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: 2018 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. With examples, the two types of Educational Measurement based on test interpretation.

The first type is norm-referenced measurement. This involves interpreting test scores by comparing an individual's performance to that of a larger group, known as the norm group. For example, if a student scores in the 80th percentile on a standardized test, it means they performed better than 80% of the group.

The second type is criterion-referenced measurement. This measures an individual's performance against a fixed set of predetermined criteria or learning standards. For example, a student may need to score at least 75% on a math test to be considered proficient in the topic, regardless of how other students perform.

2. Explain the following Concepts:
(a) Measurement
(b) Educational measurement.

Measurement refers to the process of assigning numerical or descriptive values to an individual's characteristics, behaviors, or attributes using specific tools and procedures. It is used to quantify variables like height, weight, intelligence, or performance.

Educational measurement is the application of measurement techniques specifically within the education context. It involves collecting data on students' academic achievement, attitudes, and other learning outcomes to support instructional decisions and evaluate learning.

3. (a) Define action research
(b) Give three areas in which action research can be used in education.

Action research is a reflective process conducted by educators to improve teaching and learning practices through systematic inquiry and intervention in their own environment.

One area it can be used is classroom instruction, where teachers identify challenges in student understanding and try new strategies to improve learning.

Another area is school management, where administrators assess staff collaboration or discipline policies and implement improvements.

It can also be used in curriculum development, where educators evaluate the relevance of content and modify teaching materials to fit student needs better.

4. (a) Give two types of Interview in Educational Research
(b) Describe three limitations of using interview in educational research.

The two main types of interviews are structured interviews, where the researcher uses pre-set questions, and unstructured interviews, where questions are more open-ended and flexible.

One limitation of interviews is that they can be time-consuming to conduct and analyze, especially with large groups.

Another limitation is the risk of interviewer bias, where the interviewer's tone or wording may influence the respondent's answers.

Third, respondents may give socially desirable answers rather than honest ones, especially on sensitive issues, reducing data validity.

5. (a) Briefly describe two measures of variability
(b) List down two techniques of standardizing test scores.

One measure of variability is range, which is the difference between the highest and lowest scores in a data set. It gives a simple sense of score spread.

Another is standard deviation, which shows how much individual scores deviate from the mean, providing a more accurate picture of variability.

Two techniques of standardizing test scores include z-scores, which convert raw scores into standard deviations from the mean, and T-scores, which rescale data to have a fixed mean (usually 50) and standard deviation (typically 10).

6. Identify the four functions of objective test.

Objective tests help measure learners' knowledge accurately and consistently, free from examiner bias.

They help in assessing a broad range of content quickly, since many items can be included in a single test.

They support easy and reliable scoring, especially when multiple-choice or true/false items are used.

They provide data for comparing students or groups, which is useful for evaluation and decision-making.

7. Briefly explain four effective uses of anecdotal records in assessing student achievements.

Anecdotal records help track students' behavioral and emotional development over time by recording specific observed events.

They can be used to identify patterns or recurring issues in student performance or behavior, helping inform instructional strategies.

They serve as evidence during parent-teacher conferences, providing detailed insights into a child's classroom experience.

They support individualized learning plans by highlighting unique strengths and challenges of each learner.

Find this and other free resources at: https://maktaba.tetea.org

Prepared by Maria Marco for TETEA

8. Briefly explain four advantages of using table of specifications in test construction.

A table of specifications ensures that the test content is balanced and reflects the key objectives and topics taught in class.

It improves test validity by linking items directly to instructional goals and cognitive levels.

It helps teachers distribute items appropriately across different topics and skills, avoiding overemphasis on certain areas.

It guides item writers and ensures consistency and fairness in test design.

9. Provide a brief explanation on the four qualities of a good test.

A good test is valid, meaning it measures what it is intended to measure accurately.

It is reliable, providing consistent results over time or across different assessors.

It is practical, being easy to administer, score, and interpret with available resources.

It is objective, minimizing personal bias in both the test items and the scoring process.

10. Distinguish between:
(a) Range and standard deviation

Range is a simple measure of spread, calculated by subtracting the lowest score from the highest score. It provides only the extremes of data distribution.

Standard deviation, however, shows how scores vary around the mean, considering every data point. It provides a more detailed measure of variability.

(b) Variance and standard deviation

Variance is the average of the squared deviations from the mean and gives a general sense of how much scores differ.

Standard deviation is the square root of variance and expresses variability in the same units as the original data, making it easier to interpret.

11. (a) Calculate the correlation coefficient between students' performance of the two tests (A and B) by using Spearman's Rank Order Method

Step 1: Rank the scores in both tests. If two or more scores are tied, assign the average rank for those positions.

Test A ranking:

Scores: 2, 4, 10, 9, 1, 3, 6, 5, 7, 8

Ranks: 9th highest gets rank 2, etc.

