Download More Files from VUAnswer.com

MUHAMMAD IMRAN	EDU431 Test Development & Evaluation
CURRENT AND PAST MCQs AND QUESTIONS	FINAL TERM

QUIZ NO 1 | QUIZ NO 2 | QUIZ NO 3 | GRAND QUIZ | MID TERM | FINAL TERM

FINAL TERM CURRENT AND PAST SOLVED PAPER



MUHAMMAD IMRAN	EDU431 Test Development & Evaluation
covering Topic no. 1-36. May 25, 2021.	QUIZ NO 1

1. In DOK taxonomy; graph, separate, narrate are key words for learning level: _____**Skill/concept**
2. _____determines learning progress, provides feedback to reinforce learning, and correct learning errors_____ **Formative Assessment**
3. The domain in which intellectual skills are reflected:_____ **Cognitive**
4. Norm reference assessment and criterion reference assessment are classifications of assessment on the basis of:_____ **Method of interpreting results**
5. If you achieved 97% percentile on a test, then your result has been assessed according to:_____ **Norm reference**
6. Measurement and evaluation can be used to _____._____ **all of the above**
7. The type of assessment used for identifying the problems of students is called?____ **Formative**
8. To use previous learned material in new situation is called:_____ **Application**
9. Cognitive domain was divided into subgroups by Bloom in:_____ **1956**
10. who was the founder of Pakistan is the question of _____ in Blooms taxonomy of cognitive domain._____ **Knowledge**

Download More Files from VUAnswer.com

11. In revised Bloom taxonomy; assess, defend, appraise are key words for level:___**Evaluating**
12. In revised Bloom taxonomy; build, identify, develop are key words for level:_____**Applying**
13. In revised Bloom taxonomy; classify, compare, contrast are key words for level:___ _____**Understanding**
14. Ali's writing score placed near the bottom of the class. The interpretation of this statement is:
15. Identify the lowest learning level of Bloom's cognitive domain_____**Knowledge**
16. In norm referenced test the comparison is between _____. _____**Individuals**
17. The purpose of formative evaluation is _____. _**Monitoring progress of students**
18. The classification of cognitive domain was presented by_____ **Benjamin S. Bloom**
19. Which of the following decision determines prerequisite skills, degree of mastery of course goals and mode of learning?_____ **Placement**
20. Which of the following test with time limit strict that no one is supposed to complete all items?_____ **Speed**
21. Achievement tests and aptitude tests are example of:_____ _____ _____ **Maximum performance assessment**
22. Which of the following provide a framework of categories with different hierarchical levels of outcomes?_____ **Taxonomy**
23. If you have been declared as master of content in specific domain after a test, then your result has been assessed according to: _____**Criterion-referenced**
24. The test method to compare the performance of students with other students is called: _____**criterion referenced**

Download More Files from VUAnswer.com

25. _____ determines learning progress, provide feedback to reinforce learning, and correct learning errors _____**formative assessment**

26. ____ is used when we want to assess the student's prior knowledge: _____ _____**placement assessment**

27. Persistent difficulties in learning are investigated in: _____**diagnostic assessment**

28. Assessment that monitors learning progress is the _____**formative assessment**

29. "What is your name?" Is example of which level of solo taxonomy_____ **pre structural**

30. In norm referenced test the comparison is between _____**individuals**

31. Who was the founder of Pakistan is the question of ____ in bloom's taxonomy of cognitive domain? _____**knowledge**

32. student's performance is compared with other students in: _____**norm-referenced test**

33. Explain the process of water cycle is Question of ____ level in bloom's taxonomy of cognitive domain _____**Comprehension**

34. Measurement and evaluation can be used to: _____ _____**assess the effectiveness of learning activities**

35. Procedure used to determine person best abilities are: ___**maximum performance test**

36. Assessment that monitors learning progress is: _____**formative assistant**

37. Curriculum wing also known as: _____**national bureau of Curriculum**

38. Frequently used tools of summative evaluation: _____**test**

39. Multiple choice question, matching exercise, fill in the blanks and true false is example of _____**fixed choice question**

40. The purpose of evaluation is to make judgement about: _____**Quality**

41. ___ is used when we want to assess student's prior knowledge: _**Placement assessment**

42. Assessment that monitors learning progress is: _____**formative assessment**

43. Who was the founder of Pakistan is a question of ____ in bloom's taxonomy of cognitive domain: _____**knowledge**

44. The purpose of measurement in education is to _____ _____ **Collect information for decision making**

45. procedure used to determine person best abilities are ____**maximum performance test**

46. The purpose of evaluation is to make_____ **Judgement**

47. Feature, trait or dimension to be measured is decided in the ____ element of rubric. _**Criteria**

48. In which type of question marking will be more reliable ____. **Multiple Choice Question**

49. A test very popular with classroom teacher is _____**Multiple Choice Question**

50. **The domain which deals with physical abilities and coordination objectives is:** Download More Files from VUAnswer.com _____**Psychomotor**

51. **Which of the following type of test emphasized on reading and speaking? Verbal**

52. **The assessment determines what individual can do when performing at their best is called:**_____ **Maximum Performance Assessment**

53. Which of the following is not a basic step to design rubric answer _____ _____**choose outcomes that may be measured**

54. to give an answer or make a judgement about something without being sure of all the facts is called answer _____**guess**

55. Difficulty level of an item determines answer _____**discrimination power**

56. Identify the acceptable item from given values of item discrimination: _____**0.26**

57. Identify the value of item difficulty for most _____**1.0**

58. A teacher wishes to assess students prior to instruction to measure their current level of skills. What kind of assessment is this teacher using? _____ **Placement**

59. To judge the worth or value of material for a given purpose is known as:___ _____ _____ **Evaluation**

60. Identify the lowest learning level of DOK taxonomy. _____**Extended Thinking**

61. Identify the poor item from given values of item discrimination ._____**17**

Download More Files from VUAnswer.com

62. MCQs cannot be used to assess the knowledge of _____**illusion**
63. To assess achievement at the end of instruction is answers ____**summative assessment**
64. Following are usefullNess of completion questions except: _____**easy to Mark**
65. Which of the following decision determines prerequisite skills, degree of mastery of course goals and mode of learning?_____ **Placement Assessment**
66. The process of obtaining a numerical description of the degree to which an individual processes particular characteristic is called: _____**Placement Assessment**
67. "What individual will do under natural conditions" is the: _____ _____**Typical Performance Assessment**
68. Norm reference assessment and criterion reference assessment are classifications of assessment on the basis of: _____**Method of interpreting results**
69. The way of judging how well students are doing by looking at their work during educational process is called: _____**Assessment**
70. Identify the learning level according to national curriculum that indicates what students will accomplish at the end of each of the five developmental levels in order to meet the standard. _____**Benchmarks**
71. Bring together scientific ideas to form a unique idea is called: _____**Synthesis**
72. In SOLO taxonomy; classify, list, combine are key words for learning level: _____ _____ **Multi- Structural**
73. The assessment determines what individual will do under natural conditions is termed as: _____**Typical Performance Assessment**
74. In DOK taxonomy; create, synthesize, reflect are key words for learning level**:** _____ _____ **Extended Thinking**
75. Ali's writing score placed near the bottom of the class. The interpretation of this statement is: _____**Maximum assessment**
76. Which of the following is the type of assessment by the use of classroom instruction____ _____. **placement**
77. Measure of student's performance against a certain criteria is: _____ _____**Criterion-referenced Assessment**
78. In revised Bloom taxonomy; define, name, select are key words for level: _____ _____**Remembering**
79. **List, enlist, name, define keywords are used in _____ level of Depth of Knowledge.** _____**Recall**
80. **Reliability of a test is necessity but not sufficient condition for:** _____**Validity**
81. **If you have been declared as master of content in specific domain after a test, then your result has been assessed according to: ____Criterion-referenced Assessment**
82. **Student can design a laboratory according to certain specification in which category of objective?** _____**Synthesis**

Quiz no 2

1. -------is used when we want to assess student's prior knowledge _____ _____**Placement assessment**
2. Frequently used tools of summative evaluation are---------? _____**Test**
3. What individual will do under natural conditions is the _____ _____**Typical Performance Assessment**
4. To assess achievement at the end of the instruction is_ _**Summative assessment**
5. **Student's performance is compared with other student's in_____** _____ _____ **Norm referenced test**

