EDU 431 Quiz no 1 8 June, 2020

1. Assessment that monitors learning progress ----------
A) Placement assessment
b) Formative Assessment
c) Diagnostic Assessment
d) Summative assessment

2. Which of following is the type of assessment by the use of classroom instruction……..
a) Complex performance assessment
b) Placement assessment
c) Fixed choice assessment
d) Norm reference assessment

3. To assess achievement at the end of instruction is?
A. Placement Assessment
B. Formative Assessment
C. Summative Assessment
D. Diagnostic Assessment

4. What is your name? is the example of which level of SOLO taxonomy
a) Pre-structural
b) Uni Structural
c) Multi- Structural
d) Relational

5. In norm referenced test the comparison is between _____
a) Groups
b) Interests
c) Individuals
d) None

6. The summative evaluation is……
a) Diagnostic
b) Certifying judgment
c) Continuous
d) None

7. Explain the process of water cycle is the question of -------- level in Blooms taxonomy of cognitive domain.
a) Knowledge
b) Comprehension

c) Analysis
d) Application

8. ----------is used when we want to assess student's prior knowledge
A) Placement assessment
b) Formative Assessment
c) Summative assessment
d) Diagnostic Assessment

9. Who was the founder of Pakistan is the question of----------in Blooms taxonomy of cognitive
a) Knowledge
b) Comprehension
c) Application
d) Synthesis

10. List, enlist, name, define keywords are used in …… level of Depth of knowledge
a) Recall
b) skill/concept
c) Strategic thinking
d) Extended thinking

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**
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**

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:** _____**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**

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**

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**

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**

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? _____ **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 ---------- _____ **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**

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**

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**

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**

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.

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.

---

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.

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

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

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

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.

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

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

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

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

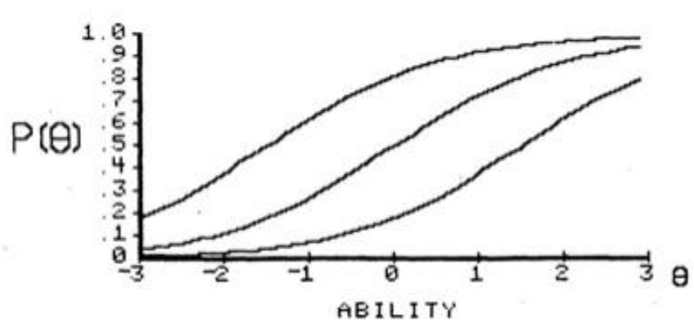
4. Keeping items and options on the same page

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).

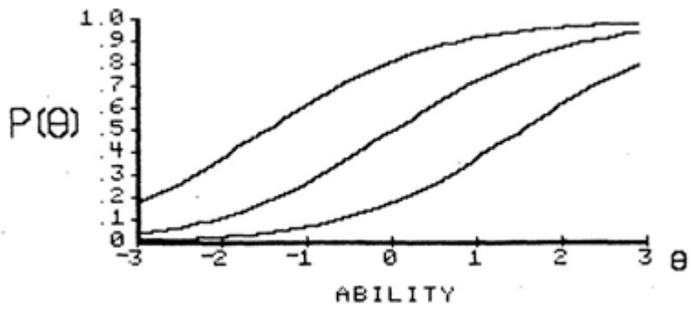


Figure 3: Three item characteristic curves with the same discrimination but different levels of difficulty (Baker, 2001, p. 10)

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

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**
11. In revised Bloom taxonomy; build, identify, develop are key words for level:_____ **Applying**
12. In revised Bloom taxonomy; classify, compare, contrast are key words for level:___ _____**Understanding**
13. Ali's writing score placed near the bottom of the class. The interpretation of this statement is:
14. Identify the lowest learning level of Bloom's cognitive domain_____ **Knowledge**
15. In norm referenced test the comparison is between _____. _____**Individuals**
16. The purpose of formative evaluation is _____. _**Monitoring progress of students**
17. The classification of cognitive domain was presented by_____ **Benjamin S. Bloom**
18. Which of the following decision determines prerequisite skills, degree of mastery of course goals and mode of learning?_____ **Placement**
19. Hina exceeded 85% of the sixth graders on the mathematics test. The interpretation of this statement is:
20.

Lecture 1: Concept of Assessment

Topic no 1: Measurement, Assessment and Evaluation

In this topic, we will discuss about the three main concepts of educational testing and measurement i.e. measurement, assessment and evaluation and the difference about them

At the end of this topic, students will be able to explain

1. What are educational measurement, assessment and evaluation?
2. Difference among educational measurement, assessment and evaluation

1. Measurement

Measurement is the process by which the attributes or dimensions of some object (both physical and abstract) are quantified. The tools used for this purpose may include test, observation, checklist, homework, portfolios, project etc. (the process of changing the person's ability into numbers is measurement). Measurement can be easily understood if we use this word to measure height and distance because these things are physical present in their existence so height can easily be measured by scale. But in the field of education, our variable are not physical and cannot be directly measured e.g. attitude, behavior, and achievement etc. these all are abstract, so there measurement is relatively difficult than those who have physical existence. The tool used for measuring the abstract variables cannot measure exactly like scale (thermometer).

So in this whole course, whenever the measurement word is used it means that tool which will be used for measuring student abilities and then converts it in numerical form.

2. Assessment

It means appraisal of something to improve quality of teaching and learning process for deciding what more can be done to improve the teaching, learning and outcomes.

3. Evaluation

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

© Copyright Virtual University of Pakistan

Let's take an example

In classroom situation, as a teacher, when you teach a chapter or unit to a class, first you made the objectives, either you will make it yourself as a teacher or you take it from curriculum document of the particular subject. Objectives are also written at the start of the chapter or book which shows that at the end of the unit what student will be able to do this, that is also referred as Student learning outcomes. So these SLOs can be checked by two aspects i. assessment ii. Evaluation. When you, as a teacher reflect on your teaching on daily basis that what you teach yesterday was a good way to teach or not, what i taught was the students need, did they understand what i teach, and then you can decide changing which things can cause the improvement in the learning process of students. This is ASSESSMENT. In the end of assessment, opinion is not about the individual, it is about the process from which the students or individuals are passing for the betterment of that process.

Evaluation is that when the learning process is complete and you want to see what are your targets or objectives and how much my students achieved those objectives , then tools are made and measures comes from that tell that how much your student learn and what is the quality of their learning.

Difference in measurement, assessment and evaluation

These terms are not different words for same concept but a process serving as prerequisite to each other and having unique purpose.

Every mechanism or process start with the measurement

In the field of education, measurement generate tools i.e. Test, Observation, Quiz, Checklist, Homework, Portfolios. They can give information about student learning. The information we got from the tool is *MEASUREMENT*

Now if that measurement is used for the process of making teaching learning process better than it is *ASSESSMENT*

e.g. if i want to decide that the procedure or way i adopted for student learning i.e. books, activities, field experience is good for their learning and they are according to the needs of student, then it is *ASSESSMENT*

© Copyright Virtual University of Pakistan

And if from these measure i decide that either their learning is getting better or not or how much student learned what i taught them and what is the quality or standard of that learning, then it is *EVALUATION*

Assessment and Evaluation don't exist in hierarchy; they both are parallel and different in purpose. Measurement is the source to move towards assessment and evaluation because it provides base and evidence to quantify teaching learning process. The quantified number has no meaning until we do assessment or evaluation. Assessment purpose is to make teaching learning process better so that student learning improve and measurement purpose is to align the learning with purpose.

Topic 2: Classroom Assessment

In this topic students will learn

1. Ways to classify the classroom assessment
2. Types of assessment in each classification

Concept of Classroom Assessment

The process of gathering, recording, interpreting, using and communicating information about a child's progress and achievement during the development of knowledge, concepts, skills and attitudes.

When we are teaching in a classroom, we are doing four things

i. Developing students' knowledge

ii. Improving their concepts

iii. Teaching the skills

iv. Making their attitudes.

And while doing so, we collect information about the development of these things, and then we record it, interpret it and use it to communicate about the learning progress of students.

This procedure is called classroom assessment

Classification of Assessment

Assessment can be classified in four ways

1. Nature of Assessment
2. Format of Assessment
3. Use in classroom instruction

© Copyright Virtual University of Pakistan

4. Method of interpreting results

1. Nature of Assessment

i. Maximum Performance Assessment

ii. Typical Performance Assessment

2. Format of Assessment

i. Fixed Choice Assessment

ii. Complex Performance Assessment

3. Use in classroom instruction

i. Placement Assessment

ii. Formative Assessment

iii. Diagnostic Assessment

iv. Summative Assessment

4. Method of interpreting results

i. Norm Referenced Assessment

ii. Criterion Referenced Assessment

Topic 3: Types of Assessment

In this topic, Student will learn

1. Types of assessment by nature
2. Types of assessment by format

1. By Nature of Assessment

i. Maximum Performance Assessment

ii. Typical Performance Assessment

i. Maximum Performance Assessment

Maximum performance assessment determines what individual can do when performing at their best. E.g. assess student in an environment when they exhibit their best performance. Procedure of this type is concerned with how well individual perform when they are motivated to obtain as high score as possible. This type of assessment includes Aptitude tests and Achievement tests.

In Achievement test , student learn by themselves or we teach them and at the end we want to see that how much student learn against our target so in this situation we make a test from which we

© Copyright Virtual University of Pakistan

can determine their best abilities. It is designed to indicate the degree of success in some past learning activity.

Aptitude test, we measure through aptitude test when we want to predict the success in future learning activity e.g. it is used when we want to see the interest of student in a particular field like medicine, sport, teaching. We know that different abilities used for going in different professions, we make a test depending on these abilities and then try to assess, in what abilities the students perform well.

ii. Typical Performance Assessment

The second category is typical performance assessment determines what individual will do under natural conditions. This type of assessment includes attitude, Interest and personality inventories, observational techniques, peer appraisal. Here emphasis is on what students will do rather than what they can do.

2. By format of assessment

i. Fixed Choice Assessment

ii. Complex Performance Assessment

i. 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. From these, we can assess student abilities of lower level learning

Fixed Choice Assessment is used for efficient measurement of knowledge and skills. This type of assessment includes standardized multiple choice questions

ii. 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.

E.g. if want to measure the student ability of writing an essay and this cannot be judged by fixed response items

Topic 4: Use of Assessment in Classroom Instruction

Placement and Diagnostic

© Copyright Virtual University of Pakistan

In this session student will learn Classification of assessment in terms of its uses in classroom instruction

3. Use in classroom instruction

i. Use of Placement Assessment
ii. Use of Diagnostic Assessment
iii. Formative Assessment
iv. Summative assessment

i. Placement Assessment

Placement Assessment determines prerequisite skills, degree of mastery of course goals and mode of learning. Placement assessment is used when we want to assess student's prior knowledge so that we can decide what the level of student is. It is associated with student's entry level performance to know either student have a sufficient knowledge required for a particular course or not. Through placement assessment, teacher can be able to know that where student should be place according to their present knowledge or skills. It determines the level of student knowledge at the beginning of session and helps teacher plan the lesson accordingly. In the classroom, the teacher can use placement assessment to assess the level of students' knowledge and skills and then make lesson plans keeping in mind the level and need of students accordingly.

It also determines the interest and aptitude of student regarding a subject and helps in selecting correct path for future.

Examples

Readiness test: It is a test used to determine the students' knowledge or concept about a particular course of instruction or what is the level of students

Aptitude test: It is used for the admission in a particular program

Pretest: It is made according to the course objectives and determines the student present knowledge about them

Self- report inventories: Determines the student level by interviewing or discussion

ii. Diagnostic Assessment

Diagnostic Assessment determines causes (intellectual physical, emotional environmental) of persistent learning difficulties. e.g. if you are having a headache , first you will try to cure it by yourself by taking a medicine and you got a relief but if you didn't got a relief by

© Copyright Virtual University of Pakistan

taking a medicine then either you change your medicine or you go to the physician or doctor. At first, doctor prescribed medicines , if you still have headache you again go to the doctor , then the doctor suggest you the tests i.e. blood test, urine test etc. and then finally by seeing the test report the doctors able to recognize the reason or cause of headache. And when doctors know the root of your headache then he will prescribe you the medicine for that cause, this is diagnosis.

Diagnosis doesn't start first day, it is for the constant or continuous problems e.g. if a student continuous to experience failure in reading or mathematics or any other subject despite the use of prescribed alternative methods , then a diagnosis is indicated. Teachers' try to find out what is the root of students failure.

Topic 5: Use of Assessment in Classroom Instruction – Formative and Summative Assessment

In this session students will learn uses of:

i. Formative Assessment

Ii. Summative Assessment

i. 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.

It is conducted during the academic session or teaching-learning process so that i can get a feedback about my way of teaching and how students are learning and decisions are made on the basis of results immediately. It is and ongoing process to modify teaching strategies on the basis of students need.

It provides feedback to teachers

. About weakness and strength of learning process

. To modify their teaching practices

© Copyright Virtual University of Pakistan

. To improve teacher-learning process

It also helps students to reflect on their weaknesses and encourages them for their successful learning. (When we tell students that their problems is their way of learning rather than their intelligence, then we tell them how to change your way of learning to learn better. With this, students can reflect on their learning process alone, and then what goes wrong with my learning process)

Formative assessment provides feedback to students who are struggling with specific content area or concept.

The main difference between formative and summative assessment is that the in formative assessment improvement is in the process of learning rather than to certify students We use different tools for formative assessment and it includes teacher made tests, custom made tests from textbook publishers, observational techniques

ii. Summative Assessment

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. It usually compares the student learning either with other students learning (norm-referenced) or the standard for a grade level (Criterion-referenced). Summative assessment includes teacher made survey test, performance, rating scales, product scales.

Topic 6: Types of Assessment

Methods of Interpreting Results

4. Method of interpreting result

i. Norm-referenced Assessment

ii. Criterion-referenced Assessment

i. Norm-referenced Assessment

© Copyright Virtual University of Pakistan

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 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

ii. 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.

