THE UNITED REPUBLIC OF TANZANIA

NATIONAL EXAMINATIONS COUNCIL OF TANZANIA

DIPLOMA IN SECONDARY EDUCATION EXAMINATION

762 EDUCATIONAL RESEARCH, MEASUREMENT AND EVALUATION

Time: 3 Hours **ANSWERS** **Year: 2015**

Instructions

1. This paper consists of section A and B.

2. Answer all questions in section A, and four questions from section B.

1. Write down four characteristics of a good measuring instrument.

A good measuring instrument should have validity, meaning it accurately measures what it is intended to measure.

A good measuring instrument should have reliability, ensuring that repeated measurements under the same conditions produce consistent results.

A good measuring instrument should be sensitive, meaning it can detect small changes in the measured attribute.

A good measuring instrument should be practical, meaning it is easy to use, cost-effective, and applicable in real-world conditions.

2. (a) What is a marking scheme?

A marking scheme is a guideline used to allocate marks to student responses in an examination or assessment. It outlines the correct answers, acceptable variations, and the corresponding marks for each part of a question to ensure fairness and consistency in grading.

2. (b) Briefly explain the importance of preparing a marking scheme.

A marking scheme ensures consistency in grading by providing clear guidelines for awarding marks, reducing subjectivity in assessment.

A marking scheme improves efficiency by allowing examiners to grade responses quickly and systematically.

A marking scheme enhances fairness by ensuring that all students are evaluated using the same criteria, minimizing bias in scoring.

A marking scheme helps students understand expectations by outlining how marks are distributed across different parts of a question.

3. State the difference between evaluation and testing in the teaching/learning process.

Evaluation is a broader process that assesses the overall effectiveness of teaching and learning, including student performance, instructional methods, and curriculum effectiveness. Testing, on the other hand, is a specific tool within evaluation that measures student achievement through exams, quizzes, and assessments. Evaluation considers multiple factors, such as classroom participation and projects, while testing focuses on specific knowledge and skills.

4. (a) Analyse five techniques that can be followed in order to construct the best true-false items.

Ensure that statements are clear and unambiguous to prevent confusion and misinterpretation by students.

Avoid using absolute terms such as "always" and "never," as they often make the statement too easy to identify as false.

Ensure that the statement is either completely true or completely false, avoiding partially correct statements that may mislead students.

Use a balanced mix of true and false statements to prevent students from guessing based on patterns rather than knowledge.

Keep statements concise and free of unnecessary complexity to ensure students focus on the core concept being tested.

4. (b) Write down three advantages of planning a test.

Planning a test ensures alignment with learning objectives, ensuring that assessment measures the intended skills and knowledge.

Planning a test allows for balanced coverage of topics, preventing overemphasis on certain areas while neglecting others.

Planning a test improves fairness by structuring questions in a logical manner, reducing ambiguity and promoting equal opportunities for all students.

5. (a) Account for four factors that must be considered before constructing a good test.

The purpose of the test must be clearly defined, whether it is for formative assessment, summative assessment, or diagnostic purposes.

The level of difficulty should be appropriate for the students' knowledge and abilities to ensure the test is neither too easy nor too hard.

The content should be relevant to the curriculum to ensure that the test assesses what has been taught in the course.

The test format should be carefully selected, including a mix of question types such as multiple-choice, short answers, and essays, depending on the assessment goals.

5. (b) Briefly explain four qualities of a good test.

A good test is valid, meaning it accurately measures the intended knowledge, skills, or abilities.

A good test is reliable, ensuring that results are consistent when administered to different groups or at different times.

A good test is fair, providing equal opportunities for all students regardless of background or ability.

A good test is practical, meaning it is easy to administer, grade, and interpret within the available time and resources.

6. (a) Suppose 0.6 is the correlation between two halves of a test. What will be the correlation coefficient for a complete test?

Using the Spearman-Brown formula, the correlation coefficient for the complete test can be calculated as follows:

r_full = (2 x r_half) / (1 + r_half)
r_full = (2 x 0.6) / (1 + 0.6)
r_full = 1.2 / 1.6
r_full = 0.75

6. (b) What interpretation can you make from the value of correlation coefficient obtained in (a) above?

A correlation coefficient of 0.75 indicates a moderately high level of reliability. This suggests that the test is fairly consistent in measuring student performance, but there may still be some variations in scores due to external factors such as test conditions or student effort.

7. (a) Mention four uses of standard deviation.

Standard deviation measures the spread of data, helping to determine how much individual scores deviate from the mean.

Standard deviation helps in comparing variability between different sets of data, allowing for meaningful analysis of performance differences.

