

UNIVERSITY OF SWAZILAND  
FACULTY OF EDUCATION  
DEPARTMENT OF EDUCATIONAL FOUNDATIONS AND MANAGEMENT  
FINAL 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 LETTER OF THE CORRECT  
RESPONSE.

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. Which of the following is not a scale of measurement?
  - a. Ratio scale
  - b. Ordinal scale
  - c. Magnitude scale
  - d. Interval scale
  
2. Reliability in a test refers to:
  - a. Dependability of the author and publisher
  - b. Homogeneity in the content of the test
  - c. Adequacy of standardization
  - d. Consistency in the results
  
3. With respect to quantification, which concept is most unlike the others?
  - a. Aptitude
  - b. Weight
  - c. Attitude
  - d. Cognitive performance
  
4. In criterion-referenced testing, we are primarily interested in comparing:
  - a. Group performance and a standard
  - b. Two or more groups
  - c. An individual's performance to a domain
  - d. Two or more individuals
  
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 student lists the departments that are in the school, the task is at which level of the Bloom, et al., taxonomy?
  - a. Application
  - b. Analysis
  - c. Knowledge
  - d. Comprehension
  
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. 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
9. 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
10. Objective items are "objective" in terms of their:
- a. Answer choices
  - b. Content
  - c. Directions
  - d. Scoring
11. 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
12. 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
13. Multiple-choice items should:
- a. Use a short stem and longer options
  - b. Have alphabetically ordered distractors
  - c. Have reasonable distractors
  - d. Provide grammatical clues to the right
14. Matching items lend themselves well to testing:
- a. Computational procedures
  - b. Knowledge of relationships
  - c. Writing skills
  - d. Application of principles
15. Selected-response items tend to have objective scoring. This increases the test's:
- a. Measurement errors
  - b. Reliability
  - c. Subjectivity
  - d. Difficulty

16. Grammatical clues are most likely to be a problem for which item format?
- Fill-in
  - Short-answer
  - Multiple-choice
  - True-false
17. Which type of item is least susceptible to guessing?
- Matching
  - Multiple-choice
  - True-false
  - Short-answer
18. A student has an average quality response to an essay item. The score received by the student on this item will tend to be:
- Higher if preceded by low quality responses
  - Higher if preceded by high quality responses
  - Unaffected by the quality of preceding responses
  - Lower if preceded by low quality responses
19. As teachers score essay responses, the score on an item:
- Tend to be highest middle in the scoring
  - Tend to be highest early in the scoring
  - Tend to be highest later in the scoring
  - Are unaffected by position in the scoring
20. Stratifying on variables such as geographic region, socio-economic status, and percentage minority:
- Ensures a random sample
  - Ensures a representative sample
  - Eliminates the need for local norms
  - Ensures recency
21. Which of the following descriptive statistics is NOT a measure of dispersion?
- Range
  - Mean
  - Standard deviation
  - Variance
22. What is the mean of the following scores? 8, 4, 9, 1, 3?
- 6
  - 4
  - 5
  - 7
23. What is the median of the following scores? 8, 4, 9, 1, 3?
- 4
  - 5
  - 6
  - 7

24. Phinda has the highest score on a test that was given to 1000 people. His percentile rank would be:
- 50
  - 1
  - 99
  - 1000
25. Percentiles, stanines, and standard scores are all examples of:
- Criterion-referenced scores
  - Norm-referenced scores
  - National scores
  - Transformed distribution scores
26. Percentiles form what kind of measurement scale?
- Nominal
  - Ratio
  - Ordinal
  - Interval
27. Which scores **cannot** be meaningfully averaged?
- Percentiles
  - Standard scores
  - Raw scores
  - All of the above
28. Stanines assume that the distribution of scores is:
- Normal
  - Discrete
  - Rectangular
  - Based on a large number of people
29. What would be a stanine of a person who was at the 55<sup>th</sup> percentile?
- 4
  - 5
  - 6
  - 7
30. Which of the following scores provides the least precision in reporting student achievement?
- T-scores
  - Percentiles
  - Z-scores
  - Stanines
31. The correlation coefficient can take on values between:
- Zero and plus infinity
  - 1.0 and +1.0 inclusive
  - Minus infinity and plus infinity
  - Zero and +1.0 inclusive

32. Which of these item difficulty indexes indicates the easiest item?
- .10
  - .30
  - .60
  - .80
33. An item has a very low difficulty index. The discrimination index for the item will be:
- Low
  - High
  - Around .5
  - We cannot tell
34. Which of the following is most likely to increase the reliability of a test?
- Score each test paper twice
  - Administer the test individually rather than in a group
  - Increase test length
  - Eliminate items that are difficult
35. The variance of a distribution of test scores is a measure of:
- Dispersion
  - Central tendency
  - Relationship
  - Location
36. Which combination of conditions is NOT possible? A test has:
- High validity, low reliability
  - Low validity, high reliability
  - High validity, high reliability
  - Low validity, high reliability
37. The validity of most concern with teacher-constructed tests in school subjects is:
- Concurrent
  - Construct
  - Content
  - Predictive
38. Which of the following is most important in determining whether a test is criterion-referenced?
- A mastery cut-off score
  - Specific instructional objectives
  - A well specified domain of knowledge or skills
  - A representative sample of examinees
39. One of the most frequent pitfalls in the use of completion-type items is the tendency:
- To require too much recall
  - To permit guessing

- c. To permit subjectivity in scoring
- d. To restrict the test to measurement of facts

40. Which of the following is the greatest weakness of essay-type items from the standpoint of usability?

- a. It does not test all that it purports to test
- b. It requires a great deal of time to grade
- c. It is time-consuming to prepare
- d. The grading is subjective

## SECTION B

Answer any three (3) questions from this section.

### Question 1 (20 marks)

- a. Discuss the purpose of measurement and testing in schools.
- b. Describe the major differences and similarities between formative and summative evaluation. Give concrete examples

### Question 2 (20 marks)

Discuss the following attributes of a good measurement instrument

- a. Reliability
- b. Validity
- c. Usability

### Question 3 (20 marks)

Define the following measures of central tendency:

- a. Mean
- b. Median
- c. The mode

Illustrate with examples.

### Question 4 (20 marks)

Describe the process of planning the test with specific reference to test objectives, content area, types of items to use, and total number of items for a test.

### Question 5 (20 marks)

Select three of the following objective-type items and for each one describe characteristics, uses, advantages, limitations, and rules for construction:

- a. True-false items
- b. Short-answer items
- c. Matching exercise
- d. Multiple-choice form