THE UNITED REPUBLIC OF TANZANIA
NATIONAL EXAMINATIONS COUNCIL OF TANZANIA
DIPLOMA IN SECONDARY EDUCATTION EXAMINATION

762 EDUCATIONAL RESEARCH, MEASUREMENT AND EVALUATION

Time: 3 Hours **ANSWERS** **Year: 2024**

Instructions.

1. This paper consists of sections A and B with a total of **Fourteen (14)** questions.
2. Answer **all** questions from section A and **four (4)** questions from section B.
3. Section A carries **forty (40)** marks and section B Carries **sixty (60)** marks.
4. Cellular phones are **note** allowed in the examination room.
5. Write your **examination Number** on every page of your answer booklet(s).

SECTION A (40 Marks)

Answer **All** questions from this section. Each question carries **four (4)** marks.

1. Describe how the "Frequency" and "Single frequency distribution" methods are applied.

   The frequency method is applied by counting the number of times a particular score or response appears within a given set of data. In the teaching and learning process, it helps teachers to quickly determine how often a certain score is obtained by students, making it easier to identify patterns in performance or attendance records.

   The single frequency distribution method involves organizing individual scores into a simple table where each possible score is listed alongside the number of students who obtained that score. This makes it easy to see how the scores are distributed and assists teachers in identifying the most common performance levels among learners.

2. Describe the major four characteristics of good quality test items.

   A good quality test item must be clear and unambiguous. This means the question should be worded in a simple, understandable manner without confusing language, so that all students clearly know what is being asked.

   The test item should also be relevant to the stated objectives of the lesson or course. This ensures that what is being tested aligns with what was taught, making the assessment a true reflection of student understanding.

   Another characteristic is fairness, meaning the test item should not favor or disadvantage any group of students based on gender, culture, or background. It should offer equal opportunity for every learner to demonstrate their knowledge.

   Finally, a quality test item should be capable of discriminating between high and low achievers. This means good students should consistently perform better on well-constructed test items compared to weaker students.

3. Explain how to use test-retest method to check the reliability of data.

   The test-retest method involves administering the same test to the same group of students on two different occasions, separated by a specific time interval. The purpose is to check whether the scores remain consistent over time.

   After both tests are conducted, the scores from the two sessions are correlated. A high positive correlation indicates that the test is reliable because it produces stable results over time.

   If there is a large difference in scores between the two administrations without any obvious reason, it suggests that the test may not be reliable. This technique is especially useful in assessing the consistency of achievement tests and attitude surveys.

4. Give four points to agree on the usefulness of the knowledge of educational measurement in the teaching and learning process.

   Educational measurement allows teachers to accurately assess the effectiveness of their teaching methods by evaluating students' performance in a systematic way. This helps in identifying areas where adjustments are needed for improved learning outcomes.

It helps in identifying individual learner differences. By analyzing students' scores, teachers can recognize which students need remedial support and which ones are progressing well, ensuring personalized attention where necessary.

The knowledge of educational measurement also assists in setting realistic and fair grading systems. It guides teachers in deciding appropriate pass marks and classifying students based on merit.

Lastly, it provides reliable data for decision-making on academic policies, promotions, or curriculum adjustments. Reliable measurements help school administrators and educators make informed, objective decisions regarding student progress.

5. Justify the statement by giving four points on the importance of cyclic action research in the teaching and learning process.

Cyclic action research helps teachers identify problems within the classroom environment and develop practical solutions through a systematic, continuous process of planning, action, observation, and reflection.

It allows teachers to implement changes immediately and observe their effects in real-time. This makes it a flexible, responsive approach to solving instructional problems or improving teaching strategies.

Another importance is that it enhances teacher professional development. As teachers engage in research cycles, they improve their skills, deepen their understanding of students' needs, and adopt new teaching approaches based on evidence.

Finally, cyclic action research fosters collaboration among teachers and school management. It encourages teamwork in addressing common educational challenges, leading to overall school improvement.