Download More Files from VUAnswer.com

6. What is your name? is the example of which level of SOLO taxonom_____ _____ **Pre-structural**
7. The purpose of the evaluation is to make ------ _____**Judgment**
8. Explain the process of water cycle is the question of -------- level in Blooms taxonomy of cognitive domain. _____**Comprehension**
9. ----------determines learning progress, provides feedback to reinforce learning, and correct learning errors _____**Formative Assessment**
10. **Assessment that monitors learning progress ----------** _____ _____**Formative Assessment**
11. The purpose of measurement in education is to ---------- _____ _____ **Collect information for decision making**
12. The summative evaluation is…… _____**Certifying judgment**
13. The purpose of evaluation is to make judgment about educational--------**Quality**
14. Multiple choice question, matching exercise, fill in the blanks and true false is the example of -------_____**Fixed choice assessment**
15. The test made to compare of student with other students is called……..__ _____**Criterion reference Norm reference**
16. List, enlist, name, define keywords are used in …… level of Depth of knowledge _____**Recall**
17. In norm referenced test the comparison is between --------- _____**Individuals**
18. Who was the founder of Pakistan is the question of----------in Blooms taxonomy of cognitive _____**Knowledge**
19. Measurement and evaluation can be used to ……..____ _____**Assess the effectiveness of learning activities**
20. How many domains of Bloom Taxonomy are…….. _____**3**
21. Procedures used to determine person best abilities are……. _____ _____**Maximum performance test**
22. Which of following is the type of assessment by the use of classroom instruction…….. _____**Placement assessment**
23. In …….. Federal government appointed university grant commission _**1976**
24. The type of essay item in which content are limited is:_____ _____ **Restricted response question**
25. Identifying relationship between two things is done in _____**Matching item**
26. How many column matchings items have...? _____**Two**
27. The incorrect option in MCQ are called _____ **Distractor**
28. The item in the column for which a match is sought is called _____ **Premise**
29. The correct option in MCQ is called _____**Answer**
30. Alternative response item is a type of …….. _____**Correct/Incorrect**
31. In which type of question marking will be more reliable? _____ _____**Multiple choice question**
32. In multiple choice items the stem of the items should be _____**Meaningful**
33. A multiple choice question is composed of question referred as _____**Stem**
34. . Instrument used for measuring sample of behavior is? _____**Test**
35. Limited to quantitative description of pupil's performance is? ___ **Measurement**
36. The purpose of the evaluation is to make? _____**Judgment**
37. The purpose of evaluation is to make judgment about educational? ___ **Quality**
38. Evaluation that monitors learning progress is? _____ **Formative evaluation**
39. A formal and systematic procedure of getting information is? _____**Test**
40. The process of obtaining numerical value is? _____**Measurement**
41. A sum of questions is? _____**Test**
42. The first step in measurement is? _____**Decision of what to measure**

Download More Files from VUAnswer.com

43. To assess achievement at the end of instruction is? _**Summative Assessment**
44. Vast of all in scope? _____**Evaluation**
45. The least in scope is? _____**Test**
46. Broader in meaning is? _____**Aims**
47. Procedures used to determine person abilities is? _____
_____**Maximum performance test**
48. In norm referenced test the comparison is between? _____ **Individuals**
49. In which question marking will be more reliable? ___**Multiple choice question**
50. Facility value of less than 0.20 means? _____**Item is difficult**
51. Objective type question have advantage over essay type because such questions? _____**Are easy to mark**
52. Discrimination value of more than 0.4 means _____**Item is acceptable**
53. Test involving the construction of certain patterns are called? _____
_____ **Performance tests**
54. In multiple choice items the stem of the items should be? _____**Meaningful**
55. Which appropriate verb will you use to make an objective behavioral? _____
_____**To construct**
56. Objectives representing the purposes of instruction of a teacher are called?
_____**Instructional**
57. Running description of active behavior of a student as observed by the teacher is? _____**Anecdotal record**
58. A test very popular with class room teacher is? _____**Multiple choices**
59. Frequently used tools of summative evaluation are? _____ **Test**
60. The most commonly used guessing correction formula to predict and control is?
_____**S=R-W/N-1**
61. The summative evaluation is? _____**Certifying judgment**
62. The difference between maximum and minimum values is? _____ **Range**
63. The number of score lying in a class interval is? _____ **Frequencies**
64. A multiple choice question is composed of question referred as? _____ **Stem**
65. In a norm referenced test which item is best? _____ **Item difficulty is near 50**
66. Which question has increasing objectivity of marking? _____
_____**Multiple type questions**
67. The most widely used format on standardized test in USA is? _____
_____**Multiple type questions**
68. Which questions are difficult to mark with reliability? _____**Unstructured essay**
69. Projective techniques are used to measure? _____**Personality**
70. Test meant for prediction on a certain criterion are called? _____**Aptitude test**
71. Kuder Richardson method is used to estimate? _____ **Reliability**
72. Value that divides the data into two equal parts is? _____**Median**
73. The test measures what we intend to measure. This quality of the test is called?
_____**Validity**
74. The length of a test is an important factor in obtaining a representative?_____
_____ _____ **Sample**
75. Median of 1,2,4,5,2,3, is ? _____**3.5**
76. The test made to compare the performance of student with the other students is called? reference _____**Norm reference**
77. The summative evaluation is used? _____**At the end of the program**
78. The appearance of normal curve resembles with? _____**Bell**
79. The alternative name of the “table of spectification” is? _____**Test Blue Print**
80. ”table of specification” helps in? _____**Test development**
81. The supply type test item is? _____**Completion items**

Download More Files from VUAnswer.com

82. Alternative response item is? _____**All above**
83. How many columns matching items have? _____**Two**
84. The item in the column for which a match is sought is? _____**Premise**
85. Identifying relationship between two things is demonstrated by? _____ _____ _ **Matching items**
86. The statement of problem in M.C .Qs is? _____**Stem**
87. The correct option in M.C.Q is? _____ **Answer**
88. The most widely applicable test item is? _____**M.C.Q items**
89. The type of essay item in which contents are limited is? _____ _____**Restricted Response Questions**
90. he list of suggested solutions in MCQs in the form of word, symbol, phrase etc. are called:_____ Alternatives
91. The ability to select organize, integrate and evaluate ideas is demonstrated by? _____**Extended Response Questions**
92. The Analysis of items is necessary in? _____**Standardized Test**
93. Which one is not the type of test of test by purpose? _____**Essay Type Test**
94. The type of the test by method is? _____ **Objective type test**
95. **Student's performance is compared with other students in?** _____ _____ _____**Norm referenced test**
96. Student performance is compared with clearly defined learning tasks in? _____**Criterion reverenced test**
97. Test that measure learning outcome of students is _____**Achievement test**
98. The tests designed to predict future performance is? _____**Aptitude test**
99. The founder of modern intelligent tests was? _____**Alfred Binet**
100. The formula to determine I.Q was presented by? _____ **Stern**
101. I.Q of a student having same physical and mental age will be? _____**100**
102. The I.Q of a student having twelve years mental age and term years physical age will be? _____**120**
103. The quality of test that measures "what it claims to measure" is? _____**Validity**
104. The characteristic of a test to discriminate between high achievers and low achievers is? _____**Differentiability**
105. If the scoring of the test is not effected by any factor, quality of test is called? _____**Objectivity**
106. The quality of test to give same scores when administered at different occasions is? _____ **Reliability**
107. If the sample of the question in the test is sufficiently large enough, the quality of test is? _____**Adequacy**
108. The quality of test showing ease of time, cost, administration and interpretation is called? _____ **Usability**
109. Facility index of an item determines? _____**Ease or difficulty**
110. High and low achievers are sorted out by? _____ **Discrimination power**
111. Test item is acceptable which its faculty index /difficulty level ranges from? _____**30-70 %**
112. Test item is very easy when value of faculty index/ difficulty level is higher than? _____**70 %**
113. Test item is very difficult when value of facility index/ difficulty level is less than? _____**30%**
114. Discrimination power of an item is acceptable when its value ranges from? _____**0.30 – 1**
115. Test item discriminates 100% when its value for discrimination is? _____**1**

116. Test item cannot discriminate low achievers and high achievers when its value is lower than? Download More Files from VUAnswer.com **0.30**

117. Assessment that monitors Learning Progress is ---------_**Formative Assessment**

118. ---------- is used when we want to assess student 's prior knowledge **Placement assessment**

119. Measurement and Evaluation can be used to ---------- _____ **All of Above**

120. The Summative evaluation is ----------_____ **Certifying Judgment**

121. Persistent Difficulties in Learning are Investigated in ---- **Diagnostic Assessment**

122. Who was the founder of Pakistan is the question of ---------- in Bloom taxonomy of cognitive domain? **Knowledge**

123. "What is your name?" is the example of which level of Solo taxonomy ---------- Download More Files from VUAnswer.com **Pre-structural**

124. ---------- determines learning progress, provides feedback to reinforce learning, and correct learning errors **Formative Assessment**

125. The purpose of formative learning is ---------- _____ **Monitoring progress of student**

126. The purpose of measurement in education is to ---------- _____ **Collect information for Decision Making**

127. The purpose of the evaluation is to make ---------- _____**Judgment**

128. "What individual will do under natural condition "is the ----------_____ **Typical Performance Assessment**

129. Multiple choice question, matching exercise, fill in the blanks and true false is the example of ..... **Fixed Choice Assessment**

130. Which of the following is the type of assessment by the use of classroom instruction? --------**Placement Assessment**

131. Curriculum Wing is also known as ---------- ___**National Bureau of Curriculum**

132. List, enlist, name, define keywords are used in --------- level of Depth of Knowledge **Recall**

133. Frequently used tools of summative evaluation are ---------- --. _____**Test**

134. Explain the process of water cycle is the question of --------- level in bloom taxonomy of cognitive domain **Comprehension**

135. Student's Performance is compared with other students in ------- _____ **Norm referenced test**

136. The test made to compare the performance of student with the other students is called ------- --. **Norm reference**

137. To Assess achievement at the end of instruction is -----**Summative Assessment**

138. Permanent difficulties in learning are investigated in---------_____ **Diagnostic Assessment**

139. The purpose of evaluation is to make judgment about educational ---------- --. **Quality**

140. In norm referenced test the comparison is between --------- --._____ **Individuals**

141. How many domains of Bloom Taxonomy ..... ? **3**

142. Procedures used to determine person best abilities are… ..... ? _____ **Maximum performance test**

143. The type of essay item in which content are limited is ---------- ? _____ **Restricted response question**

144. Identifying relationship between two things is done in ---------? ____**Matching item**

145. How many column matchings items have. ?_____ **Two**

146. Cognitive domain of Blooms taxonomy are ------- **6**

Download More Files from VUAnswer.com

147. Identifying the learning level according to national curriculum that indicates what students will accomplish at the end of each of the five development levels in order to meet standards _____**Student Learning outcomes**