© Copyright Virtual University of Pakistan

Lecture 2: Assessment, Testing and National Curriculum

In this unit, students will learn about link between curriculum and assessment. For this purpose, we precede our discussion in reference to National Curriculum of Pakistan 2006.

Topic 7: Role of national curriculum in assessment

In national curriculums of Pakistan, learning of student is classified into four levels.

- Competency
- Standards
- Benchmarks
- Student learning outcome (SLOs)

Competency

It is a key learning area. For example algebra, arithmetic, geometry etc. in mathematics and vocabulary, grammar, composition etc. in English.

Standards

These define the competency by specifying broadly, the knowledge, skills and attitudes that students will acquire, should know and be able to do in a particular key learning area during twelve years of schooling.

Benchmarks

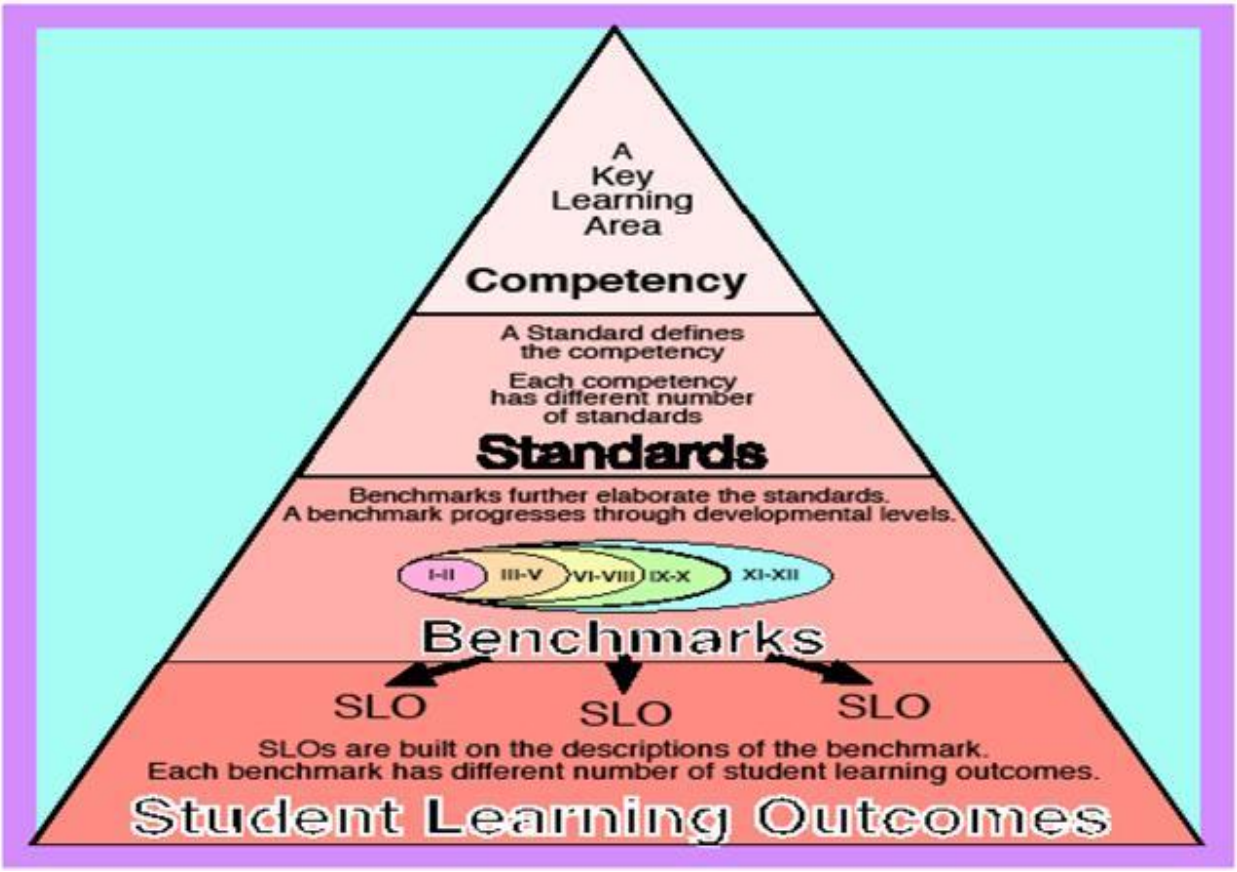
The benchmarks further elaborate the standards, indicating what the students will accomplish at the end of each of the five developmental levels in order to meet the standard.

Student learning outcomes

These are built on the descriptions of the benchmarks and describe what students will accomplish at the end of each grade. It is the lowest level of hierarchy.

Topic 8: Connecting all four levels in curriculum

© Copyright Virtual University of Pakistan



In the image above, SLOs are at the bottom which is the lowest level. All SLOs combined to make a benchmark and benchmarks convert into standards and then into competency.

Example:

Example is taken from curriculum of English subject.

Competency 1: Reading and thinking skills

Standard 1: All students will search for, discover and understand a variety of text types through tasks which require multiple reading and thinking strategies for comprehension, fluency and enjoyment.

Benchmark 1: use reading readiness strategies.

Student learning outcome:

© Copyright Virtual University of Pakistan

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.

Topic 9: Modes of Assessment in Curriculum

Curriculum document provides specific guidelines for assessment.

Example:

6.1. The two forms of assessment recommended are:

6.1.1. periodic/formative assessment through homework, quizzes, class tests and group discussions.

6.1.2. end of term/ summative assessment through final examination.

Purpose of assessment and curriculum-English 2006

The assessment system for the present curriculum should include

- A clear statement of the specific purpose(s) for which the assessment us being carried out.
- A wide variety of assessment tools and techniques to measure students ability to use language effectively.
- Criteria to be used for determining performance levels for the SLOs for each grade level.
- Procedures for interpretation and use of assessment results to evaluate the learning outcomes.

Form of suitable assessment tools- English 2006

- MCQs
- Constructed response
    - Restricted response
    - Extended response
- Performance tasks

© Copyright Virtual University of Pakistan

Lecture 3

Topic 10: Taxonomies of Educational Objectives and Assessment

In this session student will learn:

1. Concept and use of Taxonomies
2. Three popular Taxonomies

Every assessment, regardless of its purposes rests on three important pillars:

1. A model for how students present knowledge and develop competence in the subject domain
2. Tasks or situations that allow the examiner to observe the students' performance
3. Inferences from performance evidence about the quality of learning.

In developing a test to assess student learning, taxonomy provides a framework of categories with different hierarchical levels of outcomes

Popular Taxonomies

1. Bloom's taxonomy of educational objective
2. Structure of Observed Learning Outcomes (SOLO)
3. Depth of Knowledge (DOK)

Topic 11: Bloom's Taxonomy and SOLO Taxonomy

In this session students will learn

1. Levels of Structure of Observed Learning Outcomes (SOLO)
2. Levels of Depth of Knowledge (DOK)
3. Levels of Bloom's taxonomy of educational objective

Structure of Observed Learning Outcomes (SOLO)

© Copyright Virtual University of Pakistan

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

Levels of Structure of Observed Learning Outcomes (SOLO)

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

Depth of Knowledge (DOK)

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

Levels of Depth of Knowledge (DOK)

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

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.

© Copyright Virtual University of Pakistan

1. Cognitive
2. Affective
3. Psychomotor

Cognitive Domain

i. Knowledge
ii. Comprehension
iii. Application
iv. Analysis
v. Synthesis
vi. Evaluation

Affective Domain

i. Receiving
ii. Responding
iii. Valuing
iv. Organization
v. Characterization

Psychomotor Domain

i. Perception
ii. Set
iii. Guided Response
iv. Mechanism
v. Complex covert Response
vi. Adaption
vii. Origination

© Copyright Virtual University of Pakistan

Topic 12: SOLO Taxonomy

In this topic student will learn

1. Levels of Structure of Observed Learning Outcomes (SOLO).
2. Indicative verbs for each level

Levels of SOLO

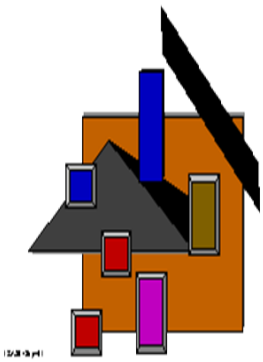
1. Pre-structural
2. Uni-Structural
3. Multi-structural
4. Relational
5. Extended Abstract

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?

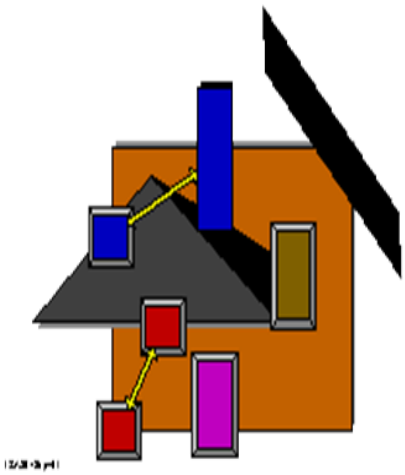


2. Uni Structural

© Copyright Virtual University of Pakistan

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.

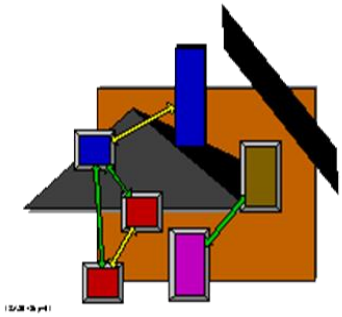
Indicative verbs: identify, memorize, do simple procedure



3. 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.



4. Relational

Student can indicate connection between facts and theory, action and purpose. Shows understanding of several components which are integrated conceptually showing

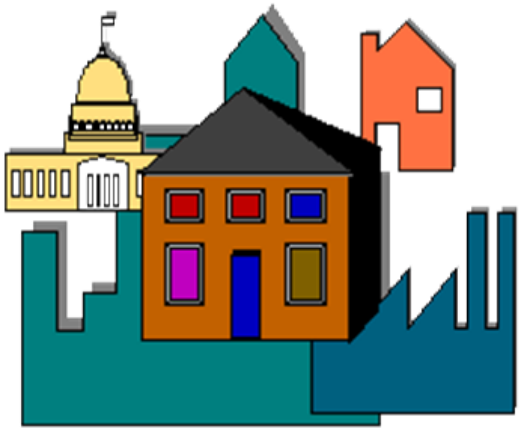
© Copyright Virtual University of Pakistan

understanding of how the parts contribute to the whole. Indicative verbs: compare/contrast, explain causes, integrate, analyze, relate, and apply.



5. 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



© Copyright Virtual University of Pakistan

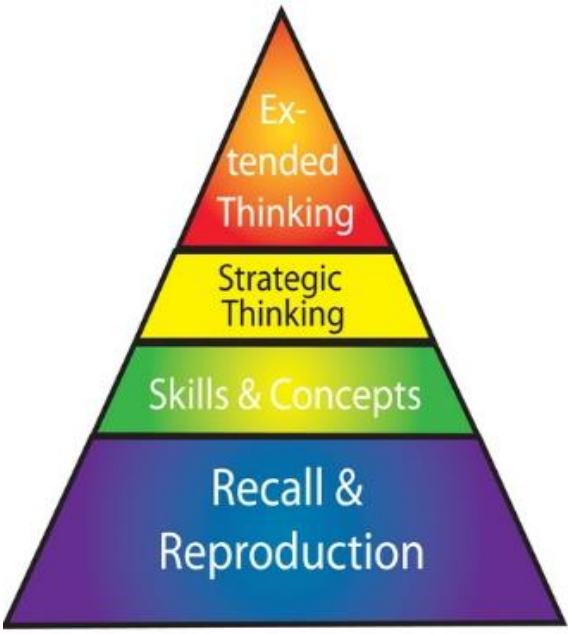
Topic 13: Depth of Knowledge

In this topic, Students will learn

1. Levels of DOK (Depth of Knowledge)
2. Key verbs for each level

Levels of DOK

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



© 2013 Heidi Butkus www.heidisongs.com

DOK measures the degree to which the knowledge bring about from students on assessments is as complex as what students are expected to know and do as stated in the curriculum.

Levels of DOK (Depth of Knowledge)

Recall

Recall of a fact, information, or procedure. The subject matter at this particular level usually involves working with facts, terms and/or properties of objects.

Key words: list, enlist, name, define etc.

Skill/Concept

It includes the engagement of some mental processing beyond recalling or reproducing a response. Use information or conceptual knowledge, two or more steps, not just recalling

Key words: Graph, separate, relate, contrast, narrate, compare etc.

Strategic Thinking

Items falling in this category demand a short-term use of higher order thinking processes, such as analysis and evaluation, to solve real-world problems with predictable outcomes.

Key words: Argue, critique, formulate

Extended Thinking

Learning outcomes to this level demand extended use of higher order thinking processes such as synthesis, reflection, assessment and adjustment of plans over time.

Key words: Create, Synthesize, Design and reflection.

Topic 14: Bloom's Taxonomy

In this topic student will learn:

1. Bloom's taxonomy of educational objectives.
2. Old and revised taxonomy.

© Copyright Virtual University of Pakistan

3. Key words for each level

Three Domains of Learning

There are three main domains of learning and all teachers should know about them and use them to construct lessons.

- Cognitive Domain
- Affective Domain
- Psychomotor Domain

1. **Cognitive or Thinking Domain**

In 2000-01 revisions to the cognitive taxonomy were spearheaded by one of Bloom's former students, Lorin Anderson, and Bloom's original partner in defining and publishing the cognitive domain, David Krathwohl. One of the major changes that occurred between the old and the newer updated version is that the two highest forms of cognition have been reversed.

Levels of Cognitive Domain

Old Cognitive Domain

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.

© Copyright Virtual University of Pakistan

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.

Topic 15: Bloom's Taxonomy (Continued)

In this topic student will learn:

1. Bloom's taxonomy of educational objectives.
2. Old and revised taxonomy.
3. Key words for each level

Analysis:

The breakdown of a concept into its constituents parts such that the relative hierarchy of the concept is made easy to understand or the relation between the parts of the concept is elaborated. Sample Question: compare and contrast the cognitive and affective domains.

