X-\bar{X})^2$	$(Y-\bar{Y})^2$
78	75	5.5	2.5	13.75	30.25	6.25
65	60	-7.5	-12.5	93.75	56.25	156.25
82	85	9.5	12.5	118.75	90.25	156.25
55	58	-17.5	-14.5	253.75	306.25	210.25
90	88	17.5	15.5	271.25	306.25	240.25
70	72	-2.5	-0.5	1.25	6.25	0.25
60	62	-12.5	-10.5	131.25	156.25	110.25
85	80	12.5	7.5	93.75	156.25	56.25

$\Sigma(X-\bar{X})(Y-\bar{Y}) = 977.5$
$\Sigma(X-\bar{X})^2 = 1108$
$\Sigma(Y-\bar{Y})^2 = 936$

$r = \Sigma(X-\bar{X})(Y-\bar{Y}) / \sqrt{[\Sigma(X-\bar{X})^2 \times \Sigma(Y-\bar{Y})^2]}$
$r = 977.5 / \sqrt{(1108 \times 936)}$
$r = 977.5 / \sqrt{1036608}$
$r \approx 977.5 / 1018.12$
$r \approx 0.96$

Correlation coefficient = 0.96, showing a very strong positive relationship.

(b) Comment on the relationship

The correlation shows that students who perform well in Mathematics also tend to perform well in Physics. This suggests the subjects may require similar cognitive abilities.

12. Discuss five factors that can reduce the reliability of a test.

One factor is ambiguous wording of test items. When questions are unclear or confusing, students may interpret them differently, leading to inconsistent answers and reduced reliability.

Another factor is poor item construction. If test items do not match the objectives or are misleading, they may fail to consistently measure the intended skills.

The lack of a clear and detailed marking scheme can also lower reliability. Without uniform marking criteria, different examiners may assign different scores to similar answers.

External environmental conditions during testing, such as excessive noise, poor lighting, or uncomfortable seating, can cause variations in student performance that are unrelated to their true ability.

Finally, having too few test items may lower reliability. Short tests do not sample the learning objectives adequately and may not represent the full range of student ability.

13. Analyze five challenges that face the process of educational evaluation in Tanzania.

One challenge is limited resources. Many schools face shortages of funds, trained staff, and materials needed for effective evaluation, which can compromise the quality of the process.

Large class sizes make it difficult for teachers to assess individual students thoroughly. This can lead to superficial evaluation and less accurate judgments about student progress.

Overemphasis on examinations in the education system tends to reduce the use of continuous assessment. As a result, only end-of-term or final scores are considered, ignoring ongoing learning progress.

Cultural and language differences can affect the fairness of evaluation. Students who are not fluent in the language of the test may underperform despite understanding the content.

Political and administrative interference can also be a challenge, particularly when examination results are influenced by non-academic considerations, undermining the credibility of the process.

14. Explain five measures that can be taken to improve the validity of national examinations.

Ensuring alignment of test content with the curriculum objectives is essential so that the examination measures what students have been taught.

Eliminating bias in questions helps make examinations fair for all candidates, regardless of gender, culture, or socioeconomic background.

Including different question types, such as multiple choice, essays, and practical tasks, helps measure a variety of skills and reduces over-reliance on one format.

Conducting pilot testing of items allows examiners to identify unclear or inappropriate questions before the actual administration, improving clarity and relevance.

Training examiners and markers ensures that assessment procedures are applied consistently, reducing errors and increasing fairness.

15. Evaluate five contributions of educational measurement to curriculum improvement.

Educational measurement identifies weaknesses in the curriculum by highlighting areas where student performance is consistently low, prompting curriculum revision.

It provides data on the effectiveness of teaching strategies, enabling educators to adopt methods that improve learning outcomes.

It helps monitor whether curriculum objectives are being met across different schools and regions, ensuring consistency in quality.

It informs policymakers on where to allocate resources, such as funding and teacher training, based on evidence from student performance.

It enhances accountability for teachers and schools, encouraging them to maintain high instructional standards.

16. Identify and explain five characteristics of a good scoring rubric.

A good scoring rubric has clear criteria that describe exactly what is being assessed, avoiding vague descriptions.

It provides well-defined performance levels that clearly distinguish between excellent, good, fair, and poor work.

It is comprehensive, covering all relevant aspects of the task so that students are assessed holistically.

It is user-friendly, meaning it is easy for both teachers and students to understand and apply.

It is reliable, producing consistent scoring results when used by different evaluators, which increases fairness and objectivity.