148. SOLO stands for: _____**Structure of Observed Learning Outcomes**

149. The process of obtaining a numerical description of the degree to which an individual processes particular characteristic is called**:** _____**Measurement**

150. Measure of student's performance against a certain criteria is **:**_____ _____**Criterion-referenced**

151. The way of judging how well students are doing by looking at their work during educational process is called _____**Evaluation**

152. Hina Exceeded 85% of the sixth graders on the mathematics test. The interpretation of this statement is: _____**Norm Referenced**

153. The Assessment determines what individual can do when performing at their best _____**Maximum Performance Assessment**

154. In DOK taxonomy, create, synthesize, reflect are key words for learning level _____**Extended thinking**

155. Which type of assessment determines the student's prerequisite skills to begin instruction _____ **Placement**

156. Identify the assessment used to find the root cause of continuous or constant problem in students learning_____ **Diagnostic**

157. Cognitive domain was divided into sub subgroups by bloom in _____**1956**

158. Degree to which a test measures an intended content area is called _____ _____**construct validity**

159. Which type of assessment measures entry level performance of student? _____**Placement**

160. The spearman-brown formula is used to estimate: _____**Split-half**

161. To judge the worth or value of material for a given purpose is known as_____ _____ **Evaluation**

162. to use the previous learned material in new situation is called _____**Application**

163. Which of the following decision determines prerequisite skills, degree of mastery of course goals and modes of learning?_____ **Placement**

164. Test designed to measure the number of items and individual can attempt correctly in a given time is referred as: _____**Speed**

165. The degree to which an instrument measures what it is supposed to measured is called _____**Validity**

166. Monitoring the outcomes with reference to the objectives is called: ___**Evaluation**

167. Which type of Assessment is designed to measure whether the students have mastered the skills presented in instruction? _____**Summative Assessment**

168. Which type of validity involves testing a group of subjects for a certain construct, and the comparing them with results obtained at some point in the future?_____ _____ **Predictive validity**

169. How well the sample of assessment tasks represents the domain of the tasks to be measured is known as _____**Content validity**

170. Reliability over different forms of assessment is known as _____**Equivalence**

171. Which of the following provides a framework of categories with different hierarchical levels of outcomes? _____**Taxonomy**

172. The procedure used to gain information about student learning and the formation of value judgement concerning learning progress is called:_____ **Evaluation**

173. Reliability within an assessment itself is known as: _____**Internal Consistency**

Download More Files from VUAnswer.com

174. ---------- refers to the assessment of participants, and summarize their development at the particular time _____**Summative assessment**

175. In SOLO taxonomy; identify, memorize, are keywords for learning level: _____**Uni-Structural**

176. Student creatively applies knowledge to construct an overall theory called: _____**Application**

177. A type of assessment done at the end of a course where the focus is on the learners receiving a grade for their work rather than receiving feedback on their progress is called _____**Summative assessment**

178. "What is your country name?" is example of which level of SOLO taxonomy._____ _____ **Pre-structural**

179. The domain which deals with physical abilities and coordination objectives is _____**Psychomotor**

180. The Degree to which a test measures an intended hypothetical construct is called: _____**Construct validity**

181. What kind of assessment, permanent difficulties in learning are investigated ? _____**Diagnostic**

182. The alternative name of the "table of specification" is ? _____**Test Blue Print**

183. Which of the following are three domains of Bloom's Taxonomy? _____ _____**Cognitive, affective, psychomotor**

184. Measure of performance which interprets an individual's relative standing in some known group is: _____**Norm referenced**

185. Which of following learning level of DOK taxonomy demands extended use of higher order thinking processes such as synthesis, reflection, assessment _____ _____ **Extended thinking**

186. Identify the lowest learning level of SOLO taxonomy._____ **Pre-Structural**

187. The purpose of the evaluation is to make judgment about educational ------ **Quality**

188. The quality of test to give same scores when administered at different occasions is called _____ **Reliability**

189. To assess achievement at the end of instruction is ---------- _____ _____**Summative assessment**

190. Which of the following test with time limit strict that no one is supposed to complete all items? _____**Speed**

191. In what kind of assessment, recurrent difficulties in learning are investigated? _____**Diagnostic Assessment**

192. In SOLO taxonomy: classify, list, combine are keywords for learning level: _____ _____ _____ **Multi- Structural**

193. The Scores of students in papers are _____**Assessment**

194. Aptitude test is example of which type of Assessment __**Placement Assessment**

195. In DOK taxonomy; argue, critique, formulate are key words for learning level: _____**Strategic Thinking**

196. Which interpretation of test results is most meaningful when test designed to measure a set of clearly stated learning tasks?_____ **Criterion referenced**

197. A tool used to interpret and grade students' work against criteria and standards._____ **Marking**

198. A scoring tool that teachers use to assess student learning._____ **Test**

199. The pattern of students' errors and misconceptions start emerging after reading a few papers that help to develop scoring rubric for extended response items which will applied to score all papers. This type of scoring is called:____**Analytic scoring**

200. Zero in raw score indicates:_____ **No absence of trait**

Download More Files from VUAnswer.com

201. A system of number or values assigned to work is known as:_____ _____ **Level of performance**

Quiz# 2

1. A list of three or more choices from which an examinee is required to choose the correct one is given in _____**MCQ items**
2. Identify the value of item difficulty for most difficult test item. _____**0.7**
3. Under Item Response Theory, the difficulty of an item describes where the item functions along the _____**ability scale**
4. Identify the excellent item from given values of item discrimination. _____ _____**0.50**
5. Identify the value of item difficulty for easiest test item. _____**0.2**
6. The correct option in MCQ is called _____**Answer**
7. Limited numbers of content area is needed to be tested in _____ _____**Restricted response items**
8. Discrimination power of an item is acceptable when its value ranges from? _____ _____**0.5-1.**
9. Identify the value of item difficulty for easiest test item. _____**0.1**
10. The ability to select organize, integrate and evaluate ideas is demonstrated by _____**Extended Response Questions**
11. The pattern of student's errors and misconceptions start emerging after reading a few papers that help to develop scoring rubric for extended response items which will applied to score all papers. _____**Analytic scoring**
12. MCQs can Not be used to assess the knowledge of _____**Illusions**
13. Identify the acceptable item from given values of item discrimination. _____**0.26**
14. Feature, trait or dimension to be measured is decided in the _..... _ element of rubric._____ **criteria**
15. Discrimination power of an item is acceptable when its value ranges from? _____**0.2-1.00**
16. Good distractor is that _____**Attracts equally low and high achievers**
17. Identify the value of item difficulty for most difficult test item. _____**0.2**
18. Identify the poor item from given values of item discrimination _____**0.15**
19. Identify the value of item difficulty for most difficult test item. _____**0.7**
20. Identify the value of item difficulty for most difficult test item. _____**0.8**
21. How many columns matching items have _____**Two**
22. The most commonly used guessing correction formula to predict and control is_____ **S=R-W/N-1**
23. Descriptors are _____**detailed description for each level of performance**
24. A test very popular with class room teacher is _____**MCQ test**
25. Guessing parameter can be measured in _____**IRT**
26. When you are ready to package and reproduce the test _____ _____ _____-_____**Objectives and test blue print**
27. There are __ __ main types of rubric _____ **two**
28. The correction option in MCQ is called _____**Distractor**
29. Item analysis focuses to find out _____**All above**
30. 40 words per minute _____**Criterion Referenced interpretation.**
31. A rubric is a __**used for assessing a particular type of work or performance.**
32. Bad distractor is that:___ **Attracts high achievers more than lower achievers.**
33. MCQs cannot be used to assess the knowledge of _____**Illusions**

Download More Files from VUAnswer.com

34. Objective type question have advantage over essay because ____**Easy to mark**
35. The problem stated in the form of question or incomplete statement in MCQs is _____**Stem**
36. Identify the poor item from given values of item discrimination _____**0.15**
37. Which of the following are major threats to the validity of test score interpretations _____**Teacher's behavior**
38. Identify the item difficulty _____**0.3**
39. Identify the excellent item from given values of item discrimination _____**0.48'**
40. Difficulty level of an item determines _____**All above**
41. A MCQ is composed of statement referred as _____ **Stem**
42. Holistic rubric are good for evaluating _____**Overall performance**
43. Good item _____**0.31''**
44. Most difficult test item _____**Lowest 0.1**
45. An item in column for which a match is sought _____--**Premise**
46. Most difficult _____**0.2**
47. Limited number of content area is needed to be tested in _____ _____**Restricted response items**
48. A steeper item characteristics curve shows that item is ____**Good discriminator**
49. Test item cannot discriminate low achievers and high achievers when value lower than_____ **0.3**
50. Example of supply response item _____**Short answer**
51. Not basic step to design rubric _____--___**Give hints to a student directly**
52. Schools professional practices _____**Material kept in a locked storage**
53. Identifying relationship between two thing _____**Matching items**
54. True-false items are also known as _____**Alternative form items**
55. Advantages of holistic rubric involves _____**All above**
56. MCq cannot be used to measure the knowledge of _____**Illusions**
57. To give an answer or make a judgement about something without being sure of all facts _____**Guess**
58. Extended response items are helpful to determine _____ _____**Length and complexity of response**

1. alternative form assessment?