Synthesis:

There is a collection of the constituents or parts of a concept so as to make a whole. This is a stage when an individual is working with the pieces and assorting them in such a way as to formulate a pattern or structure not clearly there before. Sample Question: Design a classification scheme for writing educational objectives that combines the cognitive, affective, and psychomotor domains.

Evaluation:

It is concerned with the ability to judge the value of the material for a given purpose. Judgments are made on the definite criteria. Sample Question: How far the different BISEs and universities are developing papers using Bloom's taxonomy? Support your answer with arguments.

© Copyright Virtual University of Pakistan

Topic 16: Revised version of Bloom's Taxonomy

In this topic student will learn:

1. Old and revised taxonomy.
2. Key words for each level

Revised Cognitive Domains

Levels

Remembering:

Exhibitmemory of previously learned material by recalling facts, terms, basic concepts, and answers.

Key verbs:

Choose, Define, Find, How, Label, List, Match, Name, Omit, Recall, Relate, Select, Show, Spell, Tell, What, When, Where, Which, Who, Why

Understanding:

Constructing meaning from different types of functions be they written or graphic messages, or activities.

Key verbs: Classify, Compare, Contrast, Demonstrate, Explain, Extend, Illustrate, Infer, Interpret, Outline, Relate, Rephrase, Show, Summarize, Translate

Applying:

Solve problems to new situations by applying acquired knowledge, facts, techniques and rules in a different way.

Key verbs:, Apply, Build, Choose, Construct, Develop, Experiment with, Identify, Interview, Make use of, Model, Organize, Plan, Select, Solve, Utilize.

© Copyright Virtual University of Pakistan

Analyzing:

Breaking materials or concepts into parts, determining how the parts relate to one another, or how the parts relate to an overall structure or purpose.

Key verbs: Analyze, Assume, Categorize, Classify, Compare, Conclusion, Contrast, Discover, Dissect, Distinguish, Divide, Examine, Function, Inference, Inspect

Evaluating:

Making judgments based on criteria and standards through checking and critiquing.

Key verbs: Agree, Appraise, Assess, Award, Choose, Compare, Conclude, Criteria, Criticize, Decide, Deduct, Defend, Determine, Disprove, Estimate

Creating:

Putting elements together to form a coherent or functional whole; reorganizing elements into a new pattern or structure through generating, planning, or producing.

Key verbs: Adapt, Build, Change, Choose, Combine, Compile, Compose, Construct, Create, Delete, Design, Develop, Discuss, Elaborate, Estimate, Formulate

These categories range from simple to complex and from concrete to abstract level of student's learning. It is assumed that the taxonomy represents a cumulative hierarchy, so that mastery of each simpler category is considered as prerequisite to mastery of the next, more complex one.

© Copyright Virtual University of Pakistan

Comparison of Bloom, SOLO and DoK

A comparison of hierarchical division in three taxonomies of learning objectives

SOLO Taxonomy	Bloom's Taxonomy	DOK Taxonomy
Level 1: Pre-structural Level 2: Uni-structural	Level 1 Knowledge	Level 1:REcall
Level 3:Multi-structural	Level 2:Comprehension	Level 2: Skill/Concept
Level 4: Relational	Level 3: Application	Level 3 Strategic Thinking
Level 5: Extended Response	Level 4: Analysis Level 5: Synthesis Level 6: Evaluation	Level 4: Extended Thinking

Topic 17: Instructional Objectives

In this session student will learn:

1. Instructional Objectives as Learning Outcome
2. General objectives
3. Specific objectives

Instructional Objectives as Learning Outcome

Instructional goals and objectives are stated in terms of actions to be taken

When viewing instructional objectives in terms of learning outcomes, we are concerned with products rather than process of learning

© Copyright Virtual University of Pakistan

Sources for lists of Objectives

- Professional Associations standards
- State Content Standards
- Methods books
- Year books
- Encyclopedia of educational Research
- Curriculum Frameworks
- Test manuals

Criteria of selecting the final list of objectives

Prepare tentative list of instructionally relevant learning outcomes

Review the list for

1. Completeness
2. Appropriateness
3. Soundness
4. Feasibility

General Objectives

- Stating the general objectives is selecting the proper level of generality.
- Objective should be specific enough to provide the direction for instruction but not so specific that instruction is reduced to training
- Stating general objectives in general terms, we provide for the integration of specific facts and skills into complex response

© Copyright Virtual University of Pakistan

- General statements gives teachers freedom in selecting the method and materials of instruction

List of general objectives shows the desired level of generality

- Knows basic terminology
- Understands concepts
- Relates concepts to everyday observations
- Applies principles to new situations
- Interpret graphs
- Demonstrate scientific attitude

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

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

© Copyright Virtual University of Pakistan

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

© Copyright Virtual University of Pakistan

Lecture 4: Purpose of Testing

Topic 18: Educational Decisions making

In this session student will learn:

1. Use of testing for decision making
2. Types of educational decisions

Why Test?

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

Instructional Decisions

Instructional decisions are the nuts and bolts types of decisions made in classroom by teachers.

© Copyright Virtual University of Pakistan

These are most frequently made decisions. Such decisions include deciding to

- Spend more time on specific units
- Regroup student in class for better management
- Instructional plans

Grading Decisions

Educational decisions based on grades are also made by the classroom teacher but much less frequently than instructional decisions. For most students grading decisions are most influential decision made about them

Diagnostic Decisions

Diagnostic decisions are those made about a student's strengths and weaknesses and the reasons behind them. Teachers make diagnostic decisions based on information yielded by an in-formal teacher made test

Decisions of diagnostic nature can also be made by the help of standardized tests (will be discussed in next session)

Selection Decisions

Selection decisions involves test data used in part for accepting or rejecting applicants for admission into a group, program, or institution

Placement Decisions

Placement decisions are made after an individual has been accepted in a program. They involve determining where in program someone is best suited to begin with.

Counseling and Guidance Decisions

Counseling and guidance decisions involve the use of test data to help recommend programs of study that are likely to be appropriate for the students

© Copyright Virtual University of Pakistan

Program or curriculum decision

This type of decision is taken at policy level. Where it is decided if a lesion, unit or subject will continue or abandoned for next academic session according to the national objectives of education.

Administrative Decisions

Administrative policy decisions may be made at school, district, state or national level.

Based on measurement data. This includes financial decisions of schools

Topic 19: Types of test

In this session student will learn:

- Types of written tests

How to Measure

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

Types of Written Tests

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

© Copyright Virtual University of Pakistan

- 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

© Copyright Virtual University of Pakistan

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.

Topic 20: Norm- Referenced Assessment (NRT)

In this session student will learn:

1. Characteristics of norm-referenced assessment
2. Questions included in NRT

Why Test?

General purpose of assessment is to gather information to make better and more informed decision. The utility of that information is what differentiate between types of assessments. In earlier session classification of assessment by method of interpreting results was discussed. This session will further unpack the complexity of norm and criterion referenced assessment

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.

© Copyright Virtual University of Pakistan

It provides estimate of ability in a variety of skills in much shorter time. NRT are much difficult for students to solve. On average only 50% students are able to get an item right in a test

Topic 21: Criterion Referenced Assessment

In this session student will learn:

1. Characteristics of Criterion-referenced assessment
2. Questions included in CRT

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 objectives

Topic 22: Characteristics of Criterion Referenced Assessment

In this session student will learn:

1. Characteristics of Criterion-referenced assessment
2. Questions included in CRT

© Copyright Virtual University of Pakistan

Dimensions

- Sampled content in CRT is much more comprehensive, usually three or more items are used to cover single objective.
- The meaning of the score does not depend upon on comparison with other scores.
- It flows directly from the connection between the items and the criterion.
- Items are chosen to reflect the criterion behavior. Emphasis is placed upon the domain of relevant responses.
- Number succeeding or failing or range of acceptable performance used.
- Example: 90% proficiency achieved, or 80% class reached 90% proficiency.

Topic 23: Difference between NRT and CRT

In this session student will learn:

1. Factors differentiating items of CRT and NRT

Basis of comparison

- Comparison targets
- Selection of items
- Meaning of success
- Average item difficulty
- Score distribution
- Reported scores

Comparison targets

In CRT, the examinee's performance is compared to an external standard of competence. While in NRT, examinee's performance is typically compared to that of other examinees.

Selection of items

© Copyright Virtual University of Pakistan

Items included in CRT are of specific nature and designed for the student skilled in particular subject. In NRT items are of general knowledge nature. Student should be able to answer it but superficial knowledge is sufficient to respond the item correctly

Meaning of success

In CRT, an examinee is classified as a master or non-master. There is no limit to the number of 'pass' or 'fail'. In NRT, examinee's opportunity for success is relative to the performance of the other individuals who take the test.

Average item difficulty

In CRT, the average item difficulty is fairly high. Examinees are expected to show mastery. In NRT, the average item difficulty is lower. Tests are able to spread out the examinees' and provide a reliable ranking.

Score Distributions

In CRT, a plot of the resulting score distribution will show most of the scores clustering near the high end of the score scale. In NRT, broader spread of scores is expected, with a few examinees earning very low or high scores and many earning medium scores.

Reported Scores

In CRT, classification of the examinee as master/non-master or pass/fail. In NRT, percentile ranks or scale scores are frequently used.

Topic 24: Formative Assessment

In this session student will learn:

1. Concept of Formative assessment
2. Types of formative assessment

In earlier session classification of assessment by use in classroom instruction was discussed. This session will further unpack the complexity of norm and criterion referenced assessment

© Copyright Virtual University of Pakistan

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

Topic 25: Functions of Formative Assessment

In this session student will learn:

1. Functions 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.

© Copyright Virtual University of Pakistan

Topic 26: Summative Assessment

In this session student will learn:

1. Concept of summative assessment
2. Types of summative assessment

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

© Copyright Virtual University of Pakistan

Topic 27: Functions of Summative Assessment

In this session student will learn:

1. Functions of summative assessment.

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

© Copyright Virtual University of Pakistan

Lecture 5: Table of Specification

Topic 28: Table of specification (video 28)

In this session student will learn:

1. Concept of two way table of specification
2. Two way table of specification

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

- It helps a teacher in allotting the questions to different content areas and Bloom's learning categories in a systematic manner.
- 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)
- A unit test or comprehensive exam is based on several lessons and/or chapters in a book supposedly reflecting a balance between content areas and learning levels (objectives).

Two way Table of Specification

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

Table of specification performs two important functions

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

© Copyright Virtual University of Pakistan

Learning objectives	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation	Weightage %age
1							
2							
3							
4							
5							

Activate

Topic 29: Concept of table of specification (video 29)

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

Learning objectives	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation	Weightage %age
1							
2							
3							
4							
5							

Activate

Topic 30: Elements and Appropriateness in Table of Specification (video 30)

In this session student will learn:

1. Elements in table of specification
2. Appropriateness of Table of Specification.

Elements in table of specification

Carey (1988) listed six major elements that should be attended to in developing a Table of Specifications for a comprehensive end of unit exam:

1. Balance among the goals selected for the exam (weighing objectives)
2. Balance among the levels of learning (higher order and lower order
3. The test format (objective and subjective)
4. The total number of items
5. The number of test items for each goal and level of learning
6. The enabling skills to be selected from each goal framework.

A Table of Specifications incorporating these six elements will result in a "comprehensive posttest that represents each unit and is balanced by goals and levels of learning"

Checklist for appropriateness of table of specification

1. Are the Specifications in harmony with the purpose of the test?
2. Do the specifications indicate the nature and limits of the achievement domain?
3. Do the specifications indicate the types of learning outcomes to be measured?
4. Do the specifications indicate the sample of learning outcomes to be measured?
5. Is the number of test items indicated for the total test and for each subdivision?
6. Are the types of items to be used appropriate for the outcomes to be measured?

© Copyright Virtual University of Pakistan

7. Is the difficulty of the items appropriate for the types of interpretation to be made?
8. Is the distribution of items adequate for the types of interpretation to be made?
9. If sample items are included, do they illustrate the desired attributes.
10. Do the specifications, as a whole, indicate a representative sample of instructionally relevant tasks that fits the used to be made of the results?

Topic 31: Balance among Learning Objectives and their Weight in table of specification

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.

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%

Step 2: Examine value

Then the instructor should look at the number of test items/score to be allocated to objective/content/theme 1.

© Copyright Virtual University of Pakistan

Let us assume total marks of there are 100. Then 25 marks should be allocated to questions related to objective/content/theme 1.

Step 3

Percent of instruction time = Percent of examination value (within ±2 percent, if not, redo test)

25 ±2 = 25 ±2

It can be a bit tricky if the total marks of the test are 50. Then 25% of 50 will be 12.5 marks. Point total of questions for objective / total points * on examination = % of examination value

Topic 32: Balance among the Levels of Learning Objectives in Table of Specification

In this session student will learn:

1. Practical example to develop table of specification

We have learnt to give weightage to the content area in a table of specification. Now we look at an example to develop table of specification practically. Following is the table of specification comprised of topics to be cover in test and their weightage that represent percentage of marks for each topic.

Topics/Level	Knowledge	Comprehension	Application	Marks
Pakistan Movement Time: (100/500)*100 = 20%				
Geography of Pakistan Time: (150/500)*100 = 30%				
Climate Change Time: (150/500)*100 = 20%				

© Copyright Virtual University of Pakistan

Industries Time: (50/500)*100 = 10%				
Economy Time: (50/500)*100 = 10%				
Total (Time: 500/Marks: 50)				

Let's consider that we have to develop a test of 50 marks according to the above discussed table of specification then distribution of marks for each topic is as under.

