1. **Define the term measurement? (From Topic 1)**

Measurement is the process by which the attributes or dimensions of some object (both physical and abstract) are quantified.

2. **What is the table of specifications?**

One of the tools used by the teachers to develop a blueprint for the test is called the table specification.

3. **Write any three limitations of MCQS .(From Topic 75)**
4. **Limitations of MCQS:**

i. Take a long time to construct in order to avoid arbitrary and ambiguous questions.
ii. Also require careful preparation to avoid preponderance of questions testimony recall.
iii. Provide cues that do not exist in practice.
iv. **What is the predictive validity?**

Predictive validity refers that how much test can predict future performance with reference to some defined standard.

5. **What is the diagnostic assessment? (From Topic 4)**

Diagnostic assessment is the type of the assessment that determines the causes intellectual, emotional, and environmental of persistent learning difficulties.

6. **Describe the Bloom's taxonomy of cognitive domain?**

There are six level's of the cognitive domain: 1. Knowledge 2. Comprehension 3. Application 4. Analysis 5. Synthesis 6. Evaluation

7. **What is the test security? (From Topic 106)**

Test security is the process intended or reveal flaws in the security mechanism of an information system that protect data and maintain functionality as intended.

8. **How to design rubric?(From Topic 112)**

We can design a rubric by the following some steps: 1. Identify a learning goal. 2. Choose outcomes that may be measured. 3. Develop or adapt an existing rubric. 4. Share it with students.

9. **How to assemble or packaging a test?(From Topic 103)**

Assembling Test (Packaging the Test):

Solved By Eshal MAlik

i. Grouping together items of similar format
ii. Arranging test items from easy to hard
iii. Properly spacing items
iv. Keeping items and options on the same page
v. Placing illustrations near the descriptions
vi. Checking the answer key
vii. Deciding how students will record their answers 8. Providing space for test taker's name and date
viii. Checking test directions for clarity

10. what are subjective type questions? Tell names

Subjective test items (essay questions) are constructed response type questions that can be the best way to measure the students' higher order thinking skills, such as applying, organizing, synthesizing, integrating, evaluating, or projecting while at the same time providing a measure of writing skills. Essay items can vary from very lengthy (5-10 pages), open-ended to limited or restricted response (one page or less). Therefore, essay type questions are divided into two types

i. Restricted Response Items
ii. Extended Response Items

11. High stake testing matters

High-stakes testing as the use of a summative test or an assessment (i.e., a test designed to measure student achievement after a period of instruction has been completed) to make decisions that are of prominent educational, financial, or social impact.

i. Types of High stake test
ii. Criterian Referenced
iii. Test Norm-Referenced High-Stake Tests

12. **Advantage if essay type questions**

Extended response questions have following advantages over other types of question formats: 1. Effective for assessing higher order abilities: analyses synthesize and evaluate. 2. It is comparatively less time consuming to develop such items. 3. Emphasizes essential communicational skills. 4. Guessing is eliminated.

13. **Benefits of rubric in evaluation process**

A rubric is an explicit set of criteria used for assessing a particular type of work or performance and provides more details than a single grade or mark. It is a set of scoring guidelines for evaluating student work.

Solved By Eshal MAlik

14. **Item analysis**

The item analysis is an important phase in the development of a test. In this phase statistical methods are used to identify any test items that are not working well. If an item is too easy, too difficult, failing to show a difference between skilled and unskilled examinees, or even scored incorrectly, an item analysis will reveal it.

15. **Extended response type questions with example**

It allows the students to determine the length and complexity of the response. This type of questions is most suitable for measuring higher level mental processes skills like synthesis and evaluation. Teachers prefer these questions when they want to have information about student's ability to organize, integrate, express and evaluate information.

Example: What are the teachings of Islam about the respect of elders? How do these teachings help us to make a good society? Marks will be given for correct information, your own point of view, good judgment and good presentation.

