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: 2019 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).

1. Give four characteristics of restricted response type of test items.

Restricted response items limit the scope of the answer expected from the learner. They are designed to assess specific content or objectives and often require short or structured answers.

First, they allow little freedom in the type of response. Students are expected to answer within a narrow range, usually based on a direct question or prompt that limits discussion or elaboration.

Second, they are suitable for evaluating lower-level cognitive skills. These include recall, comprehension, and application of specific knowledge rather than analysis or synthesis.

Third, they are easy to mark objectively. Since answers are more predictable and shorter, scoring can be done with high consistency and minimal subjectivity.

Fourth, they help in testing factual knowledge. They are commonly used when teachers want to assess precise facts, definitions, dates, or formulas from students.

2. Give four characteristics of a good instrument of measurement.

A good measurement instrument ensures accurate and meaningful assessment outcomes. It must possess several core characteristics to be effective.

First, it should be valid. This means it measures what it is intended to measure, ensuring the test aligns with the learning objectives or construct being evaluated.

Second, it should be reliable. A reliable instrument gives consistent results when administered under similar conditions over time or across different evaluators.

Third, it must be practical. The instrument should be easy to administer, cost-effective, and time-efficient for both the examiner and the examinee.

Fourth, it must be usable. This involves being understandable and appropriate for the target group in terms of language, difficulty level, and context.

3. Outline four sources of literature review in research.

Literature review sources provide a foundation for research by showing existing knowledge and identifying gaps.

One source is academic journals. These provide peer-reviewed research articles with up-to-date findings and are considered highly credible.

Another is textbooks. These offer broad overviews of theories and concepts in the researcher's field and are especially useful for building background knowledge.

A third source is conference proceedings. These include research presented at academic conferences and often reflect current trends or emerging issues in a field.

Finally, online databases and digital libraries, such as Google Scholar or JSTOR, offer access to a wide range of scholarly materials, including theses, reports, and articles.

4. Briefly describe four qualities of a good research problem.

A good research problem is the foundation of any effective study and must meet certain standards.

First, it should be clear and specific. The problem must be stated in a focused way, avoiding ambiguity so that readers and researchers understand the exact issue.

Second, it should be researchable. This means the problem can be investigated using empirical methods and available resources such as time, money, and tools.

Third, it must be significant. The problem should address a real need or gap in knowledge that, when resolved, contributes meaningfully to the field.

Fourth, it should be ethical and feasible. The research problem should not involve harm to participants and should be practical to carry out under existing conditions.

5. Briefly describe four scales of measurement in education.

Scales of measurement define how data is categorized and interpreted in educational assessment.

The nominal scale categorizes data without any quantitative value. Examples include gender, nationality, or types of schools.

The ordinal scale arranges data in a ranked order. For example, student positions in class (first, second, third) show order but not equal intervals between ranks.

The interval scale not only ranks data but also provides equal intervals between values. An example is temperature in Celsius, though it lacks a true zero point.

The ratio scale includes all features of interval scale and has a true zero. Examples include marks obtained in a test or time spent studying.

6. Briefly explain four differences between assessment and evaluation.

Assessment and evaluation are often used interchangeably, but they have different focuses.

Assessment is a process of gathering information about student learning, mainly for improvement. Evaluation, however, involves making judgments about the value or effectiveness of a learning activity or program.

Assessment is formative in nature, focusing on progress and helping learners improve. Evaluation is more summative, often used to make final decisions such as grading or program approval.

Assessment is continuous and conducted during the learning process. Evaluation is typically done at the end of an instructional period or program.

Assessment emphasizes feedback and diagnosis of learning difficulties. Evaluation focuses on accountability and decision-making based on performance.

7. Write four principles guiding evaluation of learners' achievement.

Evaluating learners effectively requires adherence to several guiding principles.

First, the principle of validity ensures that the evaluation measures what it is intended to measure and aligns with instructional goals.

Second, the principle of reliability guarantees consistency in evaluation results regardless of when or who conducts the assessment.

Third, the principle of fairness demands that all learners are given equal opportunity to succeed and that evaluations are free from bias or discrimination.

Fourth, the principle of usefulness implies that evaluation results must be relevant and actionable, providing feedback that can inform teaching and learning decisions.

8. Briefly explain four physical environmental factors that can affect individual's performance in a test.

The testing environment plays a major role in determining student performance.

Lighting is a key factor. Poor lighting can cause eye strain and reduce concentration, especially during long exams.

Noise is another influence. High levels of background noise, such as construction sounds or talking, can distract students and lower their ability to focus.

Temperature also matters. Extreme heat or cold in an exam room can cause discomfort and stress, affecting a student's ability to think clearly.

Seating arrangements are crucial. Uncomfortable chairs or inadequate spacing can lead to physical strain and disturb students' concentration and comfort during the test.

9. Outline four steps involved in preparation of a Table of Specifications.

A Table of Specifications (TOS) is a tool used to ensure balanced test construction.

First, identify the learning objectives. This involves reviewing the curriculum to determine the knowledge and skills to be assessed.

Second, determine the content areas. These are the main topics or units from which the test items will be drawn.

Third, decide on the cognitive levels. This means categorizing items based on Bloom's taxonomy, such as knowledge, comprehension, application, or analysis.

Fourth, allocate the number of items for each category. This ensures proportional representation of content and cognitive levels, guiding fair and balanced item distribution.

10. (a) Define the term reliability as applied to tests and examinations.