6. Explain briefly four types of validity in education.

Content validity refers to the extent to which a test represents the entire subject matter it is supposed to cover. A test with high content validity assesses all topics taught rather than focusing on a narrow section.

Construct validity deals with how well a test measures the theoretical concept it claims to assess. For example, an intelligence test should accurately measure intelligence-related abilities, not other unrelated skills.

Criterion-related validity examines how well a test's results align with other existing measures of the same ability. This can be concurrent validity, comparing current scores with another established test, or predictive validity, forecasting future performance.

Face validity involves the appearance of a test to measure what it is supposed to, based on the judgment of teachers, students, or education experts. While not scientific, it affects the acceptance and credibility of the test.

7. Explain briefly the four major characteristics of research.

Research is systematic, meaning it follows a well-organized sequence of steps from problem identification, data collection, analysis, and conclusion, ensuring orderly and structured investigation.

It is objective because findings are based on observable, measurable facts and evidence rather than personal opinions or bias, making conclusions reliable and credible.

Research is empirical, relying on real-world data gathered through observation, experiments, or surveys. Conclusions are drawn from practical evidence rather than speculation.

Lastly, research is replicable, meaning its procedures and results can be repeated by other researchers under similar conditions, producing consistent outcomes, which strengthens the validity of its findings.

8. Classify four techniques for students' assessment.

Written tests involve students responding to questions on paper or online. They assess a range of skills from knowledge recall to problem-solving and essay writing.

Practical assessments evaluate learners through hands-on activities, experiments, or demonstrations. This method is useful in science, vocational subjects, and sports.

Oral assessments include presentations, interviews, and recitations. They test communication skills, comprehension, and the ability to express ideas verbally.

Observation techniques involve teachers watching students as they perform tasks, noting behaviors, participation, or skill application. It is commonly used for assessing attitudes and practical abilities.

9. Explain four uses of the grouped frequency distribution method in processing students' scores.

It simplifies large sets of student scores into organized groups, making it easier to interpret patterns of performance across different ranges.

This method helps in identifying the most common performance category or class interval where the majority of students fall, which aids in understanding overall class performance.

It facilitates the calculation of statistical values like the mean, median, and mode by grouping similar scores, making computations quicker and clearer.

The method also assists in visualizing data through histograms or frequency polygons, helping teachers and students see score distribution trends at a glance.

10. Construct four possible research titles based on poor academic performance.

An Investigation into the Impact of Teacher-Student Relationships on Academic Performance in Secondary Schools.

The Effect of Home Environment on Students' Academic Performance in Rural Secondary Schools.

Assessing the Influence of Classroom Resources Availability on Academic Achievement in Public Secondary Schools.

An Analysis of Learners' Attitudes and Its Impact on Academic Performance in Mathematics in Selected Secondary Schools.

SECTION B (60 Marks)

Answer all questions from this question. Each question carries **fifteen (15)** marks.

11. A teacher recorded students' scores in a Biology test as follows: 70, 55, 60, 73, 65, 58, 69, 65, 58, 70, and 69.

(a) The mode is the number that appears most frequently in a data set. In this case, the scores 70, 69, 65, and 58 each appear twice. Since more than one value appears with the same highest frequency, this is a multimodal distribution. A multimodal distribution occurs when a data set has more than one mode, indicating that there are multiple common scores within the group.

(b) To find the mean, add all the scores and divide by the number of scores.

The total is 70 + 55 + 60 + 73 + 65 + 58 + 69 + 65 + 58 + 70 + 69 = 712. There are 11 scores, so the mean is 712 ÷ 11 = 64.73.
To find the median, arrange the scores in ascending order: 55, 58, 58, 60, 65, 65, 69, 69, 70, 70, 73. The median is the middle value. Since there are 11 numbers, the 6th score is the median, which is 65.

(c) A T-score is a standard score calculated using the formula $T = 50 + 10(X - M) / SD$, where X is the score, M is the mean, and SD is the standard deviation.