Alternative assessments, also referred to as performance tests or authentic assessments, are used to determine what students can and cannot do, in contrast to what they do or do not know. In other words, an alternative assessment measures applied proficiency more than it measures knowledge. Typical examples of alternative assessments include portfolios, project work, and other activities requiring some type of rubric.

What follows are some examples of alternative assessments that you might consider.

- Open book exams. ...
- Crib Sheets. ...
- Take home exams. ...
- Collaborative testing. ...
- Student portfolios. ...
- Performance Tests. ...
- Retake policies. ...
- Adding the option of explanation to an M-C test.

Advantages of Alternative Assessments

- They provide a means of assessing valued skills that cannot be directly assessed with traditional tests.
- They provide a more realistic setting for student performance than traditional tests.
- They focus on student performance and the quality of work performed by students.
- They can be easily aligned with established learning outcomes.
- Process can be costly in terms of time, effort, equipment, materials, facilities, or funds.

Download More Files from VUAnswer.com

2. construct validity?

Meaning

How well a test measures up to its claims. A test designed to measure *depression* must only measure that particular construct, not closely related ideals such as *anxiety* or *stress.*

Procedure

Develop a test framework;

1) Defining construct,
2) Identifying sub-constructs,
3) Listing indicators of each sub-construct
4) Writing test items for each indicator

Method

There are two methods to confirm construct validity of a test.

Expert judgment

1) There are experts of the field. For above example, people who are expert in essay writing will be considered to assess the construct validity of the table and table will be revised under their guidance.

2) Factor analysis

Criterion-referenced Assessment

Criterion-referenced assessment describes student performance according to a specific domain of clearly defined learning tasks e.g. adds single-digit while numbers. In this you don‘t compare student performance with other students rather you compare the performance of all students with criteria (in our case that criteria are our learning outcomes). It is most commonly used in schools to report the achievement of learning outcomes against set goals rather than other students. It grades the students to pre-defined criteria and student‘s grades represent their mastery over content. Students with same level of expertise achieve same level of grades. A cut point is determined to distinguish between failed and successful students regardless of score of highest and lowest achiever. It consists of teacher-made tests, custom made tests from the test publishers and observational techniques.

Function of formative assessment

Functions of Formative Assessment

1) Focus of measurement in formative assessment is predefined segment of instruction.
2) Limited sample of learning tasks are addressed.
3) The difficulty of item varies with each segment of instruction.
4) Formative assessment is conducted periodically during the instructional process.
5) Results of formative assessment are used to improve and direct learning through ongoing feedback.

3 suggestion for good assessment 3

A good assessment should **be reliable, valid, and free of bias**. First, reliability refers to the consistency of students' scores; that is, an assessment is reliable when it produces stable and consistent results. Reliability can come in two major forms: (1) stability and (2) alternate form reliability.

A quality assessment in education consists of four elements - reliability, standardization, validity and practicality

A good assessment is supposed to show what we have truly learned. There are four qualities of good assessments. Educators should ensure these qualities are met before assessing students. They are:

- Reliability
- Standardization
- Validity
- Practicality

function of blue print test

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

1) It helps a teacher in allotting the questions to different content areas and Bloom's learning categories in a systematic manner.
2) 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)
3) 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).

In developing a test blueprint first of all it is necessary to select some learning. Objectives and among this list of learning objectives some objectives are more important in sense that more time of instruction is spent on them while some other are less important in terms of time spent on them in classroom so in developing table of specification balance among these learning objectives is important, for this purpose we need to weigh the learning objectives for calculating their relative weightage in test.

The test blueprint is used to guide and target **item writing** as well as for test form assembly. Use of a test blueprint improves consistency across test forms as well as helping ensure that the goals and plans for the test are met in each operational test. An example of a test blueprint is provided next.

In this post, I discuss the benefits of blueprints.

- Create a shared understanding of the bigger picture.
- Understand what different teams are working on.
- Enable better communication and outcomes.
- Help identify and explain priorities.
- Stay focused by zooming in and out.
- Conclusion.

Download More Files from VUAnswer.com

learning student outcomes in national curriculum

Student learning outcome:

1) Articulate, identify and differentiate between the sounds of individual letters, digraphs and trigraphs in initial and final positions in a word.
2) Identify paragraph as a graphical unit of expansion, ¾ know that words in a sentence join to make sense in relation to each other.

Multiple choice questions

The MCQs are the most versatile type of test item available. It can measure variety of learning outcomes from simple to complex, and it is adaptable to the most type of matter content.

Despite the wide applicability of the MCQs there are SLOs, such as ability to organize and present ideas that cannot be measured with any form of selection item. For such skills constructed and restricted response questions are used.

A multiple-choice question (MCQ) is composed of two parts: **a stem that identifies the question or problem**, and a set of alternatives or possible answers that contain a key that is the best answer to the question, and a number of distractors that are plausible but incorrect answers to the question.

Essay type question

Essay question are defined as a test item which require a response composed by the examinee, usually in the form of one or more sentences , of a nature that no single response or pattern of

response can be listed as correct , and the accuracy and quality of which can be judged subjectively only by one skilled or informed in the subject.

writting essay type question steps.

1. Step 1: Prewriting. Before you start writing, you need to decide exactly what you'll write about and do the necessary research. ...
2. Step 2: Planning and outlining. ...
3. Step 3: Writing a first draft. ...
4. Step 4: Redrafting and revising. ...
5. Step 5: Editing and proofreading.

Advantages of essay type question.

- Test takers can elaborate and provide detailed answers.
- Test takers are not able to guess and select an answer.
- Can review individualized responses from each user.
- Can be used for all types of subjects.
- Takes less time to create questions.

guidelines for essay type questions?(5)

Following guidelines are for writing essay type items when developing a test.

1) Frame questions so that the examinee‘s task is explicitly defined.
2) Specify the value and an approximate time limit for each question.
3) Do not employ optional questions.
4) Employ a larger number of questions that require relatively short answers rather than only a few questions that require long answers.
5) Verify a question‘s quality by writing a trial response to the question.
6) Prepare a tentative scoring key in advance of considering examinee responses.
7) Score all answers to one question before scoring the next question.
8) Make prior decisions regarding treatment of such factors as spelling and punctuation.
9) Evaluate essay responses anonymously.

Q: State 5 rules for constructing alternative form questions.

ANS: 1. Avoid broad general statements if they are to be judged true or false.

2. Avoid trivial statements

3. Avoid the use of negative especially double negative statements.

4. Avoid to include two ideas in one statement unless cause and effect relationship are being measured

5. Avoid long complex sentences

Steps of rubric

Basic Steps to design Rubric

1. Identify a learning goal.
2. Choose outcomes that may be measured.
3. Develop or adapt existing rubric.
4. Share it with students.

Download More Files from VUAnswer.com

2.formula of instructional time of table of specification.

Step 1: Instruction Time

To do the calculations for the instruction time for columns of the table of specifications the teacher must use the following formulas for each objective in the table.

Time in class spent on objective (min) / total time for the instruction being examined (min) Percentage of instruction time

$$\text{Percentage of instruction time} = \frac{time\ spent\ on objective, content, theme\ (min)}{Total\ time\ for the\ instruction\ being\ examined\ (min)}$$

$$\text{Percentage of instruction time} = \frac{250}{1000}$$

Percentage of instruction time = 25%

3.what is test-retest. Write two advantage and two disadvantages.

Test-Retest method is a measure of stability

Test-retest reliability measures the consistency of results when you repeat the same test on the same sample at a different point in time. You use it when you are measuring something that you expect to stay constant in your sample.

Arranging test items from easy to hard

1. Increases the possibility of good start for poor starters
2. Build confidence
3. Motivates
4. Reduces test anxiety

5- advantage of writing test

Exams help a lot to bring an improvement in the individual's knowledge because they provide regular feedback to the students who acknowledge their shortcomings and work on them. Another advantage is that exams promote competition among students. They work harder to improve their knowledge and skills.

test security?

Test security is the process of protecting assessments, so that the results from those assessments can be trusted and used to make important decisions about individual competence. According to the American Education Research Association/American Psychological Association/National Council on Measurement in Education standards, test users have the responsibility to protect the security of tests. Test security is particularly important when assessments are used, for example, in selection, accountability, credentialing, or diagnosis because of the inferences drawn from the assessment's validity.

3)guessing technique for appraisal peer?

Peer appraisal **Peer appraisal**

In this procedure students rate their peers on the same rating device used by their teacher.

It depends on greatly simplified procedures.

Guess who technique

In this technique teacher use a positive or negative behavior of a student as a example. Other students from same group try to guess the statement with that characteristic correctly. Generally

Download More Files from VUAnswer.com

behaviors used for example are positive in nature to avoid any adverse effect on the student pointed out in the example.

The guess who technique is based on nomination method of obtaining peer ratings and is scored by simply counting the number of mentions each students receive on each description.

4)advantages of MCQs?