16. **What is report? Name of reports?**

Reporting is the process of communicating the knowledge and understanding gained from assessing a student's work to improve learning There are two forms of reports 1. Formal reports 2. Informal reports

17. **Explain formal and informal report**

Formal reports

Formal reports communicate about important aspects of the students' progress in the areas of academic, social, human and career development to parents and students.

Informal reports

Informal report tells parents about what their child is able to do. It provides information about the areas of learning that need further attention or development and the ways the teacher is supporting the student's learning needs.

18. **Develop 2 restricted response questions.**

Q: List the similarities and differences in the process of cell division in meiosis and mitosis?

19. **How to conduct parent teacher meeting.**

i. Make plans for the meeting.
ii. Begin the meeting in a positive manner.

Solved By Eshal MAlik

iii. Present the student's strong points before describing the areas needing improvement.
iv. Encourage parents to participate and share information.
v. Plan a course of action cooperatively.
vi. End the meeting with a positive comment.
vii. Use good human relation skills during the meeting

20. **Fixed choice assessment and complex performance assessment.**

By format of assessment

i. Fixed Choice Assessment
ii. Complex Performance Assessment

Fixed Choice Assessment

Fixed Choice Assessment is sued to measure the skills of people efficiently (means measure more skills in less time) and for this we usually use fixed choice items i.e. Multiple choice question, matching exercise, fill in the blanks and true false. It is called fixed choice because the person who is attempting the paper does not need to write the answer, just need to choice the answer.

Complex Performance

Assessment Complex Performance assessment is used for measurement of performance in contexts and the problems valued in their own right. This includes hands on laboratory experiments, projects, essays, oral presentations.

21. **State 5 rules for constructing alternative form questions.**

i. Avoid broad general statements if they are to be judged true or false.
ii. Avoid trivial statements
iii. Avoid the use of negative especially double negative statements.
iv. Avoid to include two ideas in one statement unless cause and effect relationship are being measured
v. Avoid long complex sentences

22. **Write a brief note on the restricted response essay type items?**

Pose a specific problem for which student needs to recall suitable information, organize it, derive a defensible conclusion, and express it within the given limits the questions (like 400 word, 30 min. etc.).

Example: List the similarities and differences in the process of cell division in meiosis and mitosis. A variety of learning outcomes (Linn, 2000) can be checked by using this format of essay question.

Solved By Eshal MAlik

Some of which are:

i. Analysis of relationship.
ii. Compare and contrast positions.
iii. Explain cause-effect relationship.
iv. Organize data and support a viewpoint.
v. Formulate hypotheses.
vi. Point out strengths and weaknesses.
vii. Integrate data from various resources

Teacher can use this type of questions under the following conditions when:

i. Supplying information is required instead of simple recognition.
ii. Limited numbers of content areas are needed to be tested.

If it is feared that students may help each other unfairly during the test which is difficult to control. (Especially if the number of students is large and class teacher cannot invigilate all of them as keenly as needed). Student is to decide length, content to be included, its organization and order of presentation min the answer.

23. **How to write essay type questions? (From Topic 90)**

Following guidelines are provided by Haldane and Downing (1989) for writing essay type items when developing a test.

i. Frame questions so that the examinee's task is explicitly defined.
ii. Specify the value and an approximate time limit for each question.
iii. Employ a larger number of questions that require relatively short answers rather than only a few questions that require long answers.
iv. Do not employ optional questions.
v. Verify a question's quality by writing a trial response to the question.
vi. Prepare a tentative scoring key in advance of considering examinee responses.
vii. Score all answers to one question before scoring the next question.
viii. Make prior decisions regarding treatment of such factors as spelling and punctuation.
ix. Evaluate essay responses anonymously.

24. Explain Criterion and predictive Validity.

Criterion validity refers how much test can measure some performance with reference to some defined standard. For example: The student is expected to be able to speak fluently after completing a course. So, the test is valid if it can measure spoken ability.