Reliability in testing refers to the consistency or stability of test scores across repeated administrations or different evaluators. A reliable test produces similar results under consistent conditions and minimizes errors or fluctuations due to chance or external factors.

(b) Identify two purposes of validity in testing and assessment.

Validity ensures that a test measures what it is intended to measure. One purpose of validity is to ensure accurate interpretation of scores for decision-making, such as promotion or placement. Another purpose is to maintain fairness by confirming that test items reflect the content and objectives of instruction, avoiding bias or irrelevance.

11. The following scores were obtained in a test involving 12 students. 90, 70, 80, 90, 70, 94, 80, 70, 80, 90, 96, 80. From the scores, find the following:

(a) **Mode**
The mode is the score that occurs most frequently.
In the given data:
70 occurs 3 times, 80 occurs 4 times, 90 occurs 3 times, 94 occurs once, and 96 occurs once.
Therefore, the mode is 80 because it appears most frequently.

(b) **Median**
The median is the middle value when data is arranged in ascending order.
Arranging the scores: 70, 70, 70, 80, 80, 80, 80, 90, 90, 90, 94, 96.

There are 12 scores, so the median is the average of the 6th and 7th scores.
6th score = 80, 7th score = 80.
Median = (80 + 80) ÷ 2 = 80.

(c) **Range**
The range is the difference between the highest and lowest values.
Highest score = 96, lowest score = 70.
Range = 96 − 70 = 26.

(d) **Mean**
The mean is the sum of all scores divided by the number of scores.
Sum of scores = 90 + 70 + 80 + 90 + 70 + 94 + 80 + 70 + 80 + 90 + 96 + 80 = 930.
Number of scores = 12.
Mean = 930 ÷ 12 = 77.5.

(e) **Standard Deviation**
Step 1: Find the mean = 77.5.
Step 2: Find each deviation from the mean, square it, then sum:
$(90 − 77.5)^2 = 156.25$
$(70 − 77.5)^2 = 56.25$
$(80 − 77.5)^2 = 6.25$
$(90 − 77.5)^2 = 156.25$
$(70 − 77.5)^2 = 56.25$
$(94 − 77.5)^2 = 272.25$
$(80 − 77.5)^2 = 6.25$
$(70 − 77.5)^2 = 56.25$
$(80 − 77.5)^2 = 6.25$
$(90 − 77.5)^2 = 156.25$
$(96 − 77.5)^2 = 342.25$
$(80 − 77.5)^2 = 6.25$
Sum of squared deviations = 1272.5.
Step 3: Variance = Sum of squared deviations ÷ number of scores
Variance = 1272.5 ÷ 12 = 106.04.
Step 4: Standard deviation = √Variance = √106.04 ≈ 10.3.

12. Account for five reasons of doing literature review in research.

One reason is to identify gaps in existing knowledge. This helps the researcher determine areas where new studies are needed.

Another reason is to avoid duplication of work. Reviewing past studies prevents repetition and saves time and resources.

A third reason is to refine the research problem. Literature provides background information that can shape and narrow the focus of the study.

Fourth, it helps to select appropriate research methods. By studying previous works, a researcher can choose designs and techniques that have been effective.

Fifth, it provides a theoretical foundation. Literature review establishes the conceptual framework that supports the research and links it to established theories.

13. Measurement is very crucial for educational administrators. Justify this statement by giving five points.

Measurement helps in assessing student achievement. It provides objective data on performance, which is essential for making academic decisions.

It aids in curriculum evaluation. Administrators can determine whether teaching content meets desired learning outcomes.

It supports placement decisions. Measurement results guide the assignment of learners to appropriate levels or classes based on ability.

It informs resource allocation. Data from measurement can show areas needing more support, materials, or teacher training.

It helps monitor educational progress over time. Administrators can compare results across years to evaluate improvement or decline.

14. Elaborate five strategies to consider in ensuring test validity.

One strategy is aligning test items with learning objectives. Every question should reflect the skills or knowledge intended to be assessed.

Another is using clear and unambiguous language. This ensures that students understand what is being asked without confusion.

A third is avoiding bias in questions. Items should be fair to all learners regardless of gender, culture, or background.

Fourth, include a representative sample of content. This ensures the test covers the breadth of material taught and not just a narrow portion.

Fifth, pilot testing the instrument before full use. This helps detect unclear items or issues affecting the accuracy of results.

15. Provide five technical advices on how to construct supply type of test items for effective education assessment.

One advice is to clearly define the expected length and format of answers. This prevents unnecessary variation in responses.

Another is to use precise wording in questions. This guides learners directly to the intended response without misinterpretation.

A third is to ensure the item requires the intended cognitive skill, whether recall, explanation, or problem-solving.

Fourth, avoid giving clues in the question that might lead to the answer. This maintains the integrity of the assessment.

Fifth, score responses using a marking scheme or rubric. This improves reliability by ensuring uniform evaluation of answers.

16. In five points, analyse the usefulness of measurement and evaluation in education.

Measurement and evaluation help determine the effectiveness of teaching methods. This guides teachers in improving instructional approaches.

They provide feedback to learners. Students can understand their strengths and weaknesses, allowing targeted improvement.

They assist in decision-making. Administrators and teachers use results to plan interventions, promotions, or remedial programs.

They contribute to accountability. Measurement and evaluation hold teachers, schools, and education systems responsible for learner outcomes.

They support research and policy-making. Data gathered helps in designing educational policies and programs that address real needs.