Advantages of MCQs

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

7)steps to starting general objectives?(5)

Steps for stating General Objectives

1) State each general objective as an intended learning outcome
2) Begin each general objective with a verb
3) State each general objective to include only one general learning outcome
4) State each general objective at the proper level of generality
5) Keep each general objective sufficiently free of course content so it can be used with various units of study
6) Minimize the overlap with other objectives

How parent teacher meeting conduct

5 Steps to Conducting Successful Parent-Teacher Conferences
Parent-teacher conferences are held both for discussing students' academic progress, and for building and strengthening parent-teacher partnership to achieve academic goals. Here are 5 steps that teachers can follow to execute smooth and productive parent-teacher conferences:
Keep Parents Posted: Ideally, teachers and parents should communicate before the scheduled conference date. Parents must be informed of any issues that need to be discussed – this way, the conference serves as a brainstorming session for finding solutions to improve student performance in class.
Present Evidence: Teachers can create portfolios for each student recording behaviors, class participation, academic progress, achievements, and areas that need development. At conferences, these portfolios can help parents gain a clear and complete picture about their child's academic progress.
Be Purposeful: Sometimes, the discussion may move away from the student or fixate on a particular point. Teachers can try to avoid this roadblock by forming an agenda for the conference, clearly listing the points that need to be discussed. A copy of the agenda can be shared with parents either before or during the conference and it can help teachers direct parents' attention back to the necessary points.
When scheduling conference timing, teachers can also plan additional time for miscellaneous topics, or for parents to ask questions or voice concerns. If needed, teachers can schedule an additional meeting to discuss these issues.
Be Positive: Remember to start and end the conference on a positive note. Highlight achievements and progress before discussing the areas that require improvements. Teachers can collaborate with parents and find suitable solutions to help the students with their problems. End the meeting on a positive note, as this can give parents hope and motivate them to help their child.

Define assessment

Download More Files from VUAnswer.com

Assessment is **the systematic basis for making inferences about the learning and development of students**. It is the process of defining, selecting, designing, collecting, analyzing, interpreting, and using information to increase students' learning and development.

Nature of reliability

1. Nature of reliability.

Reliability refers to the consistency of measurement

1) Reliability refers to the results obtained with an assessment instrument and not to the instrument itself.
2) An estimate of reliability always refers to particular type of consistency (stability, equivalence, internal consistency)
3) Reliability is necessary but not sufficient condition for validity.
4) Reliability is primarily statistical (range +1 and -1).

Method of assessment

1) Placement Assessment.
2) Formative Assessment
3) Diagnostic Assessment
4) Summative Assessment

Benefits of rubric scoring.

- Help clarify vague, fuzzy goals.
- Help students understand your expectations.
- Help students self-improve.
- Inspire better student performance.
- Make scoring easier and faster.
- Make scoring more accurate, unbiased, and consistent.
- Improve feedback to students.
- Reduce arguments with students.

Q: Restricted response type questions with example

ANS: 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 of 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. Some of which are;

1. Analysis of relationship.

2. Compare and contrast positions.

3. Explain cause-effect relationship.

4. Organize data and support a viewpoint.

5. Formulate hypotheses.

6. Point out strengths and weaknesses.

7. Integrate data from various resources.

Elements of Rubric

A rubric includes:

1. Score
2. Criteria
3. Level of performance
4. Descriptors

Types of written test 2

Types of Written Tests

- ✓ Verbal
- ✓ Non-verbal
- ✓ Objective
- ✓ Subjective
- ✓ Teacher Made
- ✓ Standardized
- ✓ Power
- ✓ Speed

Download More Files from VUAnswer.com

1_Solo stands for? Names of solo taxonomy?

Structure of Observed Learning Outcomes (SOLO)

1.Pre-structural

Students are simply able to acquire bits of unconnected information and respond to a question in meaningless way. Example of pre-structural level:

Question: What is your name?

Answer: What is your name?

Uni Structural

Student shows concrete understanding of the topic. But at this level is only able to respond one relevant element from the stimuli or item that is provided.

Indicative verbs: identify, memorize, do simple procedure

Multi- Structural

Student can understand several components but the understanding of each remains discreet. A number of connections are made but the significance of the whole is not determined. Ideas and concepts around an issue are disorganized and aren't related together.

Indicative verbs: enumerate, classify, describe, list, combine, do algorithms.

Relational

Student can indicate connection between facts and theory, action and purpose. Shows understanding of several components which are integrated conceptually showing understanding of how the parts contribute to the whole. Indicative verbs: compare/contrast, explain causes, integrate, analyze, relate, and apply.

Extended Abstract

Student at this level is able to think hypothetically and can synthesize a material logically. Student make connections not only with in the given subject area but understanding is transferable and generalizable to different areas. Indicative verbs: theorize, generalize, hypothesize, reflect, generate

Note on types of written test 5

Types of Written Tests

- ✓ Verbal
- ✓ Non-verbal
- ✓ Objective
- ✓ Subjective
- ✓ Teacher Made
- ✓ Standardized
- ✓ Power
- ✓ Speed

Verbal

Emphasize reading, writing, or speaking. Most tests in education are verbal tests.

Non-verbal

Does not require reading, writing or speaking ability, tests composed of numerals or drawings is example.

Objective

Refers to scoring of tests when two or more scorers can easily agree on whether the answer is correct or incorrect, the test is objective one. True false, multiple choice and matching tests are example

Subjective

Also refers to scoring. When it is difficult for two scorers to agree on whether an item is correct or incorrect, the test is a subjective one. Essay tests are the example.

Teacher Made

Constructed solely by teacher only to be used in his/her own classroom. This type of test is custom designed according to need and issues related to specific class

Standardized

Test constructed by measurement experts over a period of years. They are designed to measure broad national objectives and have a uniform set of instructions that are adhered to during each administration

Most also have tables of norms, to which a student performance may be compared to determine where the student stands in relation to a national sample of students at same level of age or grade

Power

Tests with liberal time limits that allow each student to attempt each item. Item tend to be difficult.

Speed

Tests with time limits so strict that no one is expected to complete all items. Items tend to be easy.

Difference between summative and formative assessment

Formative Assessment

Formative assessment provides feedback and information during the instructional process, while learning is taking place, and while learning is occurring. Formative assessment measures student progress but it can also assess your own progress as an instructor.

Types of Formative Assessment

1) Observations during in-class activities; of student‘s non-verbal feedback during lecture.

Download More Files from VUAnswer.com

2) Homework exercises as review for exams and class discussions)
3) Reflections journals that are reviewed periodically during the semester
4) Question and answer sessions, both formal (planned) and informal (spontaneous)
5) Conferences between the instructor and student at various points in the semester
6) In-class activities where students informally present their results
7) Student feedback collected by periodically answering specific question about the instruction and their self-evaluation of performance and progress

Functions of Formative Assessment

6) Focus of measurement in formative assessment is predefined segment of instruction.
7) Limited sample of learning tasks are addressed.
8) The difficulty of item varies with each segment of instruction.
9) Formative assessment is conducted periodically during the instructional process.
10) Results of formative assessment are used to improve and direct learning through ongoing feedback.

Summative Assessment

Summative assessment takes place after the learning has been completed and provides information and feedback that sums up the teaching and learning process. Typically, no more formal learning is taking place at this stage, other than incidental learning which might take place through the completion of projects and assignments.

Summative assessment is more product-oriented and assesses the final product, whereas formative assessment focuses on the process toward completing the product. Once the project is completed, no further revisions can be made.

If, students are allowed to make revisions, the assessment becomes formative.

Types of Summative Assessment

- ✓ Examinations (major, high-stakes exams)
- ✓ Final examination (a truly summative assessment)
- ✓ Term papers (drafts submitted during the semester would be a formative assessment)
- ✓ Projects (project phases submitted at various completion points could be formatively assessed)
- ✓ Portfolios (could also be assessed during its development as a formative assessment)
- ✓ Performances

Functions of Summative Assessment

1) Focus of measurement in summative assessment is on course or unit objectives.
2) Broad sample of all objectives is used in summative assessment.
3) This type of assessment uses wide range of difficulty while selecting items for the test.
4) Summative assessment is done at the end of the unit or the course
5) Most important functions of summative assessment is to assign grade, certification of accomplishment and evaluation of teaching

---

5.Suggestions for scoring constructed response questions. 5 marks

Suggestions for scoring constructed response questions

1. Prepare an outline of the expected answer in advance.
2. Use the scoring rubric that is most appropriate.
3. Decide how to handle factors that are irrelevant to the learning outcomes being measured.

4. Evaluate all responses to one question before going on to the next one.
5. When possible, evaluate the answers without looking at the student‘s name.
6. If especially important decisions are to be based on the results, obtain two or more independent ratings.

6.steps for stating general objectives. 5 marks.

Steps for stating General Objectives

7) State each general objective as an intended learning outcome
8) Begin each general objective with a verb
9) State each general objective to include only one general learning outcome
10) State each general objective at the proper level of generality
11) Keep each general objective sufficiently free of course content so it can be used with various units of study
12) Minimize the overlap with other objectives

Derived score 2

Derived Scores

A derived score is a numerical report of test performance on a score scale that has welldefined characteristics and yields normative meaning.

Holistic rubric 3

Holistic Scoring Rubric

These rubrics are good for evaluating overall performance on a task. All criteria are assessed as a single score. As only one score is given so holistic rubrics are easy to score.

When to use:

1. There is no correct answer/response to ask a task e.g. creative work.
2. Focus is on overall quality, proficiency or understanding of a specific content or skill.
3. The assessment is summative e.g. at the end of the semester or major.
4. Assessing significant numbers e.g. 150 student portfolio.