Topics/Level	Knowledge	Comprehension	Application	Marks
Pakistan Movement Time: (100/500)*100 = 20%				10 (20%)
Geography of Pakistan Time: (150/500)*100 = 30%				15 (30%)
Climate Change Time: (150/500)*100 = 20%				15 (30%)
Industries Time: (50/500)*100 = 10%				5 (10%)
Economy Time: (50/500)*100 = 10%				5 (10%)
Total (Time: 500/Marks: 50)				50 (100%)

© Copyright Virtual University of Pakistan

Then we have to consider the importance of each topic for cognitive level of questions according to Bloom's Taxonomy.

Topics/Level	Knowledge	Comprehension	Application	Marks
Pakistan Movement Time: (100/500)*100 = 20%	5 (50%)	2 (20%)	3 (30%)	10 (20%)
Geography of Pakistan Time: (150/500)*100 = 30%	2 (10%)	6 (40%)	7 (50%)	15 (30%)
Climate Change Time: (150/500)*100 = 20%		7 (50%)	8 (50%)	15 (30%)
Industries Time: (50/500)*100 = 10%	1 (10%)	1 (20%)	3 (70%)	5 (10%)
Economy Time: (50/500)*100 = 10%	1 (20%)	1 (20%)	3 (60%)	5 (10%)
Total (Time: 500/Marks: 50)	9 (18%)	17 (34%)	24 (48%)	50 (100%)

© Copyright Virtual University of Pakistan

Lecture 6: Selection of Test

Topic 33: Selecting pre-designed

In this session student will learn:

1. Selection of pre-designed test

Selecting Pre-designed test

Published test, supplement and complement informal classroom tests, and aid in many instructional decisions.

Published test are designed and conducted in such a manner that each and every characteristic is pre planned and known.

There are many published tests available for school use. The two most value to the instructional program are:

1. Achievement tests
2. Aptitude tests

There are hundreds of tests available for each type. Selecting the most appropriate one is important task. In some cases published tests are used by teachers. But more frequently these are used by provincial or national testing programs.

In classrooms most used published tests are:

1. Achievement tests
2. Reading test

Published tests commonly used by provincial or national testing programs are:

1. Aptitude tests
2. Readiness tests
3. Placement tests

Topic 34: Standards for selecting appropriate test

In this session student will learn:

1. Standards for selecting appropriate test

© Copyright Virtual University of Pakistan

Standards for selecting appropriate test

Test users should select tests that meet the purpose for which they are to be used and that are appropriate for intended population.

1. First define the purpose for testing and the population to be tested and select the test accordingly.
2. Investigate the potentially useful sources of information, in addition to the test scores, to validate the information provided by tests.
3. Read the materials provided by test developers and avoid using tests for which unclear or incomplete information is provided.
4. Become familiar with how and when test was developed and tried out.

Topic 35: Standards for selecting appropriate test (Continue)

In this session student will learn:

1. Standards for selecting appropriate test

Standards for selecting appropriate test

Test users should select tests that meet the purpose for which they are to be used and that are appropriate for intended population.

5. Read independent evaluations of a test and of possible alternative measures.
6. Examine specimen sets, disclosed tests or sample questions directions, answer sheets, manuals and score reports before selecting the tests.
7. Select and use only those tests for which the skills needed to administer the test and interpret scores correctly are available.

Topic 36: Fairness in selecting appropriate tests

In this session student will learn:

1. Fairness in selecting appropriate tests

Fairness in selecting appropriate tests

© Copyright Virtual University of Pakistan

1. Evaluate the procedures used by test developers to avoid potentially insensitive content or language
2. Review the performance of test takers of different races, gender, and ethnic groups when sample of sufficient size are available.
3. Evaluate the extent to which the performance differences may have been caused by inappropriate characteristics of test.
4. Use appropriately modified forms of tests or administration procedures for test takers with handicapping conditions.

© Copyright Virtual University of Pakistan

Lecture 7: Characteristics of Good Test

Topic 37: Characteristics of Good Test

In this session student will learn:

2. Characteristics of good test

We discussed different types of assessment or how the results are to be used, all assessments should possess certain characteristics. The most essential of these are:

- Validity
- Reliability
- Usability

Validity

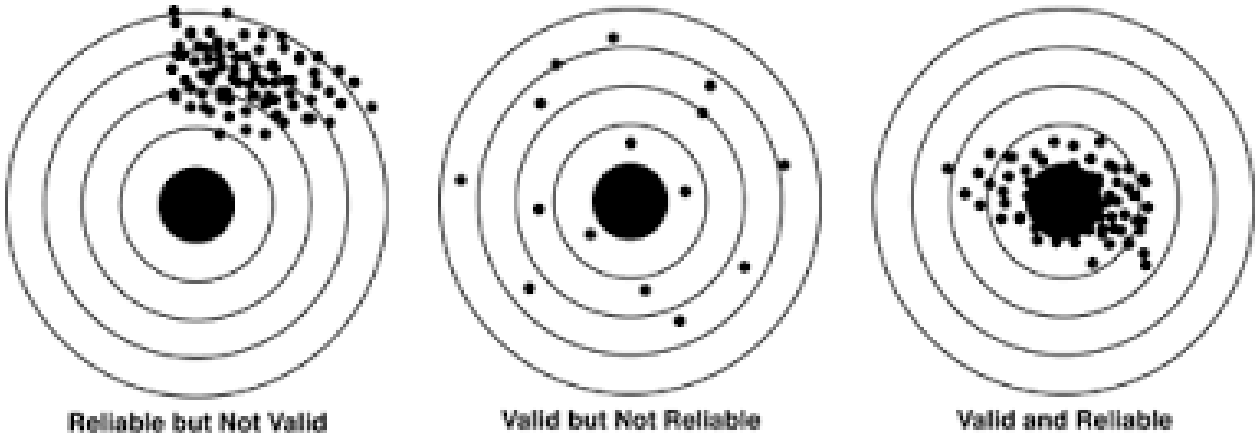
Validity is an evaluation of adequacy and appropriateness of the interpretation and uses of results. It determines if a test is measuring what it intended to measure.

Reliability

Reliability refers to the consistency of assessment results.

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.



© Copyright Virtual University of Pakistan

Usability

In addition to validity and reliability, an assessment procedure must meet certain practical requirement which includes feasibility, administration environment and availability of results for decision makers.

Topic 38: Nature of validity

In this session student will learn:

1. Nature of validity.

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 a overall evaluative judgment

© Copyright Virtual University of Pakistan

It requires an evaluation in terms of the consequences of interpretations and uses of assessment results

Topic 39: Evidences of validity: Content validity

In this session student will learn:

1. Evidences of validity
2. Concept of content validity
3. Procedure to find content validity
4. Method of ensuring content validity.

Three Evidences of Validity

Content
Construct
Criterion

Meaning

How well the sample of assessment tasks represents the domain of the tasks to be measured.

Procedure

It compares the assessment tasks to the specifications describing the task domain under consideration

Method

1. Classroom instruction determines which intended learning outcomes (objectives) are to be achieved by students
2. Achievement domain specifies and delimits a set of instructionally relevant learning tasks to be measured by an assessment
3. Instructional and assessment priorities specifies the relative importance of learning objectives to be assessed (Table of specification

© Copyright Virtual University of Pakistan

Topic 40: Evidences of validity: Construct validity

In this session student will learn:

1. Concept of construct validity
2. Procedure to find construct validity
3. Methods to ensure 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

Construct: Essay Writing

Sub-construct	**Meaning/Scope**	**Indicators**
Introduction Paragraph	It introduces the main idea, captures the interest of reader and tells why topic is important.	1. Single sentence called the thesis statement is written 2. Background information about your topic provided 3. Definitions of important terms written

© Copyright Virtual University of Pakistan

Supporting Paragraphs	Supporting paragraphs make up the main body of your essay	1. List the points about main idea of essay. 2. Write separate paragraph for each supporting point. 3. Develop each supporting point with facts, details, and examples.
Summary Paragraph	Concluding paragraph comes after you have finished developing your ideas.	1. Restate the strongest points of 2. Restate the main idea 3. Give your personal opinion or suggest a plan for action.

Method

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

1. Expert judgment
   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
   In this, we group the questions by keeping in view the responses of respondents on them.

Topic 41: Evidences of validity: Criterion validity

In this session student will learn:

1. Concept of criterion validity
2. Procedure to find criterion validity
3. Methods to ensure criterion validity

© Copyright Virtual University of Pakistan

Meaning

Demonstrates the degree of accuracy of a test by comparing it with another test, measure or procedure which has been demonstrated to be valid.

Concurrent validity

This approach allows one to show the test is valid by comparing it with an already valid test

Predictive

It involves testing a group of subjects for a certain construct, and then comparing them with results obtained at some point in the future

Procedure

Compare assessment results with another measure of performance obtained at a later date (for prediction) or with another measure of performance obtained concurrently (for estimating present status)

Method

The degree of relationship can be described more precisely by statistically correlating the two sets of scores. The resulting *correlation coefficient* provides numerical summary of relationship

Topic 42: Evidences of validity: Consequence validity

In this session student will learn:

1. Concept of consequence validity
2. Procedure to find consequence validity
3. Methods to ensure consequence validity

Meaning

How well use of assessment results accomplishes intend purposes and avoids unintended effects

Procedure

© Copyright Virtual University of Pakistan

Evaluate the effects of the use of assessment results on teachers and students. Both, the intended positive effects (e.g., increased learning) and possible unintended negative effects (e.g.,, dropout of school) need to be evaluated

Considerations

- Does the assessment artificially constrain the focus of student's study
- Does the assessment encourage or discourage exploration and creative modes of expression

Factors in Test or Assessment Itself

- Unclear directions
- Reading vocabulary and sentence structure too difficult
- Ambiguity
- Inadequate time limits (construct irrelevant variance)
- Overemphasis of easy to access aspects of domain at the expense of important, but hard to access aspects
- Test items inappropriate for the outcomes being measured
- Poorly constructed test items
- Test too short
- Improper arrangement of items
- Identifiable pattern of answers

Topic 43: Nature of reliability

In this session student will learn:

1. Nature of reliability.

© Copyright Virtual University of Pakistan

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).

Topic 44: Method of estimating reliability

In this session student will learn:

1. Method of estimating reliability

Characteristics

1. **Stability:**

Consistency over period of time

2. **Equivalence:**

Over different forms of assessment

3. **Internal consistency**:

Within the assessment itself

Determining Reliability by Correlation Methods

In determining reliability, it would be *desirable* to obtain two sets of measures *under identical conditions* and then to compare the results.

The reliability coefficient resulting from each method must be interpreted according to type of consistency being investigated

© Copyright Virtual University of Pakistan

Method to Estimate 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)

Topic 45: Method of estimating reliability: Test-retest

In this session student will learn:

1. Test-retest method

Test- Retest Method

- *Test-Retest* method is a measure of stability
- It gives the same test twice to the same group with any time interval between tests, Time interval can range from several minutes to the several years

Test- Retest	
September 25	October 15
Form A	Form A
1. Item a yes	2. Item a yes
2. Item b no	2. Item b no
3. Item c yes	3. Item c yes

© Copyright Virtual University of Pakistan

Time interval is key point in this type

- Short interval will provide inflated coefficient of reliability
- Very long interval will influence results by instability and actual changes in students over time.

Topic 46: Method of estimating reliability: Equivalent form Method

In this session student will learn:

1. Equivalent Forms method
2. Test- Retest with Equivalent Forms

Equivalent Forms method

- Equivalent Forms method is measure of equivalence
- It gives two forms of the test to the same group in close succession

September 25	September 25
Form A	Form B
1. Item a yes	2. Item d yes
2. Item b yes	2. Item e yes
3. Item c No	3. Item f No

Test- Retest with Equivalent Forms

- Test- Retest with Equivalent Forms is measure of stability and equivalence
- It gives two forms of the test to the same group with increased interval between forms

© Copyright Virtual University of Pakistan

Equivalent Forms method	
September 25	September 25
Form A	Form B
1. Item a	2. Item a
2. Item b	2. Item b
3. Item c	3. Item c
Score = 82	Score= 78
Test- Retest with Equivalent Forms	
September 25	October 15
Form A	Form B
1. Item a	2. Item a
2. Item b	2. Item b
3. Item c	3. Item c
Score = 82	Score= 74

Topic 47: Method of estimating reliability: Split Half Method

In this session student will learn:

1. Split Half Method

Split Half Method

- Split Half Method is measure of internal consistency
- It gives test once. Score two equivalent halves of test, correct correlation between halves to fit whole test by spearman –brown formula

Split Half Method		
Sum number of odd items correct	Sum number of even items correct	September 25
Item 1	Item 2	1. Item 1

© Copyright Virtual University of Pakistan

Item 3	Item 4	2. Item 2
Item 5	Item 6	3. Item 3
.	.	4. Item 4
.	.	5. Item 5
.	.	6. Item 6
.	.	
Odd	Even	
Score = 40	Score= 42	Total Score=82

Split Half Reliabilities tend to be higher than equivalent form reliabilities because split half method is based on the administration of single assessment

Topic 48: Method of estimating reliability: Kuder-Richardson Method

In this session student will learn:

1. Kuder- Richardson methods and Coefficient Alpha
2. Inter-Rater Method

Kuder- Richardson methods and Coefficient Alpha

- Kuder- Richardson methods and Coefficient Alpha is also measure of internal consistency.
- It gives test once. Score total test and apply Kuder- Richardson

As with the split half method, these formulas provide an index of internal consistency but do not require splitting the assessment in half for scoring purposes

One formula KR20 is applicable only when student responses are scored dichotomously (0 or 1). It is most useful with traditional test items scored correct or incorrect

The generalization of KR20 for assessments that have more than dichotomous, right-wrong scores is called Coefficient Alpha

© Copyright Virtual University of Pakistan

Inter-Rater Method

- Inter-Rater Method is measure of consistency of ratings
- It gives a set of students responses requiring judgmental scoring to two or more raters and have them independently score the responses

© Copyright Virtual University of Pakistan

Lecture 8: Alternate Assessment Tools

Topic 49: Anecdotal records

In this session student will learn:

1. Types of assessment tools
2. Anecdotal Record

Many outcomes in the cognitive domain, such as those pertaining to knowledge, understanding, and thinking skills, can be measured by paper pencil tests. But there are still many learning outcomes that require informal observation of natural interactions.