Solved By Eshal MAlik

Predictive validity refers that how much test can predict future performance with reference to some defined standard.

25. Explain the Taxonomy levels of Cognitive.

i. Knowledge is defined as remembering of previously learned material. This may involve the recall of a wide range of material, from specific facts to complete theories, but all that is required is the bringing to mind of the appropriate information. Knowledge represents the lowest level of learning outcomes in the cognitive domain.

Verbs: arrange, define, duplicate, label, list, memorize, name, order, recognize, relate, recall, repeat, reproduce state.

ii. **Comprehension** is defined as the ability to grasp the meaning of material. This may be shown by translating material from one form to another (words to numbers), by interpreting material (explaining or summarizing), and by estimating furture trends (predicting consequences or effects). These learning outcomes go one step beyond the simple remembering of material, and represent the lowest level of understanding.

Verbs: classify, describe, discuss, explain, express, identify, indicate, locate, recognize, report, restate, review, select, translate.

iii. Application refers to the ability to use learned material in new and concrete situations. This may include the application of such things as rules, methods, concepts, principles, laws, and theories. Learning outcomes in this area require a higher level of understanding than those under comprehension.

Verbs: apply, choose, demonstrate, dramatize, employ, illustrate, interpret, operate, practice, schedule, sketch, solve, use, write.

iv. Analysis refers to the ability to break down material into its component parts so that its organizational structure may be understood. This may include the identification of the parts, analysis of the relationships between parts, and recognition of the organizational principles involved. Learning outcomes here represent a higher intellectual level than comprehension and application becasue they require an understanding of both the content and the structural form of the material.

Verbs: analyze, appraise, calculate, categorize, compare, contrast, criticize, differentiate, discriminate, distinguish, examine, experiment, question, test.

v. **Synthesis** refers to the ability to put parts together to form a new whole. This may involve the production of a unique communication (theme or speech), a plan of operations (research proposal), or a set of abstract relations (scheme for classifying

Solved By Eshal MAlik

information). Learning outcomes in this area stress creative behaviors, with major emphasis on the formulation of new patterns or structures.

Verbs: arrange, assemble, collect, compose, construct, create, design, develop, formulate, manage, organize, plan, prepare, propose, set up, write.

vi. **Evaluation** is concerned with the ability to judge the value of material (statement, novel, poem, research report) for a given purpose. The judgements are to be based on definite criteria. These may be internal criteria (organization) or external criteria (relevance to the purpose) and the student may determine the criteria or be given them. Learning outcomes in this area are highest in the cognitive hierarchy because they contain elements of all the other categories, plus conscious value judgements based on clearly defined criteria.

Verbs: appraise, argue, assess, attach, choose compare, defend estimate, judge, predict, rate, core, select, support, value, evaluate.

26. What is the formula for the find out the item discrimination?(From Topic 97)

ID is = (Upper Group % Correct) – (Lower Group % Correct)

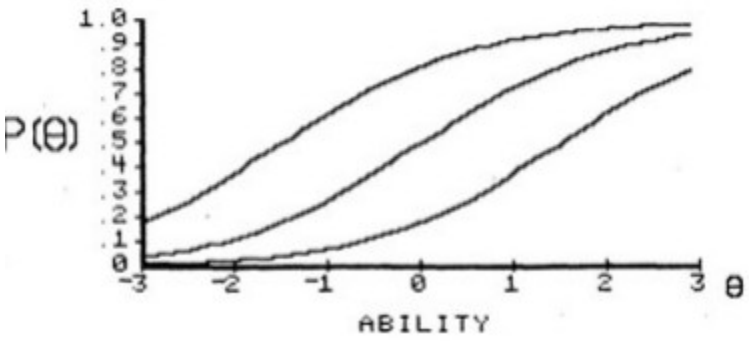
27. Describe the advantages for the essay writing test?