Standard deviation is used in grading to identify students who are significantly above or below the average performance.

Standard deviation assists in quality control and decision-making processes in various fields, including education and business.

7. (b) List down four measures of central tendency.

Mean is the average of all data points, providing a central value for the dataset.

Median is the middle value in an ordered dataset, representing the central position without being affected by extreme values.

Mode is the most frequently occurring value in a dataset, indicating the most common score.

Midrange is the average of the highest and lowest values, providing a quick estimate of the dataset's center.

8. How does formative evaluation differ from summative evaluation?

Formative evaluation is conducted during the learning process to provide continuous feedback for improvement, while summative evaluation takes place at the end of an instructional period to measure overall achievement. Formative evaluation focuses on identifying weaknesses and adjusting teaching strategies, whereas summative evaluation determines final grades and assesses overall competency.

9. (a) Differentiate between predictive validity and concurrent validity.

Predictive validity refers to the extent to which test scores accurately forecast future performance, such as using entrance exams to predict academic success.

Concurrent validity measures how well a test correlates with an established measure taken at the same time, such as comparing new test results with existing standardized assessments.

9. (b) Give four methods that are used in estimating the reliability of a test.

Test-retest reliability involves administering the same test twice to the same group and measuring consistency in results.

Split-half reliability divides the test into two halves and compares scores from each half to assess internal consistency.

Parallel-form reliability involves giving two different but equivalent versions of a test to the same group and comparing scores.

Inter-rater reliability measures the level of agreement between different graders when scoring subjective responses.

10. (a) What is educational research?

Educational research is a systematic investigation aimed at improving educational practices, policies, and outcomes. It involves data collection, analysis, and interpretation to solve problems and enhance teaching and learning methods.

10. (b) Briefly explain three characteristics of educational research.

Educational research is systematic, meaning it follows a structured process with clear objectives and methodologies.

Educational research is empirical, relying on data collection and analysis rather than personal opinions or assumptions.

Educational research is objective, ensuring that findings are based on evidence and free from bias to produce reliable conclusions.

11. (a) Explain the meaning of correlation coefficient.

The correlation coefficient is a statistical measure that indicates the strength and direction of a linear relationship between two variables. It ranges from -1 to 1. A correlation coefficient close to 1 signifies a strong positive correlation, meaning that as one variable increases, the other also increases. A value close to -1 represents a strong negative correlation, indicating that as one variable increases, the other decreases. A value near 0 suggests no significant relationship between the two variables.

11. (b) Consider the following tests for students who were selected as a sample for item analysis, 10 from higher achievers and 10 from the low achievers, where A was the correct response.

To analyze the given data, we use the discrimination index formula, which measures how well a test item differentiates between high and low achievers. The difficulty index can also be calculated to determine how easy or hard the item was. The correct responses from both groups are compared, and their differences provide insight into whether the item was fair and effective in distinguishing student performance levels.

12. (a) Describe four types of assessment.

Diagnostic assessment is conducted before instruction begins to evaluate students' prior knowledge, strengths, and weaknesses. It helps teachers design lessons tailored to student needs.

Formative assessment occurs during the learning process and provides continuous feedback to students and teachers. It includes quizzes, assignments, and class participation to monitor progress.

Summative assessment takes place at the end of a learning period, such as final exams or standardized tests. It evaluates the overall achievement of learning objectives.

Ipsative assessment compares a student’s current performance with their previous performance rather than with others. This helps in tracking individual progress and self-improvement.

12. (b) Explain five uses of formative assessment.

Formative assessment helps identify learning gaps by revealing which concepts students struggle with, allowing for targeted interventions.

It provides immediate feedback to students, helping them understand their mistakes and improve their learning before final evaluations.

It allows teachers to adjust their teaching methods based on student performance, ensuring lessons are effective and engaging.

It encourages student engagement by promoting active participation in learning activities, discussions, and group work.

It supports differentiated instruction by allowing teachers to modify lessons to meet the diverse needs of learners with varying levels of understanding.

13. (a) Distinguish between dependent variable and independent variable.

A dependent variable is the outcome being measured in an experiment. It depends on changes in the independent variable. For example, in a study examining the effect of study hours on test scores, the test scores are the dependent variable.

An independent variable is the factor that is manipulated to observe its effect on the dependent variable. In the same study, study hours are the independent variable because they are intentionally varied to measure their impact on test scores.

13. (b) Analyse four characteristics of a good hypothesis and research.

A good hypothesis is testable, meaning it can be examined through experiments or observations to determine its validity.

A good hypothesis is specific and clearly defines the relationship between variables, avoiding vague statements.