First, find the mean (M = 64.73 as computed earlier) and the standard deviation (SD).
Calculating SD:

Variance = $\Sigma(X - M)^2 / N$
= $(70-64.73)^2 + (55-64.73)^2 + (60-64.73)^2 + (73-64.73)^2 + (65-64.73)^2 + (58-64.73)^2 + (69-64.73)^2 + (65-64.73)^2 + (58-64.73)^2 + (70-64.73)^2 + (69-64.73)^2$
= 27.73 + 94.73 + 22.33 + 68.33 + 0.07 + 45.23 + 18.23 + 0.07 + 45.23 + 27.73 + 18.23
= 368.93

Variance = 368.93 / 11 = 33.54
SD = $\sqrt{33.54} \approx 5.79$
Now, T-score for the highest score (73):
T = 50 + 10(73 - 64.73) / 5.79
= 50 + 10(8.27 / 5.79)
= 50 + 10(1.43)
= 50 + 14.3
= 64.3

T-score for the lowest score (55):
T = 50 + 10(55 - 64.73) / 5.79
= 50 + 10(-9.73 / 5.79)
= 50 + 10(-1.68)
= 50 - 16.8
= 33.2

So, the highest score standardises to a T-score of 64.3, and the lowest to 33.2.

2. Explain six features to adhere to when organising the report.
A good report must have a clear and descriptive title. The title should reflect the content of the report, helping the reader quickly understand what the report is about.

The report should include an introduction that states the objectives, background, and scope of the report. This section provides context and guides the reader on what to expect.

It should be logically structured with headings and subheadings. Proper organisation makes it easy for readers to navigate different sections of the report.

Another feature is accuracy and completeness. The report should present facts, data, and interpretations precisely and should cover all aspects of the topic being addressed.

A report must maintain objectivity and neutrality. It should avoid personal bias or opinions unless it is a recommendation section, ensuring the findings are based on evidence.

Finally, it should conclude with a summary of the key findings and recommendations, giving clear and practical solutions based on the report's findings.

3. Clarify six points on the importance of carrying out regular assessments to tutors.

Regular assessments help tutors to monitor the academic progress of their students. Through continuous evaluation, tutors can identify students' strengths and weaknesses in specific subjects.

They enable tutors to improve their teaching methods. By assessing student outcomes, tutors can reflect on which teaching strategies are effective and which need to be modified.

Assessments also help in setting realistic academic goals for both tutors and students. They offer data that guides goal-setting and performance expectations in the classroom.

Another importance is that they motivate students to maintain consistent study habits. Knowing that their understanding will be regularly evaluated encourages students to stay engaged in their studies.

Regular assessments contribute to the identification of individual learning needs. Tutors can plan remedial lessons or special assistance for learners who lag behind in certain areas.

Finally, assessments assist educational authorities in evaluating the overall effectiveness of the curriculum and instructional quality within schools, using assessment results to guide policy decisions.

4. Give six descriptions that support the statement that the research process is a systematic and overlapping process.

The research process follows a systematic sequence of steps, starting from problem identification, literature review, hypothesis formulation, data collection, analysis, and conclusion. Each stage builds upon the previous one in a structured manner.

Despite having a sequence, the process is overlapping because some stages interact with or influence others. For example, data analysis might prompt a revision of the hypothesis or lead to additional data collection.

Another overlapping aspect is the literature review. Although it is typically conducted at the beginning, new relevant literature might emerge during data collection or analysis, requiring the researcher to revisit this step.

The development of research instruments and data collection often occur simultaneously. While instruments are being tested, preliminary data can reveal flaws that necessitate adjustments before full data collection continues.

Data interpretation sometimes overlaps with data analysis. As trends or patterns become visible during analysis, interpretation might begin before all data is fully processed, influencing further analytical focus.

Finally, writing the research report is also overlapping. Many researchers begin drafting background or methodology sections while still collecting or analyzing data, progressively refining the report as findings emerge.