I. Effective for assessing higher order abilities: analyses synthesize and evaluate.
II. It is comparatively less time consuming to develop such items.
III. Emphasizes essential communicational skills.
IV. Guessing is eliminated

28. Explain Item difficulty in IRT.

The application of Item difficulty in IRT is defined as the ability at which the probability of success on the item is .5 on a logit scale, which is also known as threshold difficulty (see Fig. 3). An item's level of difficulty is another factor affecting an individual's probability of responding in a particular way. An item that has a high level of difficulty will be less likely to be answered correctly by an examinee with low ability than an item that has a low level of difficulty (i.e., an easy item).

Solved By Eshal MAlik



29. What is the Rubric, Write the element of the rubric?

Rubric:

A rubric is an explicit set of criteria used for assessing a particular type of work or performance and provides more details than a single grade or mark. It is a set of scoring guidelines for evaluating student work. Scores or levels identified in a scoring rubric must be descriptive, not merely judgmental in nature. Example: Define the level of rubric as “Writing is clear and thoughts are complete” as compared to “Excellent.”

Elements of Rubric: A rubric includes: 1. Score 2. Criteria 3. Levels of performance 4. Descriptors

30. What is the self-assessment?

Self-assessment: Self-assessment is a process of formative assessment during which students reflect on and evaluate the quality of their work and their learning, judge the

31. What is Analytic rubric?

Each criterion is assessed separately by using different descriptive ratings.

Each criterion is given a separate score.

Final score is made up of adding each component parts.

It takes more time to score but gives detailed more feedback.

32. **Describe the Advantages of Self-Assessment.**

I. Self-assessment can promote learner responsibility and independence.
II. Self-assessment tasks encourage student ownership of the learning.
III. Self-assessment tasks shift the focus from something imposed by someone else to a potential partnership.
IV. Self-assessment emphasizes the formative aspects of assessment.

Solved By Eshal MAlik

V. Self-assessment encourages a focus on process.
VI. Self-assessment can accommodate diversity of learners‟ readiness, experience and backgrounds.

33. What is the Syllabus criteria for the higher education in Pakistan?
34. **What are the objective test item types?**

There are three types of items which are included in the category of selection type items or the objective test items types 1. : multiple choice items, 2. alternative form or 3. true-false and matching.

35. **Define Reliability with examples.**

Reliability refers to the extent to which assessment results are consistent. The assessment that maintains the consistency in results is known as reliable.

“Reliability is the extent to which a measurement instrument or procedure yields the same results as on repeated trails (Carmines and Zeller, 1979)” For example, a calculator cannot be reliable if it gives wrong calculations or results. The weight of an object is not reliable if machine is not working properly. Same is the case of the results obtained from assessments.

36. What is meant by the term of unreliable results?

The unreliable results of assessment mean inconsistency. Keeping these views in mind, we can trust the results if these are consistent. It is the evidence ofan accurate system of assessment. Keeping in view the significance of reliability, different methods are used to attain reliability. For example, testretest procedure, the alternative–test form procedure and the split halves procedure.

37. Define the evaluation?

Evaluation is process of making a value judgment against intended learning outcomes and behavior, to decide quality and extent of learning.

38. What is the item difficulty?

Item difficulty is the is a measure of the proportion of examinees who responded to an item correctly.

39. What are the importance of the student self-assessment?

I. Clarifying, sharing, and understanding learning intentions and criteria for success–getting the students to really understand what their classroom experience will be and how their success will be measured.

Solved By Eshal MAlik

II. Engineering effective classroom discussions, activities, and learning tasks that elicit evidence of learning– developing effective classroom instructional strategies that allow for the measurement of success.
III. Providing feedback that moves learning forward– working with students to provide them the information they need to better understand problems and solutions.
IV. Activating learners as instructional resources for one another– getting students involved with each other in discussions and working groups can help improve student learning.
V. Activating learners as owners of their learning– self-regulation of learning leads to student performance improvement.