A good hypothesis is based on existing knowledge and theories, ensuring it is grounded in research and logical reasoning.

A good hypothesis is measurable, allowing for data collection and analysis to support or refute it.

A good research study is objective and free from bias, ensuring that findings are based on evidence rather than personal opinions.

14. (a) Outline four factors to be considered before constructing a table of specification.

The learning objectives must be defined to ensure the test aligns with the intended outcomes of the curriculum.

The content coverage should be balanced to fairly represent all topics taught during the instructional period.

The cognitive levels of questions should be considered, ensuring a mix of lower and higher-order thinking skills based on Bloom's taxonomy.

The weight assigned to different topics should be proportional to their importance in the syllabus to maintain a fair assessment.

14. (b) Explain three main components of the table of specification.

The content area defines the topics and subtopics covered in the test to ensure a comprehensive assessment.

The cognitive level categorizes questions based on difficulty, ensuring a balance of knowledge, comprehension, application, and analysis.

The test item distribution specifies the number and type of questions allocated to each topic, ensuring fairness in assessment.

15. (a) Both subjective and objective test items are useful in measuring students' performance. Explain five limitations for each.

Subjective test items require human judgment in scoring, leading to inconsistency and potential bias in evaluation.

Subjective test items are time-consuming to grade because they require detailed analysis of student responses.

Subjective test items may encourage rote memorization rather than deep understanding, as students might focus on recalling expected answers.

Subjective test items lack standardization, making it difficult to compare performance across different students or institutions.

Subjective test items can be influenced by factors such as handwriting, grammar, and organization, which may not reflect actual knowledge.

Objective test items encourage guessing, as students can select an answer without fully understanding the concept.

Objective test items do not assess deep thinking or creativity, as they primarily measure recognition and recall rather than application.

Objective test items require careful construction to avoid ambiguity and misleading options, which can affect validity.

Objective test items may not be suitable for complex topics that require explanation or problem-solving skills.

Objective test items may lead to test anxiety, as students are limited in their ability to express knowledge beyond predefined choices.

15. (b) Explain five advantages of an essay test.

Essay tests encourage critical thinking by allowing students to analyze, evaluate, and synthesize information rather than just recall facts.

Essay tests provide flexibility in responses, allowing students to explain concepts in their own words and demonstrate depth of understanding.

Essay tests assess higher-order cognitive skills, such as application and problem-solving, making them more effective for evaluating complex learning.

Essay tests allow partial credit, enabling students to earn marks even if their answers are not completely correct, which is fairer than multiple-choice formats.

Essay tests reduce the chances of guessing, as students must construct their responses rather than choosing from pre-set options.

16. (a) Explain five advantages of an essay test.

Essay tests promote originality and creativity by allowing students to express unique ideas and perspectives on a topic.

Essay tests provide insight into a student's ability to organize thoughts logically and present coherent arguments.

Essay tests enable teachers to evaluate students' reasoning and justifications, which are not possible in objective tests.

Essay tests can cover a wide range of topics within a subject, allowing for comprehensive assessment.

Essay tests encourage students to develop strong communication skills by practicing written expression and argumentation.

16. (b) Describe five procedures for scoring essay-type questions.

Establish clear grading criteria by defining key points, structure, and expected content for each answer.

Use a rubric to ensure consistency and fairness in evaluating responses across different students.

Grade all responses to one question before moving to the next to maintain uniformity in scoring.

Provide constructive feedback by highlighting strengths and areas for improvement in student responses.

Double-check borderline cases to ensure that students receive fair marks, especially when responses are close to the expected answer.

17. (a) Outline five steps involved in the educational evaluation process.

Defining objectives ensures that the evaluation aligns with learning goals and instructional aims.

Selecting appropriate evaluation methods involves choosing tools like tests, observations, and surveys to gather data.

Collecting data provides measurable evidence of student learning and instructional effectiveness.

Analyzing data helps identify patterns, strengths, and weaknesses in student performance.

Making decisions based on findings allows for curriculum adjustments and policy improvements in education.

18. (a) Differentiate facility index from discrimination power index of test items.

The facility index measures the percentage of students who answered a test item correctly, indicating its difficulty level.

The discrimination power index evaluates how well an item differentiates between high-achieving and low-achieving students.

18. (b) Describe the importance of item analysis. Provide five points.

Item analysis helps improve test quality by identifying poorly performing questions.

Item analysis assists teachers in refining instruction by revealing common student misconceptions.

Item analysis enhances fairness by ensuring that questions are neither too easy nor too difficult.

Item analysis aids in developing reliable assessments by ensuring items measure intended knowledge accurately.

Item analysis provides insights into student learning trends, helping educators improve curriculum design.