5 recommendations for effective HST 5

Recommendations for effective High Stake Testing

- Protection against high stake decisions based on single test
- Adequate resources and opportunity to learn
- Validation for each intended separate use
- Full disclosure of likely consequences
- Alignment between test and curriculum
- Validity of passing scores and achievement levels
- Appropriate attention towards language difference between examinees
- Appropriate attention towards examinees with disabilities
- Careful adherence to explicit rules for determining which students are to be tested
- Ongoing evaluation for intended and unintended effects of high stake testing

State the steps of specific learning objectives 5

Download More Files from VUAnswer.com

Specific Learning Outcomes

Each General objective must be defined by a sample of specific learning outcome to clarify how students can demonstrate that they have achieved general objective. Until the general objective are further defined in this manner they will not provide adequate direction for assessment

Steps for Stating Specific Outcomes

1) List beneath each general objective a representative sample of specific learning outcome that describes terminal performance students are expected to demonstrate
2) Begin each specific learning outcome with an action verb that specifies observable performance
3) Make sure that each specific learning outcome is relevant to the general objective it describes
4) Include enough SLOs to describe adequately the performances of students who have attained the objectives.
5) Keep the SLOs sufficiently free of course content so that the list can be used with various units of study
6) Consult reference materials for the specific components of those complex outcomes that are difficult to define
7) Add a third level of specificity to the list of outcomes, if needed

What is Mcqs?2

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:

A Stem

No. of options or alternatives

Stem

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

SOLO stands for and write its levels? 2

Structure of Observed Learning Outcomes (SOLO)

Levels of Structure of Observed Learning Outcomes (SOLO)

1) Pre-structural
2) Uni-structural
3) Multi-structural
4) Relational
5) Extended Abstract

Norm-referenced Assessment 5 marks

Norm-referenced Assessment measures the students‘ performance according to relative position in some known group. E.g. ranks tenth in classroom group of 50 or top 5 students in a class. Relative

Download More Files from VUAnswer.com

position means a point defined with reference to another position (where a student stands compared to other students)

Rather than reporting student‘s achievement it reports students standing among other students.

NRT is utilized to discriminate between a certain groups of students. It‘s never used for certification or issuing grades to students. Position of the student is generally represented by percentile scarce noting the percentage of students achieving same or low scarce in the test. Examples of NRT are NTS or CSS exams. E.g. if you achieved 97% percentile on NTS test, it means there are 96% who scored lower then you.

Norm referenced test includes items with average difficulty and high discriminating power. This provides large spread of scores which makes it easy to declare the relative positions of students. Because the purpose of NRT is not to certify grades, so the test made for it must be of average difficulty means test items should not be very easy or very tough. If all the items are difficult then no student will be able to solve it and we are not able to discriminate who is good students. And if all the items are easy then even the low ability students can solve it then we are not able to discriminate it.

Norm referenced test includes standardized aptitude and achievement tests, teacher-made survey tests, interest inventories, adjustment inventories

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.

Nature of validity-----5

Following are different points that describe the nature of validity.

1. Appropriateness of the interpretation of the results

Validity is referred as validity of test‖ but it is in fact validity of the interpretation and use to be made of the results.

2. Validity is matter of degree

It does not exist on all or none basis. It is best considered in term of categories that specify degree, such as high, moderate or low validity

3. Validity is specific to some particular use or interpretation

No assessment is valid for all purposes. An arithmetic test may have high degree of validity for computational skill and low degree for arithmetical reasoning

4. Validity is a unitary concept

Validity does not have different types. It is viewed as a unitary concept based on different kind of evidences

5. Validity involves an overall evaluative judgment

Portfolio

Systematic collection of students work into portfolios can serve a variety of instructional and assessment purposes. The value of portfolios depend heavily on the clarity of purpose the guidelines for the inclusion of materials, and the criteria to be used in evaluating portfolio.

Portfolio of students work

A portfolio is collection of student work selected to serve a particular purpose such as documentation of student growth. It is purposeful; collection of pieces of student‘s work.

Key steps in defining and using portfolios

1. Specify purpose.
2. Provide guidelines for selecting portfolios.
3. Define student‘s role in selection and self-evaluation.
4. Specify evaluation criteria.
5. Use portfolios in instruction and communication.

Strengths of portfolios

1) The can be readily integrated with the instruction.
2) Provide opportunity to student‘s to show what they can do.
3) Encourage to become reflective learner.
4) Help in setting goal and self-evaluation
5) Help teacher and student to collaborate and reflect on student‘s progress.
6) Effective way to communicate with parents.
7) Provide mechanism for student centered and student-directed conferences with parents.
8) Provide concrete examples of students development and current skills.

Weaknesses of portfolios

1) Can be time consuming to assemble.
2) Hard to use in summative assessment.
3) Difficult to compare results.
4) Very low reliability.

Evaluation 2 marks

Evaluation is process of making a value judgment against intended learning outcomes and behavior, to decide quality and extent of learning. Evaluation is always related to your purpose, you aligned your purpose of teaching with what students achieved at the end, with their quality and quantity of learning

Download More Files from VUAnswer.com

Classical Test Theory (CTT) and Item Analysis 5

Classical Test Theory (CTT) has relatively weak theoretical assumptions, which make it easy to apply in many testing situations. Relatively weak theoretical assumptions not only characterize CTT but also its extensions (e.g., generalizability theory).

CTT‘s major focus is on test-level information, item statistics (i.e., item difficulty and item discrimination) are also an important part of the CTT model.

Guidelines for essay type items

- Following guidelines are for writing essay type items when developing a test.

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

Reliability 3 marks

Reliability

Reliability refers to the consistency of assessment results.

Reliability refers to the results obtained with an assessment instrument and not to the instrument itself.

An estimate of reliability always refers to particular type of consistency (stability, equivalence, internal consistency)

Reliability is necessary but not sufficient condition for validity.

Reliability is primarily statistical (range +1 and -1).

Methods of reliability

Test-Retest (stability)

Equivalent Forms (equivalence)

Test- Retest with Equivalent Forms (stability and equivalence)

Split Half (Internal Consistency)

Kuder- Richardson(Internal Consistency)

Cronbach Alpha (Internal Consistency)

Inter-rater Reliability (consistency of rating)

Validity

Validity is an evaluation of adequacy and appropriateness of the interpretation and uses of results. It determines if a test is measuring what it intended to measure.

Reliability vs Validity

Reliability of measurement is needed to obtain the valid results, but we can have reliability without validity. Reliability is necessity but not sufficient condition for validity

Q: analytic rubric? In which conditions it is used?

ANS: 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.

When to use:

1. Several faculty are collectively assessing student work.

2. Outside audience will be examining rubric scores.

3. Profiles of specific strength/weakness are desired.

Q: Extended response type questions with example

ANS: 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.

Q: assembling or packaging the test

ANS: Assembling 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

Assembling Test (Packaging the Test)

1. Grouping together items of similar format

2. Arranging test items from easy to hard

3. Properly spacing items

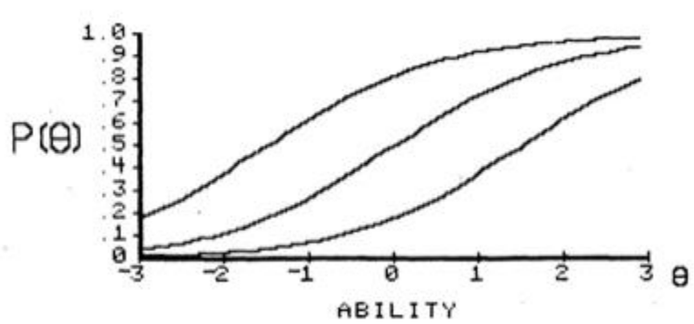
4. Keeping items and options on the same page

5. Placing illustrations near the descriptions

6. Checking the answer key

Q: What is Item difficulty in IRT? Explain with graph.

ANS: 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).



Score

A system of number or values assigned to a work often combines with a level of performance. High numbers are for best performance like 4, 5 or 6 whereas down to 1 or 0 are the lowest score in a performance assessment.

Criteria

It tells us that which feature, trait or dimension to be measured and include a definition and example to make clear the meaning of each trait to be assessed.

Level of Performance

There are adjectives to describe the performance levels. These levels tell students what they are expected to do.

Descriptors can be used with levels of performance to achieve objectivity but they can be used without them as well.

Descriptors

These are details for each level of performance to reliable and unbiased scoring.

Formative Assessment

Formative Assessment determines learning progress, provides feedback to reinforce learning, and correct learning errors. When we assess student during classroom instruction with a purpose to have a feedback that how can we make our teacher learning process better, that is formative assessment. In this assessment, we are not assessing what students learnt or not rather we assess the process behind the students learning .The process behind the student learning includes teaching method , book, . If we make all these things according to the needs of students then learning will improve.

Summative Assessment----5

Summative assessment comes at the end of the instructional session (course of unit). It is designed to measure the extend of achievement of intended learning outcomes. The primary utility of this type of assessment is to assign grades and certifying the level of mastery and expertise in certain subject. It is usually done through teacher made achievement tests or alternative assessment techniques like portfolio to summarize the overall performance of the student at the end of the session. It is not compulsory to done it at the end of the semester, in semester system there is midterm and final term in one semester, and these both are summative assessment.

Sociometric technique

Sociometric technique is method for assessing the social acceptance of individual students and the social structure group. It is based on students choice of companion for some group situation.