Types of assessment tools

Learning outcomes aspects of development can generally be assessed by

1. Observing students as they perform and describing or judging that behaviors (Anecdotal record).
2. Asking their peers about them and assessing social relationships (Peer appraisal).
3. Questioning them directly and assessing expressed interests (Self-appraisal).
4. Measuring progress by recorded work (portfolio).

Anecdotal records

Impressions gained through observation are apt to provide an incomplete and biased picture, however unless we keep an accurate record of our observations. Method to do so is called anecdotal records.

Anecdotal records are factual descriptions of meaning incidents and events that the teacher observes.

Topic 50: Effective use of Anecdotal Records

In this session student will learn:

1. Effective use of Anecdotal Record

© Copyright Virtual University of Pakistan

Effective use of Anecdotal Records

One should keep in mind the following points to use anecdotal records effectively.

1. Determine in advance what to observe but be alert of unusual behavior.
2. Analyze observational records for possible sources of bias.
3. Observe and record enough of the situation to make behavior meaningful.
4. Make record of the incident as soon after the observation is possible.
5. Limit each anecdote to a brief description of a single incident.
6. Keep the factual description of the incident and your interpretation of it separate.
7. Record both positive and negative behavioral incidents.
8. Collect a number of anecdotes on a student before drawing inferences concerning typical behavior.
9. Obtain practice in writing anecdotal records.

Topic 51: Advantages and limitations of Anecdotal Records

In this session student will learn:

1. Advantages and limitations of Anecdotal Records

Advantages of Anecdotal Records

Following are the advantages of anecdotal records.

1. It depicts actual behaviors in natural situations.
2. Facilitate gathering evidence on the events that are exceptional but significant.
3. Beneficial for students with less communication skills.

Limitations of Anecdotal Records

Following are the advantages of anecdotal records.

1. It takes long time to maintain.
2. Subjective in nature.
3. Anxiety may lead to wrong observation.

© Copyright Virtual University of Pakistan

Topic 52: Peer appraisal

In this session student will learn:

1. 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.

Techniques of peer appraisal

There are two widely used techniques in this area are:

1. Guess who technique
2. Sociometric technique

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 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.

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:

© Copyright Virtual University of Pakistan

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.

Topic 53: Portfolio

In this session student will learn:

1. Portfolio
2. Weakness and strengths of portfolio

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.

© Copyright Virtual University of Pakistan

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.

Topic 54: Purpose of Portfolio

In this session student will learn:

1. Purpose of Portfolio
2. Guidelines for portfolios entries.

Purposes of portfolios

Fundamentally two global purposes for creating portfolios of students work: for student's assessment and instruction. It can be used to showcase student's accomplishment and document the progress.

Instructional purposes:

When primary purpose is instruction, the portfolio might be used as means of:

1. Helping students develop and refine self-evaluation skills.
2. Providing teacher with more reflecting information regarding students' progress.
3. Set criteria of excellence between teacher and student.
4. Student directed conferences with parents.
5. Access to student thought process and awareness of standards.

© Copyright Virtual University of Pakistan

6. Teachers' student to communicate with different audience.

Assessment purposes:

When emphasis is on assessment it is important to distinguish between formative and summative roles of assessment.

1. It can be used for formative purposes to measure progress.
2. Basis for certifying accomplishment.
3. For system accountability mechanism.

Current accomplishment and progress

When the focus is on accomplishments, portfolios usually are limited to finished work and may cover only a relatively small period of time.

When focus is on demonstrating growth and development the time frame is longer. It will include multiple version of same work over time to measure progress.

Showcase and documentation portfolios

It contains student selected entries. It demonstrate students ability to choose his best work which demonstrates his ability to do a task.

It intended to provide evidence about breath as well as depth of learning. It needs to be more inclusive and not just limited to special strength of student.

Finished and working portfolios

It implies that work is complete for specific audience. A job application portfolio for example. It is finished product for specific audience.

Guidelines for portfolios entries

Guidelines should specify:

1. The uses that will be made of the portfolio.
2. Who will have to do it?

© Copyright Virtual University of Pakistan

3. What type of work is appropriate to include.
4. What criteria will be used in evaluating the work?
5. Should define timeline for the portfolios.
6. Minimum and maximum numbers of entries.

© Copyright Virtual University of Pakistan

Quiz # 1

The test method to compare the performance of students with other students is called: **criterion-referenced**

The purpose of formative evaluation is: **monitoring progress of students**

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

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

Persistent difficulties in learning are investigated in: **diagnostic assessment**

Assessment that monitors learning progress is the **formative assessment**

"What is your name?" Is example of which level of solo taxonomy **pre structural**

In norm referenced test the comparison is between **individuals**

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

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

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

Persistent difficulties in learning are investigated in: **diagnostic assessment**

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

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

Assessment that monitors learning progress is: **formative assistant**

Curriculum wing also known as: **national bureau of Curriculum**

Frequently used tools of summative evaluation: **test**

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

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

____ is used when we want to assess student's prior knowledge: **Placement assessment**

Assessment that monitors learning progress is: **formative assessment**

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

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

The test method to compare the performance of students with other students: **norm referenced test**

Tanveer Online Academy

Purpose of formative evaluation is ___. **Monitoring students' progress**

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

The purpose of evaluation is to make **Judgement**

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

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

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

The domain which deals with physical abilities and coordination objectives is: Psychomotor

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

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

Test designed to measure broad national objectives and have a uniform set of instructions that are adhered to during each administration. Standardized

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

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

Difficulty level of an item determines answer discrimination power

Identify the acceptable item from given values of item discrimination: 0.26

Identify the value of item difficulty for most 1.0

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

To judge the worth or value of material for a given purpose is known as: Evaluation

Identify the lowest learning level of DOK taxonomy. Extended Thinking

Identify the poor item from given values of item discrimination .17

Holistic rubric are good for evaluating answer overall performance on a task

MCQs cannot be used to assess the knowledge of illusion

Tanveer Online Academy

To assess achievement at the end of instruction is answers summative assessment

Following are usefullNess of completion questions except: easy to Mark

Which of the following decision determines prerequisite skills, degree of mastery of course goals and mode of learning? Placement Assessment

The process of obtaining a numerical description of the degree to which an individual processes particular characteristic is called: Placement Assessment

"What individual will do under natural conditions" is the: Typical Performance Assessment

Norm reference assessment and criterion reference assessment are classifications of assessment on the basis of: Method of interpreting results

The way of judging how well students are doing by looking at their work during educational process is called: Assessment

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

Bring together scientific ideas to form a unique idea is called: Synthesis

In SOLO taxonomy; classify, list, combine are key words for learning level: Multi- Structural

The assessment determines what individual will do under natural conditions is termed as: Typical Performance Assessment

In DOK taxonomy; create, synthesize, reflect are key words for learning level: Extended Thinking

Ali's writing score placed near the bottom of the class. The interpretation of this statement is: Maximum assessment

Student creatively applies knowledge to construct an overall theory is called?

Which of the following is the type of assessment by the use of classroom instruction_____. placement

Measure of student's performance against a certain criteria is: Criterion-referenced Assessment

In revised Bloom taxonomy; define, name, select are key words for level: **Remembering**

List, enlist, name, define keywords are used in _____ level of Depth of Knowledge. Recall

Reliability of a test is necessity but not sufficient condition for: Validity

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

Student can design a laboratory according to certain specification in which category of objective? Synthesis

Tanveer Online Academy

EDU431 Quiz# 2 and Quiz 3

"Types 40 words per minute without error" is an example of: Classical test theory

Which of the following is helpful for more consistent scoring across students and grades? Analytic Rubric

Guessing parameter can be measured in: IRT

Zero in raw score indicates:

A tool used to interpret and grade students' work against criteria and standards. Rubric

Identify the excellent item from given values of item discrimination. 0.50

Identify the excellent item from given values of item discrimination. 0.48

Psychometric properties of items are measured for: High Stake test

Test item is very difficult when value of difficulty level is less than? 0.3

Which of the following term used to describe the performance levels in scoring rubric? Criteria

"Types 40 words per minute without error" is an example of Classical test theory

Identify the value of item difficulty for most difficult test item. 0.1

Identify the value of item difficulty for easiest test item. 0.9 or above

Identify the acceptable item from given values of item discrimination. 0.26 or 0.29

Which of the following is not a basic step to design rubric? Give hints to a student directly

Descriptors are _____. Details description for each level of performance

Extended response items are helpful to determine: Length and complexity of response

MCQs cannot be used to assess the knowledge of: Illusion

A system of number or values assigned to work is known as: Score

Item analysis focuses to find out?

A scoring tool that teachers use to assess student learning. test

A set of criteria that provides an interpretation of the stated objectives (performance, behavior, quality).

Tanveer Online Academy

Which of the following are major threats to the validity of test score interpretations_____?

Which of the following is appropriate to measure student's abilities to synthesize when no definite answer can be pre-specified? Holistic Rubric

A set of criteria that provides an interpretation of the stated objectives (performance, behavior, quality).

When you are ready to package and reproduce the test _____. Objectives and test blueprint

Multiple test forms should be assembled so that _____. Students don't cheat

Test item cannot discriminate low achievers and high achievers when its value is lower than? 0.3

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 Rubric

Identify the good item from given values of item discrimination. 0.3 - 0.39

A rubric is _____? used for assessing a particular type of work or performance

Identify the poor item from given values of item discrimination. 0.15,0.17

There are _____ main types of rubric. Two

Scoring of essay items is a _____. Time Consuming

Holistic rubric are good for evaluating _____. overall performance on a task

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

A test form is statistically equated to another test form to make the resulting test scores _____. Comparable

Difficulty level of an item determines

The detail of each level of performance in scoring rubric is known as: Descriptors

Following are the decisions made on the basis of high stake testing EXCEPT: Measure learning process during session

Advantages of holistic rubric involves _____. Quick scoring

Test that measure learning outcomes of students is called: Achievement test

Analytical rubric works best for: Restricted Response items

Tanveer Online Academy

Which of the following should be school's professional security practices? Materials must be kept in a locked storage

Identifying relationship between two things is in: Matching Items

Feature, trait or dimension to be measured is decided in the _____ element of rubric. Criteria

To give an answer or make a judgment about something without being sure of all the facts is called: Guess

Which of the following option includes in reproducing the test? Knowing the photocopying machine

Which type of test typically designed for norm-referenced interpretations? Standardized

Discrimination power of an item is acceptable when its value ranges from? : 0.20 or higher

Which of the following are elements of scoring rubric? 1. Score 2. Criteria 3. Level of performance 4. Descriptors

The most commonly used guessing correction formula to predict and control is_____. S=R-W/N-1

Zero in raw score indicates?

When the test is ready, the next step is to _____ it. administer

Which type of test is mostly used in survey test batteries? Standardized achievement

Objective type question has advantage over essay type because such questions?
A. Are easy to prepare
B. Are easy to solve
C. Are easy to mark
D. None
Answer is = C

Discrimination value of more than 0.4 means
A. Item is good
B. Item is acceptable
C. Item is weak
D. None
Answer is = B

Test involving the construction of certain patterns are called?
A. Intelligence test
B. Performance tests

Online Academy

C. Scholastic test
D. None
Answer is = B

In multiple choice items the stem of the items should be?
A. Large
B. Small
C. Meaningful
D. None
Answer is = C

Which appropriate verb will you use to make an objective behavioral?
A. To know
B. To appreciate
C. To understand
D. To construct
Answer is = D

Objectives representing the purposes of instruction of a teacher are called?
A. Performance
B. Instructional
C. Attainment
D. None
Answer is = B

Running description of active behavior of a student as observed by the teacher is?
A. Anecdotal record
B. Autobiography
C. Interview
D. None
Answer is = A

True-False Items are also known as: Alternatives form items

A test very popular with class room teacher is?
A. True false test
B. Multiple choices
C. Matching
D. Completion test
Answer is = B

Frequently used tools of summative evaluation are?
A. Test
B. Teacher observation
C. Daily assignment
D. None

Answer is = A

The most commonly used guessing correction formula to predict and control is?
A. S=R-W
B. S=R-W/N-1
C. S=R-w/2-1
D. None
Answer is = B

The summative evaluation is?
A. Diagnostic
B. Certifying judgment
C. Continuous
D. None
Answer is = B

The difference between maximum and minimum values is?
A. Mean
B. Mode
C. Range
D. None
Answer is = C

The number of scores lying in a class interval is?
A. Mid-point
B. Quartiles
C. Class
D. Frequencies
Answer is = D

A multiple-choice question is composed of question referred as?
A. Stem
B. Distracter
C. Foil
D. Response
Answer is = A

In a norm referenced test which item is best?
A. Item difficulty is near zero
B. Item difficulty is near 100
C. Item difficulty is near 70
D. Item difficulty is near 50
Answer is = D

Which question has increasing objectivity of marking?
A. Unstructured essay

B. Structured essay
C. Short answer
D. Multiple type questions
Answer is = D

The most widely used format on standardized test in USA is?
A. Unstructured essay
B. Structured essay
C. Short answer
D. Multiple type questions
Answer is = D

Which questions are difficult to mark with reliability?
A. Unstructured essay
B. Structured essay
C. Short answer
D. Multiple type questions
Answer is = A

Projective techniques are used to measure?
A. Aptitude
B. Intelligence
C. Knowledge
D. Personality
Answer is = D

Tanveer Ch

MCQS for EDU431

Educational Assessment Mcqs For PSC Headmaster & Teaching Exams

Assessment in Education

Instrument used for measuring sample of behavior is?

A. Test

B. Measurement

C. Assessment

D. Evaluation

Answer is = A

Limited to quantitative description of pupil’s performance is?

A. Test

B. Measurement

C. Assessment

D. Evaluation

Answer is = B

The purpose of the evaluation is to make?

A. Decision

B. Prediction

C. Judgment

D. Opinion

Answer is = C

The purpose of evaluation is to make judgment about educational?

