UNIVERSITY OF SWAZILAND

FACULTY OF EDUCATION

DEPARTMENT OF EDUCATIONAL FOUNDATIONS AND MANAGEMENT

SUPPLEMENTARY EXAMINATION PAPER 2011/2012

TITLE OF PAPER: MEASUREMENT AND TESTING

COURSE NUMBER: EDF 321 PGCE

TIME ALLOWED: THREE (3) HOURS

INSTRUCTIONS: THERE ARE TWO SECTIONS IN THIS PAPER – SECTIONS A AND B.

ANSWER ALL ITEMS 1 TO 40 IN SECTION A. THERE IS ONLY ONE CORRECT ANSWER TO EACH ITEM. PUT A CIRCLE AROUND THE CORRECT RESPONSE ON THE ANSWER SHEET PROVIDED.

YOU ARE ADVISED TO SPEND 40 MINUTES ONLY IN SECTION A.

ANSWER THREE QUESTIONS IN SECTION B. USE ANSWER BOOK PROVIDED.

TOTAL MARK ALLOCATION: 100

Do not open until told to do so by the chief invigilator

Student ID# _____

1. The crucial aspects of the evaluation of school learning is:
    a. Measuring growth
    b. Measuring status
    c. Preparing tests
    d. Judging adequacy

2. The ultimate value of any given instrument lies in:
    a. The interpretability of its outcomes
    b. Its validity
    c. Its reliability
    d. The equality of its units

3. In criterion-referenced testing, we are primarily interested in comparing:
    a. Group performance and a standard
    b. Two or more groups
    c. And individual's performance to a domain
    d. Two or more individuals

4. Diagnostic tests are used primarily for:
    a. Summative evaluation
    b. University admission
    c. Grouping for instructional purposes
    d. Identification of individual weaknesses

5. Two terms which may be considered synonymous are:
    a. Assessment and evaluation
    b. Measurement and evaluation
    c. Measurement and assessment
    d. Testing and assessment

6. When a first grader explains the process by which a caterpillar changes into a butterfly, she is operating at which level of the taxonomy?
    a. Analysis
    b. Comprehension
    c. Knowledge
    d. Application

7. When a student can correctly divide any 4-digit number by any 2-digit number, the student, the student is operating at which level of the Bloom, et al., taxonomy?
    a. Knowledge
    b. Comprehension
    c. Analysis
    d. Application

8. The items on a test should be determined primarily by a consideration of:
    a. The ability levels of the students
    b. Our expectations of student performance

c. The purpose of the test
d. The attention span of students

9. An assignment to construct a table of specifications for a science test is at which level of the cognitive taxonomy?
   a. Evaluation
   b. Application
   c. Synthesis
   d. Comprehension

10. Multiple-choice and true-false items are preferred to essay items on the basis of their:
   a. Scope and speed
   b. Scoring reliability and aptitude validity
   c. Scoring reliability and content coverage
   d. Objectivity and accuracy

11. Objective items are "objective" in terms of their:
   a. Answer choices
   b. Content
   c. Directions
   d. Scoring

12. The most difficult part of writing multiple-choice items is:
   a. Estimating how long the test will take
   b. Thinking of reasonable distractors
   c. Having only one right answer
   d. Making the options roughly the same length

13. The problems of guessing the right answer on a multiple-choice item can be lessened by:
   a. Making the right answer more obvious
   b. Increasing the number of options
   c. Writing items at higher taxonomic levels
   d. Varying the length of the options

14. Which type of test would produce the highest score for a student who guessed on every item?
   a. Multiple-choice with four options
   b. Multiple-choice with five options
   c. True-false
   d. Matching

15. Good true-false items:
   a. Are more often false than true
   b. Tend to be no longer than one line
   c. Must be clearly true or false
   d. Have a penalty for guessing

16. Matching items lend themselves well to testing:

a. Computational procedures
b. Knowledge of relationships
c. Writing skills
d. Application of principles

17. The most desirable length of a matching item for typical classroom testing is:
    a. 9 to 12 premises
    b. 2 to 4 premises
    c. 5 to 8 premises
    d. 13 to 15 premises

18. Selected-response items tend to have objective scoring. This increases the test's:
    a. Measurement errors
    b. Reliability
    c. Subjectivity
    d. Difficulty

19. The combination of item formats which can accommodate the largest number of items in a specified testing time is:
    a. Multiple-choice and matching
    b. Multiple-choice and true-false
    c. Multiple-choice and short-answer
    d. Matching and short-answer