40. What is the alternative form of test?

Alternative form of question requires students to select any one of the two given categories. The categories may be True-false, yes-no, correct-incorrect or fact-opinion (Airasian, 1994). Alternative form of items is most suitable for measuring lower level learning outcomes.

41. What is the maximum and Typical performance?

Maximum Performance:

What a person can do. For example how you drive on a drivers test. Happens when someone knows they are being observed.

Typical Performance:

what a person will do – ie. how you drive normally. To observe typical performance, you have to do it when people don't know they are being observed or be there on enough occasions that people get used to the examinee, or don't know what they are being tested on.

42. What is face validity and content validity?

Face Validity: - Does the test appear to test what it is supposed to be tested? - A drawing test must be related to figures, so the question paper should appear with figures. A test that asks for filling some blanks must have blanks on paper. A multiple choice question must have options along with it.

Content Validity: Content validity refers to the contents a test is suppose to measure. For instance, a book has ten chapters the chapters are the contents of the book.

43. Describe the curriculum in Pakistan at higher level education?

I. In 1973's constitution, Government placed education on the concurrent list.
II. Federal Government took the responsibility to determine the curriculum text boards, policy, planning and standards of education.

Solved By Eshal MAlik

III. In December, 1976 Federal Government appointed University Grant commission (now HEC) as the competent authority to look after the curriculum revision work at bachelor level and onwards .

IV. It also look after the degrees, certificates and diplomas awarded by degree colleges, universities and other institutions of HEC.

44. Describe the item difficulty in CTT?

At the item level, the Classical Test Theory (CTT) model is relatively easy. CTT does not raise a complex theoretical model to relate an examinee's ability to the probability of success on a particular item. Instead, CTT collectively considers a pool of examinees and empirically examines their success rate on an item (assuming it is dichotomously scored). This success rate of a particular pool of examinees on an item, well known as the p value of the item, is used as the index for the item difficulty (Linden& Hambleton, 2004). (Actually, it is an inverse indicator of item difficulty, with higher value indicating an easier item). In CTT, the item difficulty index p (p value), is the proportion of examinees correct on an item, expresses item difficulty on an Item.

Item difficulty in CTT is simply calculated by the percentage of students that correctly answered the item as refers to the p value which range from .00 to 1.00. The values closer to 1 more easy will be the item and conversely the values near to .00 the more difficult will be the item. The values lie somewhere in the middle i.e. 0.4 to 0.6 will refer to moderate item difficulty index. Item difficulty: the percentage of students that correctly answered the item. The range is from 0% to 100%, or more typically written as a proportion of 0.0 to 1.00. The higher the value, the easier will be the item.

Calculation: Divide the number of students who got an item correct by the total number of students who answered it.

Ideal value: Slightly higher than midway between chance (1.00 divided by the number of choices) and a perfect score (1.00) for the item. For example, on a four- alternative, multiplechoice item, the random guessing level is 1.00/4 = 0.25; therefore, the optimal difficulty level is .25 + (1.00 - .25) / 2 = 0.62. On a true-false question, the guessing level is (1.00/2 = .50) and, therefore, the optimal difficulty level is .50+(1.00- .50)/2 = 0.75. P-values above 0.90 are very easy items and should be carefully reviewed based on the instructor's purpose. For example, if the instructor is using easy "warm-up" questions or aiming for student mastery, than some items with p values above .90 may be warranted. In contrast, if an instructor is mainly interested in differences among students, these items may not be worth testing. P-values below 0.20 are very difficult items and should be reviewed for possible confusing language, removed from subsequent exams, and/or identified as an area for re-instruction. If almost all of the students get the item wrong, there is either a problem with the item or

Solved By Eshal MAlik

students were not able to learn the concept. However, if an instructor is trying to determine the top percentage of students that learned a certain concept, this highly difficult item may be necessary. The optimal item difficulty depends on the question-type and on the number of possible distractors.