A. Quanitiy

B. Quality

C. Teme period

D. Age

Answer is = B

Evaluation that monitors learning progress is?

A. Placement evaluation

B. Formative evaluation

C. Diagnostic evaluation

D. Summative evaluation

Answer is = B

A formal and systematic procedure of getting information is?

A. Test

B. Measurement

C. Assessment

D. Evaluation

Answer is = A

The process of obtaining numerical value is?
A. Test
B. Measurement
C. Assessment
D. Evaluation
Answer is = B
A sum of questions is?
A. Test
B. Measurement
C. Assessment
D. Evaluation
Answer is = A
The first step in measurement is?
A. Decision of what to measure
B. Development of the test
C. Administering the test
D. None
Answer is = A
The purpose of formative evaluation is?
A. Decision of what to measure
B. Development of the test
C. Administering the test
D. Monitoring progress of students
Answer is = D
To assess achievement at the end of instruction is?
A. Placement Assessment
B. Formative Assessment
C. Summative Assessment
D. Diagnostic Assessment
Answer is = C
Vast of all in scope?
A. Test
B. Measurement
C. Assessment
D. Evaluation
Answer is = D
The least in scope is?
A. Test
B. Measurement

C. Assessment
D. Evaluation
Answer is = A
Permanent difficulties in learning are investigated in?
A. Placement Assessment
B. Formative Assessment
C. Summative Assessment
D. Diagnostic Assessment
Answer is = D
Broader in meaning is?
A. Aims
B. Objectives
C. Instructional objectives
D. Specific Objectives
Answer is = A
Procedures used to determine person abilities is?
A. Maximum performance test
B. Typical performance test
C. Norm performance test
D. Criterion performance test
Answer is = A
In norm referenced test the comparison is between?
A. Groups
B. Individuals
C. Areas
D. Interest
Answer is = B
In which question marking will be more reliable?
A. Completion
B. Short answer
C. Multiple choice question
D. Essay
Answer is = C
Facility value of less than 0.20 means?
A. Item is too easy
B. Item is difficult
C. Item is acceptable
D. Item is easy
Answer is = B

Objective type question have advantage over essay type because such questions?
A. Are easy to prepare
B. Are eay to solve
C. Are easy to mark
D. None
Answer is = C
Discrimination value of more than 0.4 means
A. Item is good
B. Item is acceptable
C. Item is weak
D. None
Answer is = B
Test involving the construction of certain patterns are called?
A. Intelligence test
B. Performance tests
C. Scholastic test
D. None
Answer is = B
In multiple choice items the stem of the items should be?
A. Large
B. Small
C. Meaningful
D. None
Answer is = C
Which appropriate verb will you use to make an objective behavioral?
A. To know
B. To appreciate
C. To understand
D. To construct
Answer is = D
Objectives representing the purposes of instruction of a teacher are called?
A. Performance
B. Instructional
C. Attainment
D. None
Answer is = B
Running description of active behavior of a student as observed by the teacher is?
A. Anecdotal record
B. Autobiography

C. Interview
D. None
Answer is = A
A test very popular with class room teacher is?
A. True false test
B. Multiple choices
C. Matching
D. Completion test
Answer is = B
Frequently used tools of summative evaluation are?
A. Test
B. Teacher observation
C. Daily assignment
D. None
Answer is = A
The most commonly used guessing correction formula to predict and control is?
A. S=R-W
B. S=R-W/N-1
C. S=R-w/2-1
D. None
Answer is = B
The summative evaluation is?
A. Diagnostic
B. Cerifying judgment
C. Continous
D. None
Answer is = B
The difference between maximum and minimum values is?
A. Mean
B. Mode
C. Range
D. None
Answer is = C
The number of score lying in a class interval is?
A. Mid oint
B. Quartiles
C. Class
D. Frequencies
Answer is = D

A multiple choice question is composec of question referred as?
A. Stem
B. Distracter
C. Foil
D. Response
Answer is = A
In a norm referenced test which item is best?
A. Item difficulty is near zero
B. Item difficulty is near 100
C. Item difficulty is near 70
D. Item difficulty is near 50
Answer is = D
Which question has increasing objectivity of marking?
A. Unstructured essay
B. Structured essay
C. Short answer
D. Multiple type questions
Answer is = D
The most widely used format on standardized test in USA is?
A. Unstructured essay
B. Structured essay
C. Short answer
D. Multiple type questions
Answer is = D
Which questions are difficult to mark with reliability?
A. Unstructured essay
B. Structured essay
C. Short answer
D. Multiple type questions
Answer is = A
Projective techniques are used to measure?
A. Aptitude
B. Intelligence
C. Knowledge
D. Personality
Answer is = D
Test meant for prediction on a certain criterion are called?
A. Aptitude test
B. Intelligence

C. Knowledge
D. Personality
Answer is = A
Kuder Richardson method is used to estimate?
A. Reliability
B. Validity
C. Objectivity
D. Usability
Answer is = A
Value that divides the data into two equal parts is?
A. Mean
B. Median
C. Mode
D. None
Answer is = B
The test measures what we intend to measure. This quality of the test is called?
A. Reliability
B. Validity
C. Objectivity
D. Usability
Answer is = B
The length of a test is an important factor in obtaining a representative?
A. Mean
B. Median
C. Mode
D. Sample
Answer is = D
Meidan of 1,2,4,5,2,3, is ?
A. 2
B. 5
C. 3.5
D. None
Answer is = C
The test made to compare the performance of student with the other students is called?
A. Criterion reference
B. Norm reference
C. Achievement
D. None
Answer is = B

The summative evaluation is used?
A. At the end of the program
B. At the middle of the program
C. At the start of the program
D. None
Answer is = A
The appearance of normal curve resembles with?
A. U
B. Bell
C. V
D. None
Answer is = B
The alternative name of the “table of spectification” is?
A. Test Blue Print
B. Test Construction
C. Test Administration
D. Test Scoring
Answer is = A
”table of specification” helps in?
A. Test development
B. Test Construction
C. Test Administration
D. Test Scoring
Answer is = A
The supply type test item is?
A. True / False items
B. Matching items
C. M.C.Q items
D. Completion items
Answer is = D
Alternative response item is?
A. True / False items
B. Right / wrong
C. Correct / incorrect
D. All above
Answer is = D
How many columns matching items have?
A. One
B. Two

C. Four
D. Five
Answer is = B
The item in the column for which a match is sought is?
A. Premise
B. Response
C. Destructor
D. None
Answer is = A
Identifying relationship between two things is demonstrated by?
A. True / False items
B. Matching items
C. M.C.Q items
D. Completion items
Answer is = B
The statement of problem in M.C .Qs is?
A. Premise
B. Response
C. Stem
D. None
Answer is = C
The correct option in M.C.Q is?
A. Answer
B. Premise
C. Response
D. Destructor
Answer is = A
The incorrect options in M.C.Q are?
A. Answer
B. Premise
C. Response
D. Destructor
Answer is = D
The most widely applicable test item is?
A. True / False items
B. Matching items
C. M.C.Q items
D. Completion items
Answer is = C

The type of essay item in which contents are limited is?
A. Restricted Response Questions
B. Extended Response Questions
C. Matching items
D. M.C.Q items
Answer is = A
The ability to select organize, integrate and evaluate ideas is demonstrated by?
A. Restricted Response Questions
B. Extended Response Questions
C. Matching items
D. M.C.Q items
Answer is = B
The Analysis of items is necessary in?
A. Standardized Test
B. Essay Type Test
C. Objective type test
D. Norm referenced test
Answer is = A
Which one is not the type of test of test by purpose?
A. Standardized Test
B. Essay Type Test
C. Objective type test
D. Norm referenced test
Answer is = B
The type of the test by method is?
A. Standardized Test
B. Essay Type Test
C. Objective type test
D. Norm referenced test
Answer is = C
Student’s performance is compared with other students in?
A. Standardized Test
B. Essay Type Test
C. Objective type test
D. Norm referenced test
Answer is = D
Student performance is compared with clearly defined learning tasks in?
A. Standardized Test
B. Essay Type Test

C. Criterion reverenced test
D. Norm referenced test
Answer is = C
Test that measure learning outcome of students is
A. Achievement test
B. Aptitude test
C. Criterion reverenced test
D. Norm referenced test
Answer is =A
The tests designed to predict future performance is?
A. Achievement test
B. Aptitude test
C. Criterion reverenced test
D. Norm referenced test
Answer is =B
The founder of modern intelligent tests was?
A. Alfred Binet
B. Stern
C. Gulford
D. None
Answer is = A
The formula to determine I.Q was presented by?
A. Alfred Binet
B. Stern
C. Gulford
D. None
Answer is = B
I.Q of a student having same physical and mental age will be?
A. 100
B. 120
C. 50
D. 140
Answer is = A
The I.Q of a student having twelve years mental age and tem years physical age will be?
A. 100
B. 120
C. 50
D. 140
Answer is = B

The quality of test that measures “what it claims to measure” is?
A. Validity
B. Differentiability
C. Objectivity
D. Reliability
Answer is = A
The characteristic of a test to siscriminate between high achievers and low achievers is?
A. Validity
B. Differentiability
C. Objectivity
D. Reliability
Answer is = B
If the scoring of the test is not effected by any factor, quality of test is called?
A. Validity
B. Differentiability
C. Objectivity
D. Reliability
Answer is = C
The quality of test to give same scores when administered at different occasions is?
A. Validity
B. Differentiability
C. Objectivity
D. Reliability
Answer is = D
If the sample of the question in the test is sufficiently large enough, the quality of test is?
A. Adequacy
B. Differentiability
C. Objectivity
D. Reliability
Answer is = A
The quality of test showing ease of time, cost, administration and interpretation is called?
A. Usability
B. Differentiability
C. Objectivity
D. Reliability
Answer is = A
Facility index of an item determines?
A. Ease or difficulty
B. Discrimination power

C. Objectivity
D. Reliability
Answer is = A
High and low achievers are sorted out by?
A. Ease or difficulty
B. Discrimination power
C. Objectivity
D. Reliability
Answer is = B
Test item is acceptable which its faculty index /difficulty level ranges from?
A. 30-70 %
B. 70 %
C. 30%
D. None
Answer is =A
Test item is very easy when value of faculty index/ difficulty level is higher than?
A. 30-70 %
B. 70 %
C. 30%
D. None
Answer is =B
Test item is very difficult when value of facility index/ difficulty level is less than?
A. 30-70 %
B. 70 %
C. 30%
D. None
Answer is =C
Discrimination power of an item is acceptable when its value ranges from?
A. 0.30 – 1
B. 1
C. 0.30
D. None
Answer is = A
Test item discriminates 100% when its value for discrimination is?
A. 0.30 – 1
B. 1
C. 0.30
D. None
Answer is = B

Test item cannot discriminate low achievers and high achievers when its value is lower than?

A. 0.30 – 1

B. 1

C. 0.30

D. None

Answer is = C

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 _**criteria**_ element of rubric.

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 ____M**CQ 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 __**two**__ main types of rubric
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.**

Bad distractor is that: **Attracts high achievers more than lower achievers**.

MCQs cannot be used to assess the knowledge of

Illusions

Objective type question have advantage over essay because

Easy to mark

The problem stated in the form of question or incomplete statement in MCQs is

Stem

Identify the poor item from given values of item discrimination

0.15

Which of the following are major threats to the validity of test score interpretations

Teacher's behavior

Identify the item difficulty **0.3**

Identify the excellent item from given values of item discrimination

0.48'

Difficulty level of an item determines

All above

A MCQ is composed of statement referred as
Stem

Holistic rubric are good for evaluating

Overall performance

Good item

0.31''

Most difficult test item

Lowest 0.1

An item in column for which a match is sought

Premise

Most difficult

0.2

Limited number of content area is needed to be tested in

Restricted response items

A steeper item characteristics curve shows that item is

Good discriminator

Test item cannot discriminate low achievers and high achievers when value lower than

0.3

Example of supply response item

Short answer

Not basic step to design rubric

Give hints to a student directly

Schools professional practices

Material kept in a locked storage

Identifying relationship between two thing

Matching items

True-false items are also known as

Alternative form items

Advantages of holistic rubric involves

All above

MCq cannot be used to measure the knowledge of

Illusions

To give an answer or make a judgement about something without being sure of all facts

Guess

Extended response items are helpful to determine

Length and complexity of response

- D: >0.4 – excellent

Topic 110: Item Characteristic Curve (ICC) in Item Response Theory

In this session student will learn:

1. Item Characteristic Curve (ICC)in Item Response Theory

Item Characteristic Curve (ICC) in IRT

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;

Methodological properties of an ICC

1. 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 high-ability examinees
2. The second 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
3. 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.