This form was used to measure student‘s acceptance as seating companions, work companions and play companions.

There are few important principles of sociometric choosing

71

1. The choices should be real choices that are the natural part of classroom activities.
2. The basis for the choice and restriction on the choosing should be made clear.
3. All students should be equally free to participate in the activity or situation.

4. The choice of each student make must be kept confidential.

5. The choices should actually be used to organize or rearrange the group

Name of Psychomotor Domain

i. Perception ii. Set

iii. Guided Response iv.

Mechanism

v. Complex covert Response vi.

Adaption vii. Origination

Theory (IRT) and Item Analysis

Item Response Theory‖ (IRT) presents a model for expressing the association between an individual's response to an item and the underlying latent variable (often called "ability" or "trait")
The latent variable, expressed as theta (θ), is a continuous one-dimensional construct that explains the covariance among item responses
People at higher levels of latent trait have a higher probability of responding correctly or endorsing an item.

High stake testing

The use of test and assessment to make decisions that are of prominent educational, financial, or social impact.

Decision based on high stake testing

Promotion to next grade

Awarding of diploma or degree

Evaluation of school performance

Incentives and accountability of school staff

Holistic

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.

Bloom taxonomy

Bloom’s Taxonomy of Learning Objectives Bloom‘s Taxonomy was presented by Benjamin Bloom in the 1956, consists of a framework with most common objectives of classroom instruction.

Bloom’s Taxonomy of Learning Objectives

Those dealing in three different domains and further sub categories in these domains.

Cognitive

Affective

Psychomotor

Cognitive Domain:

Knowledge

Comprehension

Application

Analysis

Synthesis

Evaluation

Affective Domain

Receiving

Responding

Valuing

Organization

Characterization

Psychomotor Domain

Perception

Set

Guided Response

Mechanism

Complex covert Response

Adaption

Origination

High Stake Testing-----5

Why does it matter?

The use of test and assessment to make decisions that are of prominent educational, financial, or social impact.

Decision based on high stake testing

1. Promotion to next grade

2. Awarding of diploma or degree

3. Evaluation of school performance

4. Incentives and accountability of school staff

High Stake Testing in Pakistan

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

1. Grade 5 and 8 examination by Punjab Examination Commission.

2. Grade 10 and 12 examination by BISEs

Test construction in High Stake testing

1. The format of items are same as used in CRT in classroom

2. Item of developed by professional employed in dedicated organization

Item banks are developed

4. Psychometric properties of item are tested

5. The process goes on round the year

Criticism on high stake testing

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

Example of knowledge comprehension and application level

Knowledge:

It is defined as the remembering of previously learned material. This may involve the recall of a wide range of facts, procedures principals and generals, the recall of procedures and the processes.

Sample Question: Define the 6 levels of Bloom's taxonomy of the cognitive domain.

Comprehension:

It is defined as the ability to grasp the meaning of the material. Individual can make use of the content or idea being communicated without necessarily related it to other content and seeing its fullest implications.

Sample Question: explain the purpose of Bloom's taxonomy of the cognitive domain.

Application:

It refers to the ability to use the previously learned material in new and concrete situations. The abstractions may be in the shape of universal ideas, rules of methods.

Sample Question: write an instructional objective for each level of Bloom's taxonomy.

Item characteristic curve

To analyze items using IRT, the main thing need to consider is item characteristic curve (ICC). The item characteristic curve is considered as the basic building block of item response theory.

It gives a picture of:

The item difficulty

Discrimination power

The probability of answering correctly by guessing

Advantages of anecdotal record:

Following are the advantages of anecdotal records.

It depicts actual behaviors in natural situations.

Facilitate gathering evidence on the events that are exceptional but significant.

Beneficial for students with less communication skills.

Steps to design rubric

Basic Steps to design Rubric

Identify a learning goal.

Choose outcomes that may be measured.

Develop or adapt existing rubric.

Share it with students

Test security

Educational decisions

In the classroom, decisions are constantly being made. Teachers face huge numbers of dilemmas every day. These decisions can be of following nature

Types of Educational Decisions

- Instructional
- Grading
- Diagnostic
- Selecting
- Placement
- Counseling and Guidance
- Program or Curriculum
- Administrative

These types of decisions are taken at different levels. Some are decided at Board/Administrative level while some are taken at school management level and other are taken in classrooms by teachers

Scoring Rubrics for Essay Type Items

Scoring of essay items is a time consuming and difficult process. Reliability of the test demands that scoring should be consistent not only by the rater at different times but by two independent raters as well.

Scoring rubrics are descriptive scoring schemes that are developed by teachers or other evaluators to guide the analysis of writing products or processes of student‘s piece of writing or any performance

Alternatives

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

The correct alternative in each item is called the answer while all incorrect or less appropriate alternatives are called distractors or foiled and the student‘s task is to select the correct or best alternative from all the options.

Formative assessment

Formative assessment provides feedback and information during the instructional process, while learning is taking place, and while learning is occurring. Formative assessment measures student progress but it can also assess your own progress as an instructor.

Types of formative assessment

- Observations during in-class activities; of student‘s non-verbal feedback during lecture.
- Homework exercises as review for exams and class discussions)
- Reflections journals that are reviewed periodically during the semester
- Question and answer sessions, both formal (planned) and informal (spontaneous)
- Conferences between the instructor and student at various points in the semester
- In-class activities where students informally present their results
- Student feedback collected by periodically answering specific question about the instruction and their self-evaluation of performance and progress

Written test

In classroom assessment different forms of assessments are utilized. Each form of test has its own benefits and disadvantages. Most common type of assessment used in classrooms is written assessment

Analytical 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.

Analytic scoring requires a list of major elements that students are expected to include in an ideal answer of a question.

Analytic rubric is used to score an essay type of question on different points and significantly it works best on the restricted response type essay question.

Advantages

More detailed feedback.

Scoring more consistent across students and grades.

Disadvantages

Time consuming to score.

Norm referenced and criterion reference

NRT

Type of test which tells us where a student stands compared to other students. It helps determining a student‘s place or rank among a group of similar students. Such kind of test is called norm-referenced test Dimensions

It provides estimate of ability in a variety of skills in much shorter time.

NRT tend to be general. It measures variety of skills at same time but fails to measure them thoroughly.

It‘s hard to make decisions regarding the mastery of student‘s skill in subject

Criterion Referenced Assessment

A second type of test tells us about student‘s level of proficiency in or mastery of some skill or set of skills. This is achieved by comparing a student‘s performance to a standard mastery called a criterion. Test that yields such information is called Criterion Referenced Test

Dimensions

CRT tends to be specific. It measures particular set of skill at one time and focus on level of achievement of that skill. CRT gives clear picture regarding the mastery of student‘s skill in subject.

It measures skill more thoroughly so naturally it takes more time comparing to NRT in measuring the mastery of said skill

Items included in CRT are relatively easier. Around 80% of the students are expected to respond item correctly in the test

CRT compares students ‘performance to standards indicative of mastery.

Breadth of content sampled is narrow and covers very few objective

Q: what are subjective type questions? Tell names

ANS: 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

Q: High stake testing matters

ANS: 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.

Types of High stake test

Criterian Referenced Test

Norm-Referenced High-Stake Tests

Q: How to design rubric?

ANS: 1. Identify a learning goal.

2. Choose outcomes that may be measured.

3. Develop or adapt an existing rubric.

4. Share it with students

Q: Restricted response items names

ANS: Restricted response item: 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.)

Some of which are;

1. Analysis of relationship.

2. Compare and contrast positions.

3. Explain cause-effect relationship.

4. Organize data and support a viewpoint.

5. Formulate hypotheses.

6. Point out strengths and weaknesses.

7. Integrate data from various resources

3 marks

Q: Advantage if essay type questions

Ans: 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

Q: State any three limitations of mcqs

Ans: Limitations of MCQs. 1. Take a long time to construct in order to avoid arbitrary and ambiguous questions. 2. Also require careful preparation to avoid preponderance of questions testimony recall. 3. Provide cues that do not exist in practice.

Q: Benefits of rubric in evaluation process

ANS: 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.

Q: Item analysis

ANS: 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.

5 marks:

Q: How self assessment should be implemented

ANS: Boud (1995) argues that the way in which self-assessment is implemented is critical to its acceptance by students. According to Boud (1995), the implementation process needs to include:

A clear rationale: what are the purposes of this particular activity?

Explicit procedures—students need to know what is expected of them.

Reassurance of a safe environment in which they can be honest about their own performance without the fear that they will expose information which can be used against them.

Confidence that other students will do likewise, and that cheating or collusion will be detected and discouraged (Boud, 1995, p.182).

Q: What is report? Name of reports?

ANS: 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

Q: Explain formal and informal report

ANS: 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.

Q:

ANS: A derived score is a numerical report of test performance on a score scale that has well-defined characteristics and yields normative meaning. Most common type of derived scores are 1. Grade norms

2. Percentile norms

3. Standard norms

ANS: 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

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

ANS: Q1. Give examples to identify different forms of governments, their advantages and disadvantages over each other. Which form of government is most suitable in the socioeconomic and cultural context of our country? Your response will be graded on accuracy of information, power of arguments, conciseness and orderliness of presentation. It is suggested to keep your response limited to 5000-6000 words.

Q2. What are the teachings of Islam about the respect of elders? How do these teachings help us to make a good society?