10 → Rank 1

9 → Rank 2

8 → Rank 3

7 → Rank 4

6 → Rank 5

5 → Rank 6

4 → Rank 7

3 → Rank 8

2 → Rank 9

1 → Rank 10

So:

Simbomboye (2) → 9

Mugo (4) → 7

Kinani (10) → 1

Magu (9) → 2

Muturubhai (1) → 10

Magoke (3) → 8

Kinye (6) → 5

Dude (5) → 6

Mwike (7) → 4

Kime (8) → 3

Test B ranking:

Scores: 3, 5, 10, 9, 2, 1, 4, 7, 6, 8

10 → Rank 1

9 → Rank 2

8 → Rank 3

7 → Rank 4

6 → Rank 5

5 → Rank 6

4 → Rank 7

3 → Rank 8

2 → Rank 9

1 → Rank 10

So:

Simbomboye (3) → 8

Mugo (5) → 6

Kinani (10) → 1

Find this and other free resources at: https://maktaba.tetea.org

Prepared by Maria Marco for TETEA

Magu (9) → 2
Muturubhai (2) → 9
Magoke (1) → 10
Kinye (4) → 7
Dude (7) → 4
Mwike (6) → 5
Kime (8) → 3
Step 2: Find difference in ranks (d) and square (d²):
Simbomboye: 9 − 8 = 1 → 1² = 1
Mugo: 7 − 6 = 1 → 1² = 1
Kinani: 1 − 1 = 0 → 0² = 0
Magu: 2 − 2 = 0 → 0² = 0
Muturubhai: 10 − 9 = 1 → 1² = 1
Magoke: 8 − 10 = −2 → 4
Kinye: 5 − 7 = −2 → 4
Dude: 6 − 4 = 2 → 4
Mwike: 4 − 5 = −1 → 1
Kime: 3 − 3 = 0 → 0
Σd² = 16
Step 3: Apply Spearman's formula:
$\rho = 1 - [(6 \times \Sigma d^2) / (n(n^2 - 1))]$
n = 10
$\rho = 1 - [(6 \times 16) / (10 \times (100 - 1))]$
$\rho = 1 - [96 / 990]$
$\rho = 1 - 0.09697$
$\rho \approx 0.903$

Correlation coefficient = 0.903, indicating a very strong positive relationship.

(b) Give reasons on whether the two subjects indicate the same students' ability or not

The high correlation coefficient of 0.903 suggests that performance in one test is closely related to performance in the other. This means students who performed well in Test A also tended to perform well in Test B. Therefore, the two subjects likely measure similar abilities or cognitive skills, although they may not be identical in content.

12. Analyze five factors affecting subjectivity of a test.

One factor is ambiguous questions. If test items are unclear, different students and markers may interpret them differently, increasing subjectivity.

Another factor is the use of essay-type items. These allow for varied student expression and require examiner judgment, which may vary between markers.

Marker bias can also affect subjectivity. Personal beliefs, expectations, or moods of the examiner may unconsciously influence scoring.

Lack of a marking scheme can lead to inconsistent scoring. Without clear criteria, different markers may assign different scores to similar answers.

Cultural differences between the test designer and test-takers may influence interpretation of questions and answers, leading to subjective judgments.

13. Explain five objectives of examinations in Tanzania.

One objective is to measure students' mastery of the curriculum, ensuring they have acquired the intended knowledge and skills.

Another is to provide feedback to teachers and learners about the effectiveness of teaching and learning processes.

Examinations are also used for selection purposes, such as determining eligibility for further education or employment.

They help maintain educational standards by setting benchmarks that students are expected to meet.

They serve as a basis for certification, confirming that learners have met the required competencies for their level.

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

Align the test with learning objectives so that it measures the intended knowledge and skills.

Ensure content coverage by including items from all important topics taught during the course.

Use appropriate item formats that suit the skills being tested, such as practical tasks for performance skills.

Avoid bias by ensuring items are fair to all students regardless of gender, culture, or background.

Pilot test the instrument to identify and correct unclear items before using it in the actual assessment.

15. Evaluate five roles of item analysis in curriculum development.

Item analysis helps identify strengths and weaknesses in test items, guiding improvements in future assessments.

It provides feedback on how well the curriculum objectives are being achieved through student performance.

It helps determine whether test items discriminate effectively between high- and low-performing students.

It informs revisions in teaching methods and learning materials by highlighting areas where students struggle.

It ensures that assessments remain fair, reliable, and valid, supporting overall curriculum quality.

16. Identify five characteristics of norm-referenced measurement.

Norm-referenced measurement compares an individual's performance to a group standard rather than fixed criteria.

Find this and other free resources at: https://maktaba.tetea.org

Prepared by Maria Marco for TETEA

Scores are usually expressed in percentiles, ranks, or standard scores to show relative standing.

It is useful for selection decisions, such as admission to programs or awarding scholarships.

The focus is on ranking students rather than determining mastery of specific skills.

The distribution of scores typically follows a normal curve, with most students clustering around the average.