Item Characteristic Curve (ICC) in IRT

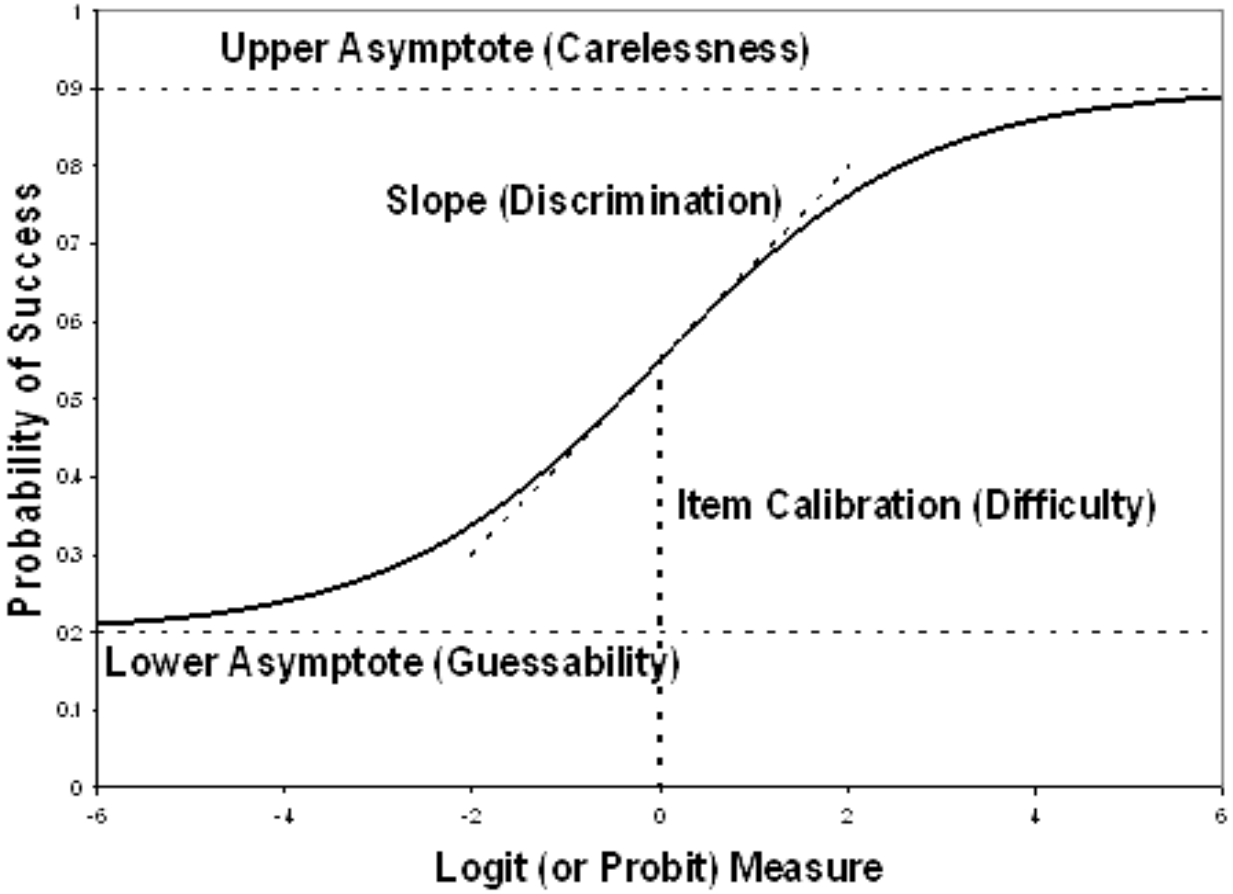
It gives a picture of:

- The item difficulty
- Discrimination power
- The probability of answering correctly by guessing

© Copyright Virtual University of Pakistan

As the item difficulties are defined in relation to ability levels, on the same scale. If we know person ability then we can predict how that person is likely to perform on an item, without administering the item to the person.

A typical item characteristic curve (ICC)



Topic 111: Item Difficulty in Item Response Theory

In this session student will learn:

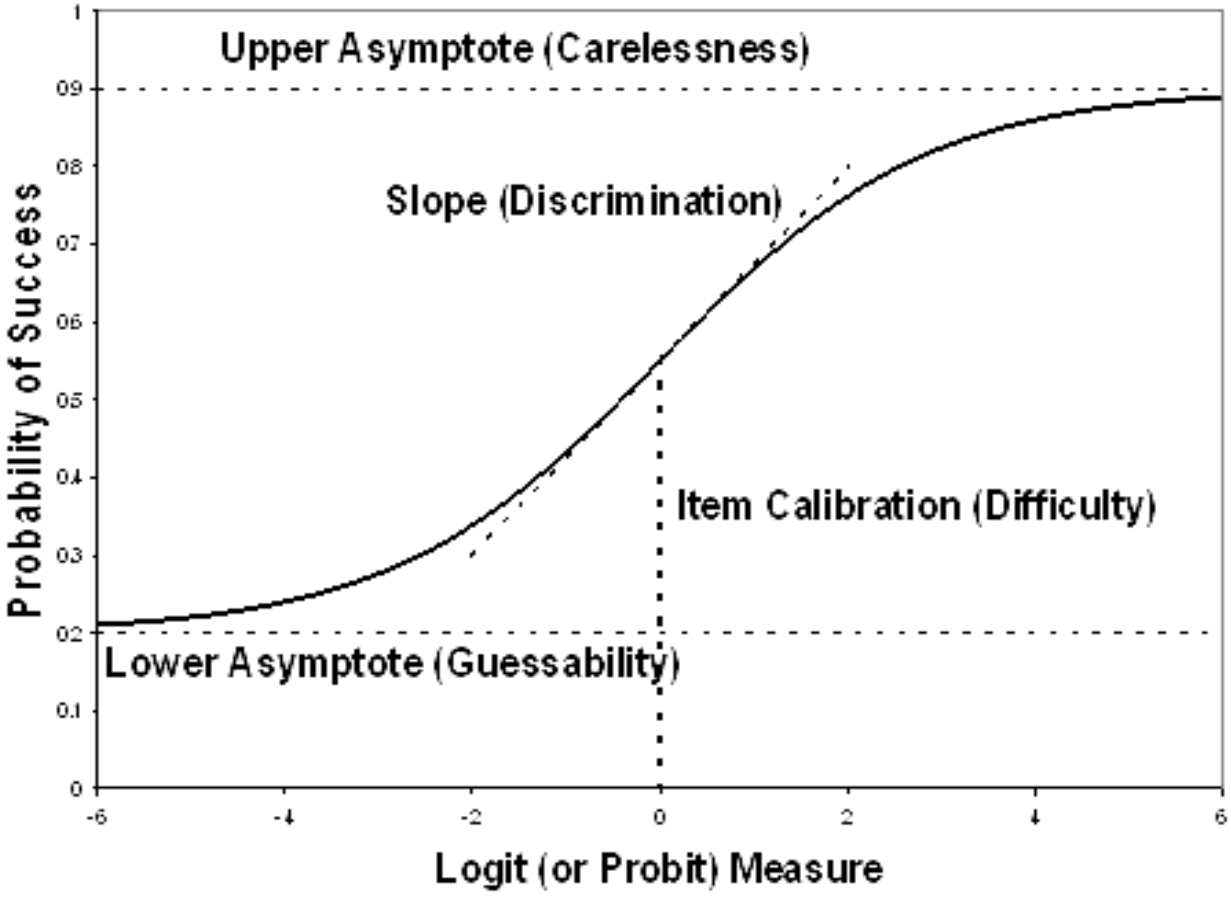
1. Item Difficulty in Item Response Theory

Item Difficulty in Item Response Theory

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.

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).

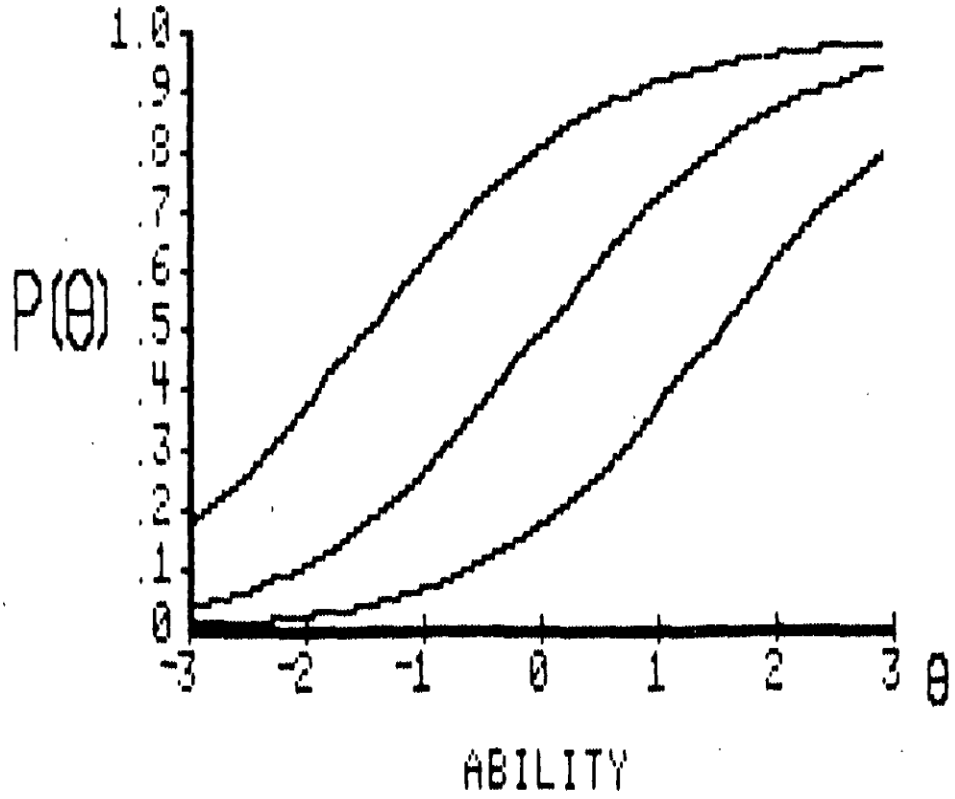
A typical item characteristic curve (ICC)



© Copyright Virtual University of Pakistan

Three item characteristic curves are presented on the same graph. All have the same level of discrimination but differ with respect to difficulty.

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.

© Copyright Virtual University of Pakistan

Even at the highest ability level shown (+3), the probability of correct response is only 0.8 for the most difficult item.

Topic 112: Item Discrimination in Item Response Theory (IRT)

In this session student will learn:

1. Item Discrimination in Item Response Theory (IRT)

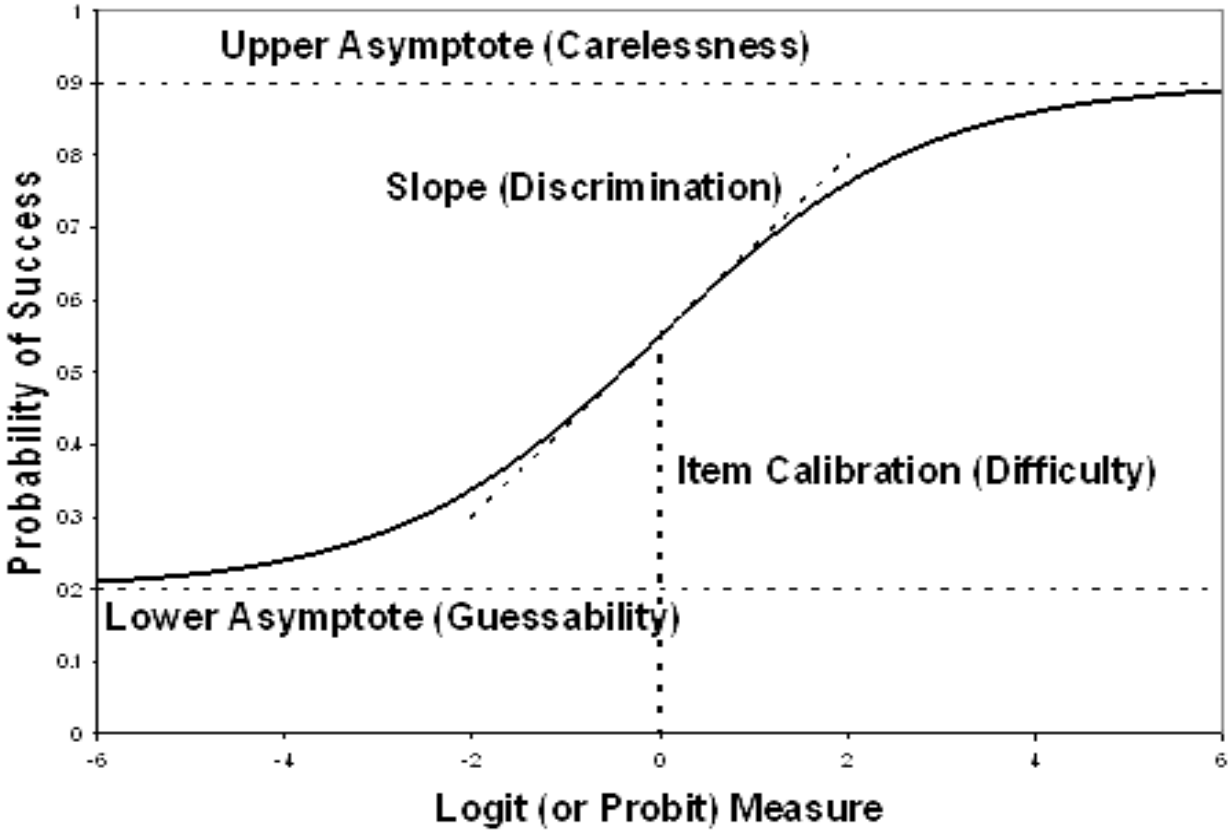
Item Discrimination in Item Response Theory (IRT)

The items on a test might also differ in terms of the degree to which they can differentiate individuals who have high trait levels from individuals who have low trait levels.

This property is reflected in the steepness of the item.

The steeper the curve, the better it discriminates.