45. Explain the charactertics of the Curve.

The item characteristic curve is considered as the basic building block of item response theory; all the other constructs of the theory depend upon this curve. Therefore, significant attention is devoted to this curve and its role within the theory. There are two methodological properties of an item characteristic curve that are used to describe it: one is the difficulty which under item response theory describes the item functions along the ability scale. For example an easy item functions among the low- ability examinees and a hard item functions among the highability examinees; thus, difficulty is a location index. The second technical property is discrimination, which describes how well an item can differentiate between examinees having abilities below the item location and those having abilities above the item location (Baker, 2001). An item characteristic curve is the graphical representation of the probability of answering an item correctly with the level of ability on the construct being measured. It gives you a picture of:

I. The item difficulty
II. Discrimination power c.
III. The probability of answering correctly by guessing

46. What is the Prohibited Professional Test Security Practices?

I. Give students access to test content or secure test questions before test administration.
II. Copy reproduces or uses in any manner inconsistent with test security measures of secure assessment material.
III. Review actual test items before, during or after the test administration.
IV. Fail to follow directions for the distribution and return of secure material or fail to account for any secure materials before, during or after test administration.
V. Leave a testing room unsupervised at any time.
VI. Make test answers available to students or permit the use of any supplemental or reference materials during test.
VII. Assist a student directly or indirectly (gestures, pointing, prompting etc.,)
VIII. Disclose or discuss the content of the tests with students, teachers, parents, other educators or community members before, during or after testing. 9. Making any changes to student's responses.

47. Standardized Achievement Test

1. **Standardized Achievement Test Achievement Test**

Solved By Eshal Malik

A standardized achievement test a fixed set if items to measure defined achievement domain, specific direction for administrating and scoring test, and norms based on representative groups of individuals. Most published achievement test are called standardized achievement tests. These typically are norm referenced tests. Quite a few criterion referenced achievement tests are also been published. Achievement tests are used as part if broader assessment system or alone. They provide relatively inexpensive means of measuring broad achievement goals. Standardized achievement tests are often customized to include characteristics of both norm and criterion referenced test. Standard content and procedure makes it possible to give an identical test to individuals in different places at different times. Equivalent forms are included in many standardized tests, which make it possible to repeat the test without fear that the test takers will remember the answers from first testing.

48. Characteristics of Standardized Achievement Test

I. The test items are of highly technical quality. They have been developed by educational and test specialists, pretested and selected on the basis of difficulty discriminating power, and relationship to a clearly defined and rigid set of specifications.
II. Directions for administering and scoring are so precisely stated that the procedures are standard for different users of test.
III. Norms based on national samples of students in the grade where the test is intended for use are provided as aids in interpreting the test scores.
IV. Equivalent and comparable forms of the test are usually provided, information concerning the degree to which the forms are comparable.
V. A test manual is used as guides for administering the test, evaluating its technical qualities, and interpreting and using the results.

49. Standardized Test versus Informal Classroom Test

Standardized tests and carefully constructed test are common in many ways. The main differences between the two types are:

I. The nature of the learning outcomes and the content measured.
II. Quality of test items
III. Reliability of the tests
IV. Procedure for administering and scoring
V. Interpretation of scores

The standardized tests inflexibility makes it less valuable for those purposes for which the informal classroom tests is so admirably suited.

Solved By Eshal MAlik

50. Table of specification One of the tools used by teachers to develop a blueprint for the test is called ―Table of Specification‖ in other words Table of Specification is a technical name for the blue print of test. It is the first formal step to develop a test.

Concept of Table of Specification

I. It helps a teacher in allotting the questions to different content areas and Bloom's learning categories in a systematic manner.
II. The blueprint is meant to insure content validity. Content validity is the most important factor in constructing an achievement test. (will be discussed in later unit)
III. A unit test or comprehensive exam is based on several lessons and/or chapters in a book supposedly reflecting a balance between content areas and learning levels (objectives).