20. A student has an average quality response to an essay item. The score received by the student on this item will tend to be:
    a. Higher if preceded by low quality responses
    b. Higher if preceded by high quality responses
    c. Unaffected by the quality of preceding responses
    d. Lower if preceded by low quality responses

21. As teachers score essay responses, the score on an item:
    a. Tend to be highest middle in the scoring
    b. Tend to be highest early in the scoring
    c. Tend to be highest later in the scoring
    d. Are unaffected by position in the scoring

22. The extent to which a norm group is comparable to the group being tested determines the norm's:
    a. Stability
    b. Equivalence
    c. Relevance
    d. Representativeness

23. Which of these statistics is usually the largest for a given set of test scores?
    a. Mean
    b. Standard deviation
    c. Range
    d. Not necessarily any of these

24. Thembeni scores at the median. What would be his score in percentiles?
    a. 25
    b. 50
    c. 75
    d. 100

25. Percentiles, stanines, and standard scores are all examples of:
    a. Criterion-referenced scores
    b. Norm-referenced scores
    c. National scores
    d. Transformed distribution scores

26. A standard score of -5, when expressed as a T-score is:
    a. -5
    b. 5
    c. 95
    d. 45

27. Percentiles form what kind of measurement scale?
    a. Nominal
    b. Ratio
    c. Ordinal
    d. Interval

28. What level of measurement are standard scores?
    a. Nominal
    b. Ratio
    c. Ordinal
    d. Interval

29. In a normal distribution a percentile rank of 30 would be a stanine of:
    a. 2
    b. 3
    c. 4
    d. 5

30. The normal curve equivalent score:
    a. Can take on values 1 through 99
    b. Has a mean of 100
    c. Is based on an ordinal scale
    d. Has a standard deviation of 10

31. Which difficulty index is best for an item on a norm-referenced test?
    a. .10
    b. .50
    c. .80
    d. .20

32. Which of the following discrimination indexes indicates an item with a problem?
    a. -.50
    b. .75
    c. .40
    d. none of the above

33. An item has a discrimination index of -.40. This means that the item:
    a. Should be revised or eliminated
    b. Has a satisfactory difficulty index
    c. Is too easy for the good students
    d. Has a satisfactory discrimination index

34. Which of the following is most likely to increase the reliability of a test?
    a. Score each test paper twice
    b. Administer the test individually rather than in a group
    c. Increase test length
    d. Eliminate items that are difficult

35. Which of the following is most important in determining whether a test is criterion-referenced?
    a. A mastery cut-off score
    b. Specific instructional objectives
    c. A well specified domain of knowledge or skills
    d. A representative sample of examinees

36. An objective test is best determined in terms of:
    a. Responses given by testees
    b. The uniformity of scoring procedure
    c. High validity it has
    d. The number of options it contains

37. The most important advantage of the true-false test is:
    a. Its wide sampling
    b. Its elimination of guessing through correction
    c. The validity of the items
    d. Its high diagnostic value

38. Statistics is a tool by means of which we can:
    a. Synthesize the masses of quantitative data
    b. Estimate the extent of statistical data
    c. Predict the occurrence of phenomena
    d. Calculate the probability of occurrence of events

39. Reliability in a test refers to:
    a. Adequacy of standardization
    b. Consistency in the results
    c. Homogeneity in the content of the test
    d. Dependability of the author and publisher

40. When a distribution of achievement test scores is positively skewed, one might infer that the test was:
    a. Too hard
    b. Too easy
    c. Unreliable
    d. Poorly standardized
    e. Incorrectly administered

SECTION B

Answer any three (3) questions from this section.

Question 1 *(20 marks)*

Discuss the need for evaluation in education in relation to:

a. instructional decisions
b. curricular decisions
c. selection decisions
d. placement decisions
e. diagnosis

Question 2 *(20 marks)*

Define the following measures of variability;

a. The range
b. Variance
c. Standard deviation

Question 3 *(20 marks)*

Discuss the following issues about instructional objectives:

a. How to state general instructional objectives
b. How to state specific objectives
c. How to select appropriate objectives
d. What to consider when stating instructional objectives

Question 4 *(20 marks)*

Identify five (5) types of tests and write short notes about each one of them.

Question 5 *(20 marks)*

Name and discuss four scales of measurement.