ANS: Basic Steps to Design Rubric

1. Identify a learning goal.

2. Choose outcomes that may be measured.

3. Develop or adapt an existing rubric.

4. Share it with students.

ANS: 1. Make plans for the meeting.

2. Begin the meeting in a positive manner.

3. Present the student's strong points before describing the areas needing improvement.

4. Encourage parents to participate and share information.

5. Plan a course of action cooperatively.

6. End the meeting with a positive comment.

7. Use good human relation skills during the meeting

ANS: 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.

Q.4: What is the predictive validity?

Answer:

Predictive Validity:

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

Q.5: What is the diagnostic assessment? (From Topic 4) Diagnostic assessment:

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

Q.6: Describe the Bloom's taxonomy of cognitive domain? Answer: Cognitive Domain:

There are six levels of the cognitive domain:

1. Knowledge
2. Comprehension
3. Application
4. Analysis
5. Synthesis
6. Evaluation

Q.7: What is the test security? (From Topic 106) Answer:

Test Security:

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.

Q.8: How to design rubric?(From Topic 112) Answer:

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.

Q.9: How to assemble or packaging a test?(From Topic 103) Answer:

Assembling Test (Packaging the Test):

1. Grouping together items of similar format
2. Arranging test items from easy to hard
3. Properly spacing items
4. Keeping items and options on the same page
5. Placing illustrations near the descriptions
6. Checking the answer key
7. Deciding how students will record their answers
8. Providing space for test taker's name and date 9. Checking test directions for clarity
10. Proofreading the test

Q.10: Write a brief note on the restricted response essay type items?

Answer:

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. Some of which** are;

1. Analysis of *relationship.*
2. *Compare and contrast* positions.
3. Explain *cause-effect* relationship.
4. *Organize data* and support a viewpoint.
5. *Formulate hypotheses*.
6. Point out *strengths and weaknesses.*
7. *Integrate data* from various resources

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

1. *Supplying information is required* instead of simple recognition.
2. *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.

Q.11: How to write essay type questions? (From Topic 90) Answer:

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

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

Q.12: Explain Criterion and predictive Validity.

Answer:

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.

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

Q.1: Explain the Taxonomy levels of Cognitive.

Answer:

Bloom's Taxonomy of Cognitive Development

Bloom identified six levels within the cognitive domain, from the simple recall or recognition of facts, as the lowest level, through increasingly more complex and abstract mental levels, to the highest order which is classified as evaluation. A description of the six levels as well as verb examples that represent intellectual activity are listed here.

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, and reproduce state.

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, and translate.

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.

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 because 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, and test.

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 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.

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.

Q.2: What is the formula for the find out the item discrimination?(From Topic 97)

Answer:

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

Q.3: What are the guidelines for the essay type questions? (Repeat)

Answer:

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

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

Q.4: Describe the advantages for the essay writing test?

Answer:

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

Q.5: Explain Item difficulty in IRT.

Answer:

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).

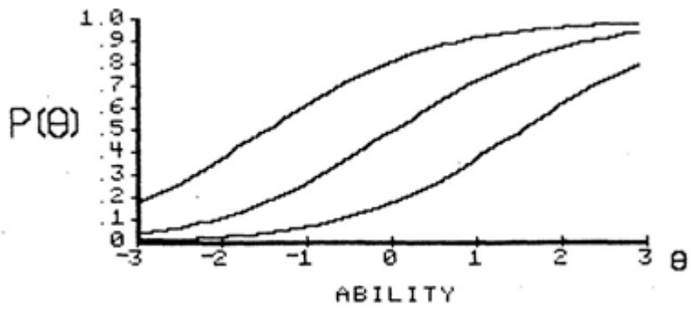


Figure 3: Three item characteristic curves with the same discrimination but different levels of difficulty (Baker, 2001, p. 10)

In Figure 2, three item characteristic curves are presented on the same graph. All have the same level of discrimination but differ with respect to difficulty. The left- hand curve represents an easy item because the probability of correct response is high for low-ability examinees and approaches 1 for high-ability examinees. The center curve represents an item of medium difficulty because the probability of correct response is low at the lowest ability levels, around 0.5 in the middle of the ability scale and near 1 at the highest ability levels. The right-hand curve represents a hard item. The probability of correct response is low for most of the ability scale and increases only when the higher ability levels are reached. Even at the highest ability level shown (+3), the probability of correct response is only 0.8 for the most difficult item.

Q.6: What is the Rubric, Write the element of the rubric?

Answer:

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

Q.7: What is the self-assessment?

Answer:

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 degree to which they reflect explicitly stated goals or criteria, identify strengths and weaknesses in their work, and revise accordingly.

Q.8: What is Analytic rubric?

Answer:

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.

Q.9: Describe the Advantages of Self-Assessment.

Answer:

Advantages:

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

- Self-assessment can accommodate diversity of learners‟ readiness, experience and backgrounds.

- **Q.10: Why use peer assessment?**

Answer:

- Falchikov (2007) reminds us that peer learning builds on a process that is part of our development from the earliest years of life (it is the practice of formal education and the centrality of the teacher that makes us lose sight of this).
- Peer feedback can encourage collaborative learning through interchange about what constitutes good work.
- If the course wants to promote peer learning and collaboration in other ways, then the assessment tasks need to align with this. It is also important to recognize the extra work that peer learning activities may require from students through the assessment. Boud,

  o Cohen & Sampson (1999) observe that “if students are expected to put more effort into a course through their engagement in peer learning activities, then it may be necessary to have this effort recognized through a commensurate shift in assessment focus” (p.416).

- Peer learning draws on the “cognitive apprenticeship model” (Kvale, 2006).
- Students can help each other to make sense of the gaps in their learning and understanding and to get a more sophisticated grasp of the learning process.
- The conversation around the assessment process is enhanced. Research evidence indicates that peer feedback can be used very effectively in the development of students‟ writing skills.
- Students engaged in commentary on the work of others can heighten their own capacity for judgment and making intellectual choices.
- Students receiving feedback from their peers can get a wider range of ideas about their work to promote development and improvement.
- Peer evaluation helps to lessen the power imbalance between teachers and students and can enhance the students‟ status in the learning process.
- The focus of peer feedback can be on process, encouraging students to clarify, review and edit their ideas.
- It is possible to give immediate feedback, so formative learning can be enhanced. Peer assessment processes can help students learn how to receive and give feedback which is an important part of most work contexts.

- Peer assessment aligns with the notion that an important part of the learning process is gradually understanding and articulating the values and standards of a “community of practice” (Wenger, 1999, cited in Falchikov, 2007, p.129). Drawing on Wenger‟s ideas,
    - Falchikov suggests that “learning involves active participation in a „community of practice‟ in which members of the community determine and structure their own practices, and construct identities in relation to these communities” (2007, p.129). Peer commentary in the assessment process initiates into the community to hear, experiment with and gradually internalize the norms of the community

Q.12: What are the objective test item types?

Answer:

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.

Q.13: Define Reliability with examples.

Answer:

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.

Q14: What is meant by the term of unreliable results? Answer:

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 of an accurate system of assessment. Keeping in view the significance of reliability, different methods are used to attain reliability. For example, test retest procedure, the alternative–test form procedure and the split halves procedure.

Q.1: Define the evaluation?

Answer:

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

Q.2: What is the item difficulty?

Answer:

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

Q.3:

What are the importance of the student self- assessment?

Answer:

1. **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.
2. **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.
3. **Providing feedback that moves learning forward**– working with students to provide them the information they need to better understand problems and solutions.
4. **Activating learners as instructional resources for one another**– getting students involved with each other in discussions and working groups can help improve student learning.
5. **Activating learners as owners of their learning**– self-regulation of learning leads to student performance improvement

Q.4:

What is the alternative form of test?

Answer:

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.

Q.5:

What is the maximum and typical performance?

Answer:

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.

Q.6:

What is the restricted response item?

Answer:

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.).

Q.7:

What is face validity and content validity?

Answer:

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.

Q.8:

Describe the curriculum in Pakistan at higher level education?

Answer:

- In 1973's constitution, Government placed education on the concurrent list.
- Federal Government took the responsibility to determine the curriculum text boards, policy, planning and standards of education.
- 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
- It also look after the degrees, certificates and diplomas awarded by degree colleges, universities and other institutions of HEC.

Q.9:

Describe the item difficulty in CTT?

Answer:

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 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. Many test experts believe that for a maximum discrimination between high and low achievers, the optimal levels (adjusting for guessing) are:

Q.10:

Explain the character tics of the Curve.

Answer:

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: a. The item difficulty

b. Discrimination power

c. The probability of answering correctly by guessing

Q.11:

What are the steps for packing test material?

Answer:

- Grouping together items of similar format
- Arranging test items from easy to hard
- Properly spacing items
- Keeping items and options on the same page
- Position illustration near description (if diagram given)

- Randomness of answer key
- Determine how students record answers (separate sheet or same sheet)
- Providing space for test taker‘s details
- Proofread the test (typographical and grammatical error)
- Test directions **Q.12:**

What is the Prohibited Professional Test Security Practices? Answer:

1. Give students access to test content or secure test questions before test administration.
2. Copy reproduces or uses in any manner inconsistent with test security measures of secure assessment material.
3. Review actual test items before, during or after the test administration.
4. 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.
5. Leave a testing room unsupervised at any time.
6. Make test answers available to students or permit the use of any supplemental or reference materials during test.
7. Assist a student directly or indirectly (gestures, pointing, prompting etc.,)
8. 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 response