Two way Table of Specification

A Table of Specifications consists of a two-way chart or grid relating instructional objectives to the instructional content.

Table of specification performs two important functions

I. Ensures the balance and proper emphasis across all content areas covered by teacher
II. It ensures the inclusion of items at each level of the cognitive domain of Bloom's Taxonomy.

51. Structure of Observed Learning Outcomes (SOLO)

The taxonomy of Structure of Observed Learning Outcomes (SOLO) was initially developed by Biggs and Collis in 1982, and then well described in Biggs and Tang in 2007. It carries five different levels of competency of learners Levels of

Structure of Observed Learning Outcomes (SOLO) 1. Pre-structural 2. Uni-structural 3. Multi-structural 4. Relational 5. Extended Abstract

52. Depth of Knowledge (DOK) DOK (Depth of Knowledge) was presented by Webb in 1997, giving four levels of learning activities Levels of Depth of Knowledge (DOK)

1. Recall 2. Skill/Concept 3. Strategic Thinking 4. Extended Thinking

53. Multiple Choice Items

The multiple choice items is generally recognized as the most widely applicable and useful type of objectives test items that can measure knowledge, comprehension and as well as application level learning outcomes.

Multiple choice items consist of two parts:

Solved By Eshal MAlik

I. A Stem Stem

The problem may be stated as direct question or an incomplete statement and is called the stem of the item.

II. No. of options or alternatives

Alternatives The list of suggested solutions which may include words, numbers, symbols or phrases are called alternatives (also called choices or options).

54. Advantages of MCQs

i. Ensure objectivity, reliability and validity; preparation of questions with colleagues provides constructive criticism.
ii. Increase significantly the range and variety of facts that can be sampled in given time.
iii. Provide precise and unambiguous measurement of the higher intellectual processes.
iv. Provide detailed feedback for both student and teachers.
v. MCQ are easy and rapid to score.

55. Decision based on high stake testing

i. Promotion to next grade
ii. Awarding of diploma or degree
iii. Evaluation of school performance
iv. Incentives and accountability of school staff

56. High Stake Testing in Pakistan

In school education in Pakistan, high stake testing is conducted at each level from primary to higher secondary.

i. Grade 5 and 8 examination by Punjab Examination Commission.
ii. Grade 10 and 12 examination by BISEs

Test construction in High Stake testing

- The format of items are same as used in CRT in classroom
- Item of developed by professional employed in dedicated organization
- Item banks are developed
- Psychometric properties of item are tested
- The process goes on round the year

57. **Criticism on high stake testing**

High stake tests have same issues as classroom test but much larger impact and consequences.

58. Recommendations for effective High Stake Testing

Solved BY Eshal MAlik

i. Protection against high stake decisions based on single test
ii. Adequate resources and opportunity to learn
iii. Full disclosure of likely consequences
iv. Alignment between test and curriculum
v. Appropriate attention towards language difference between examinees
vi. Careful adherence to explicit rules for determining which students are to be tested

59. **Institutions involve in assessment**

i. Examination commission
ii. Board of intermediate and Secondary education (BISEs)
iii. Board of Technical Education

60. **National Education Assessment System (NEAS)**

National Education Assessment system (NEAS) works for promoting quality learning among children of Pakistan by carrying out fair and valid national assessment.

Objectives of NEAS

i. Informing policy
ii. Monitoring standards
iii. Identifying correlation of achievement
iv. Directing teachers‘ efforts and raising students‘ achievement

61. **National Education Assessment system**

The National Education Assessment System has been institutionalized in Pakistan at national level with the cooperation of provincial and area Assessment Centers.

Objectives of NEAS

i. Informing policy: the extent to which geography and gender are linked to inequality in student performance.
ii. Monitoring standards: How well the curricula are translated into knowledge and kills
iii. Identifying correlation of achievement: the principal determinants of student performance
iv. Directing teachers‘ efforts and raising students‘ achievement: assisting teachers to use data to improve student performance.

62. **Test- Retest Method**

Test-Retest method is a measure of stability

It gives the same test twice to the same group with any time interval between tests, Time interval can range from several minutes to the several years.

Solved By Eshal MAlik

63. Guidelines for constructing Restricted response essay type items

In this session student will learn:

Guidelines for constructing Restricted response essay type items.

Guidelines for constructing Restricted response essay type items 1. Statements should not be quoted directly from the text.

i. Evaluate essay responses anonymously.
ii. Frame questions so that the examinee‘s task is explicitly defined.
iii. Specify the value and an approximate time limit for each question.
iv. Employ a larger number of questions that require relatively short answers rather than only a few questions that require long answers.
v. Do not employ optional questions.
vi. Verify a question‘s quality by writing a trial response to the question.
vii. treatment of such factors as spelling and punctuation

64. Limitations of essay type items.

These advantages come with the following limitations:

i. The scoring is unreliable and time consuming.
ii. Limited sampling of the content is possible.

65. Scoring Rubrics for Essay Type Items

Scoring rubrics provide at least two benefits in the evaluation process.

i. they support the examination to the extent to which the specified criteria has been reached.
ii. They provide feedback to students concerning how to improve their performance.

66. Types of Scoring Rubrics

i. Types of Scoring Rubrics for Essay Type Items

There are two general methods for scoring subject-matter essays: the analytic (scoring key, point or rating) method and the holistic (also called global, sorting or rating) method.

Holistic Scoring Rubric for Essay Type Items

The holistic scoring rubric requires making a judgment about the overall quality of each student‘s response to an item. No need to mark each specific content element that a student included in the answer.

Solved BY Eshal MAlik

Holistic rubric is probably more appropriate for extended response essay type items involving a student‘s abilities to synthesize and create and when no definite answer can be prespecified.

67. Theories of Test Development

There are two widely perceived theories in psychosocial measurement, classical test theory (CTT) and item response theory (IRT) as both of these theories represents two different measurement frameworks.

ii. Item difficulty
iii. Item discrimination
iv. Chance of guessing

Classical test theory is more focused on finding the true score of an individual on a test and called as true score theory

Observed Score= True score + Error

Item response theory is known as latent trait theory as its focus is to find the item characteristics as a function of ability.

68. Appropriate time for Item Analysis

There are two stages at which items can be analyzed

i. After administration, but before scoring
ii. After scores have been reported

69. Test theories in item analysis

i. Classical Test Theory (CTT)
ii. Item Response Theory (IRT)

70. Types of Guessing?

There are two types of guessing

i. Blind guessing

Where an examinee chooses an answer at random among the alternatives offered.

ii. Informed guessing

Informed guessing: Where the examinee draws upon all his knowledge and abilities to choose the answer most likely to be correct.

71. Issues in test administration

Solved By Eshal MAlik

1. Cheating 2. Poor testing conditions 3. Test anxiety 4. Errors in test scoring procedure.\

72. Recording test items

The card contain information concerning the instructional objectives, specific objective, difficulty index, discrimination index and the content measured by the item should be prepared for each item to maintain its record.

73. Packing the test

Once you have measureable instructional objectives, test blueprint, written test items matching instructional objectives, so you are ready to

Package the test and

Reproduce the test

74. Reproducing the test

In schools it is usually by photocopying and quality of test copies may vary considerably.

Reproducing the Test Reproduction of test involves:

1. Knowing the photocopying machine 2. Specifying copying instructions Filing original test

75. Things to remember when administering test

Test is ready, get students ready.

1. Maintain a positive attitude
2. Maximize achievement motivation
3. Equalize advantages
4. Avoid surprises
5. Clarity the rules 6
6. Rotate distribution

Solved By Eshal MAlik