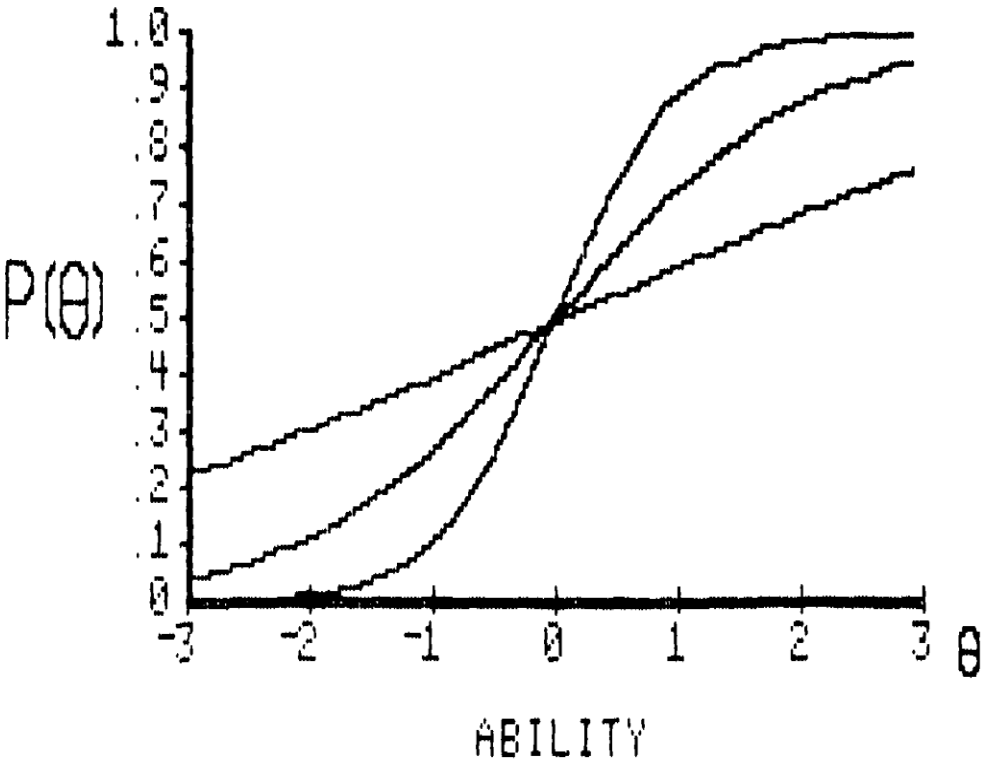
A typical item characteristic curve (ICC)



© Copyright Virtual University of Pakistan

Items with steep ICC are more discriminating as compare to relatively flatter curves.

The figure contains three item characteristic curves having the same difficulty level but differing with respect to discrimination.



The upper curve has a high level of discrimination since the curve is quite steep in the middle where the probability of correct response changes very rapidly as ability increases.

The figure contains three item characteristic curves having the same difficulty level but differing with respect to discrimination.

The middle curve represents an item with a moderate level of discrimination. The slope of this curve is much less than the previous curve and the probability of correct response changes less dramatically as the ability level increases.

The figure contains three item characteristic curves having the same difficulty level but differing with respect to discrimination.

© Copyright Virtual University of Pakistan

The probability of correct response is near zero for the lowest-ability examinees and near 1 for the highest ability examinees. The third curve represents an item with low discrimination.

The figure contains three item characteristic curves having the same difficulty level but differing with respect to discrimination.

The third curve represents an item with low discrimination. The curve has a very small slope and the probability of correct response changes slowly over the full range of abilities shown.

Topic 113: Probability of Guessing in Item Response Theory (IRT)

In this session student will learn:

1. Probability of Guessing in Item Response Theory

Probability of Guessing in Item Response Theory

Guessing means giving an answer or making a judgment about something without being sure of all the facts.

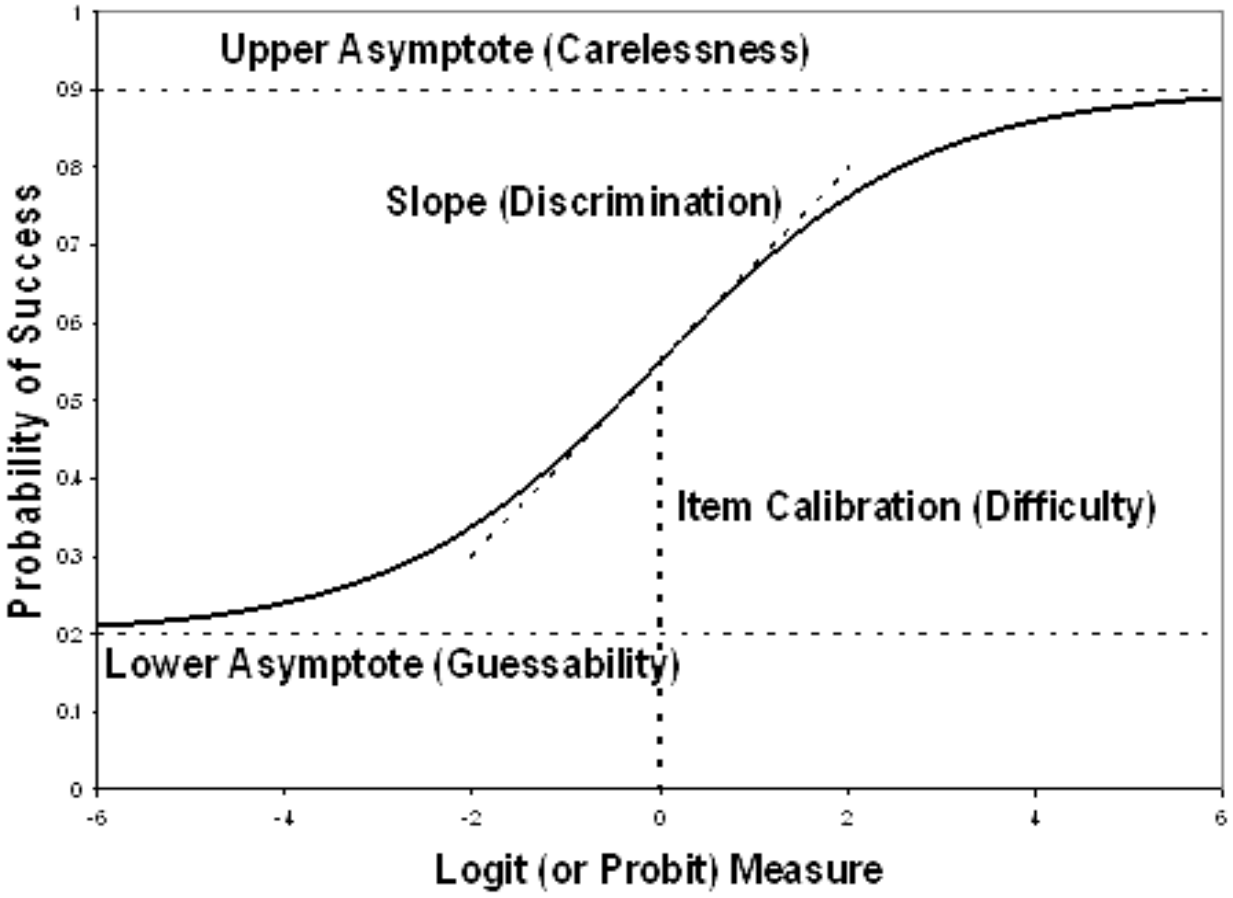
Guessing is a standard test-taking strategy presented to examinees taking a multiple choice assessment.

If test scores are based simply on the number of questions answered correctly, then a random guess increases the chance of a higher score.

In IRT this parameter of an item is also known as G (guessing) parameter which allow to detect the potential possibility of guessing in an item

A typical item characteristic curve (ICC)

© Copyright Virtual University of Pakistan



Examinees guess because they do not have adequate knowledge or ability to provide correct answer

There are two types of guessing

- Blind guessing
- Informed guessing

Blind guessing:

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

Informed guessing:

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

© Copyright Virtual University of Pakistan

Item writers should be conscious of guessing and not write item that could be prone to guessing.

IRT method of item analysis should be employed to eliminate those items prone to guessing

© Copyright Virtual University of Pakistan

Lecture 13: Administering Test

Topic 114: Concept and importance of assembling and administration of tests

In this session student will learn:

1. Concept and Importance of Assembling and Administering Test

We have spent a lot of time in learning writing instructional objectives, table of specification, selection of questions format, writing questions, aligning questions and objectives to ensure GOOD TEST.

But imagine after all this hard work, if care is not taken while assembling all good work to a test actually presented to students or test administration is not done the way it should be then whole hard work will be wasted.

Issues in test administration

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

What to do?

It is equally important to control all the factors other than test itself, to collect trustable evidence of student learning by addressing test administration issues.

The way in which the test is administered is very important to meet the goal of producing highly valid, reliable results.

Once a test is ready then next step is to administer it.

Teacher has to help the students psychologically by maintaining a positive test-taking attitude, clarifying the rules, penalties for cheating, reminding them to check their copies, minimizing distractions and giving time warnings.

© Copyright Virtual University of Pakistan

Cheating, poor testing conditions, and test anxiety, as well as errors in test scoring procedures contribute to invalid test results.

Accurate achievement data are very important for planning curriculum and instruction.

Test scores that overestimate or underestimate students' actual knowledge and skills cannot serve these important purposes.

So it is worth a little more time to properly assemble and administer a test.

Topic 115: Assembling Test: Recording items

In this session student will learn:

1. Recording items

The preparation of test items is greatly facilitated if the items are properly recorded.

Recording test items

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

Topic 116: Assembling Test: Packing the test

In this session student will learn:

1. Packing the test

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

- package the test and
- reproduce the test

Assembling Test (Packaging the Test)

Packing of test involves:

© Copyright Virtual University of Pakistan

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. Position illustration near description (if diagram given)
6. Randomness of answer key
7. Determine how students record answers (separate sheet or same sheet)
8. Providing space for test taker's details
9. Proofread the test (typographical and grammatical error)
10. Test directions

1. Grouping together items of similar format

Makes easy to understand

Save response time

Once set of instruction per format is enough

2. Arranging test items from easy to hard

Increases the possibility of good start for poor starters

Build confidence

Motivates

Reduces test anxiety

Topic 117: Assembling Test: Packing the test (Continue 1)

In this session student will learn:

1. Packing the test

Following are some more points.

© Copyright Virtual University of Pakistan

3. Space the item for easy reading

Enough space between lines and questions

Suitable answer space

Standard font type and size

4. Keeping items and options on the same page

Minimizes the likelihood of misprint

Saves respondent from unnecessary hassle

Save test time.

5. Position illustration near description (if diagram given)

Put all related questions on same page

Diagram above the question

Topic 118: Assembling Test: Packing the test (Continue 2)

In this session student will learn:

1. Packing the test

Following are some more points.

6. Randomness of answer key

Check key, equal distribution of correct answer, use all options equally

Avoid guessing

Avoid confusion

7. Determine how students record answers (separate sheet or same sheet)

Brings uniformity attempt with clarity

Prepares for standardized tests

8. Providing space for test taker's details

© Copyright Virtual University of Pakistan

Identification ensured

Makes easy to combine different parts of test (if applicable)

Topic 119: Assembling Test: Packing the test (Continue 3)

In this session student will learn:

1. Packing the test

Following are some more points.

9. Proofread the test (typographical and grammatical error)

Save time during test

Avoid clues/confusion

Increase confidence in test

10. Test directions

Conveys expectation

Help in time management

Conveys scoring policy and priorities

Topic 120: Assembling Test: Reproducing the test

In this session student will learn:

1. Reproducing the test

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

Reproducing the Test

Reproduction of test involves:

1. Knowing the photocopying machine
2. Specifying copying instructions

© Copyright Virtual University of Pakistan

3. Filing original test

1. Knowing the photocopying machine
   - Expert operator
   - Toner quality
   - Atomization facility in machine
   - Ensure uniformity of legibility, shades etc.
2. Specifying copying instructions
   - Randomly checking of every 19th copy
   - Paper size
   - Margins
   - Ordering of pages
   - Stapling
   - Packing

3. Filing original test
   - Master copy for reference and reuse
   - Use as reference in random checking
   - Keep with you during test

Topic 121: Assembling Test: Things to remember when administering test

In this session student will learn:

1. Things to remember when administering test

Test is ready, get students ready.

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

© Copyright Virtual University of Pakistan

7. Remind students to check their copies
8. Monitor students
9. Minimize distributions
10. Give time warning
11. Collect test uniformly

Maintain a positive attitude

1. Assure test is from taught
2. Students convenience and limitations addressed
3. Guide to reduce anxiety
4. Eliminate non-actors affecting achievement

Maximize achievement motivation

1. Encourage to do best, mitigate the fear
2. Highlight value of giving best
3. Reduce panic
4. Encourage serious thinking

Topic 122: Assembling Test: Things to remember when administering test (Continue 1)

In this session student will learn:

1. Things to remember when administering test

Equalize advantages (test-wise)

1. Discourage guessing
2. Discourage no answer
3. Discourage multiple answers
4. Instructions like, don't spend more time on difficult items, do easy questions first, must read after completing, etc.

Avoid surprises

Advance notice of tests, discuss test structure, and ways of preparing for test

© Copyright Virtual University of Pakistan

1. Brings out stable achievement
2. What students can perform the best

Topic 123: Assembling Test: Things to remember when administering test (Continue 2)

In this session student will learn:

1. Things to remember when administering test

Clarity the rules

1. Time limits
2. Restroom policy
3. Special requirement of students
4. Test distribution
5. Consulting others
6. Borrowing things

Rotate distribution

1. Left to right, right to left
2. Front to back, back to front
3. Multiple person distribution
4. Starting at same time

Remind students to check their copies

1. Order of pages
2. Number of pages
3. Quality of print
4. Replacement if needed before start of test
5. Ensuring recording of name and date

© Copyright Virtual University of Pakistan

Topic 124: Assembling Test: Things to remember when administering test (Continue 3)

In this session student will learn:

1. Things to remember when administering test

Monitor students (test invigilation)

1. Penalties of cheating
2. Disturbing others
3. Seating position/posture
4. Protecting test from others
5. Cheating material handing
6. Jurisdiction of invigilation staff

Minimize distributions

1. Avoid noise
2. Avoid instruction after start of test
3. Avoid in-out movement
4. No talk between invigilation staff- silence

Time warning and test collection

1. Half time
2. 2/3 time
3. 30, 15 and 5 min reminder
4. Use minimum words to announce time warning
5. Give time to close
6. Announce “stop writing”
7. Collecting test in reverse order of distribution

© Copyright Virtual University of Pakistan

Lecture 14: Scoring Test

Topic 125: Scoring Criteria: Scoring Rubric

In this session student will learn:

1. Scoring criteria
2. Scoring rubric

Scoring Criteria

Planning how responses will be scored leads to rethinking and clarifying the questions so that students have a clearer idea of what is expected.

Clear specification of scoring criteria in advance of administering essay questions can contribute to improved reliability and validity of the assessment.

Scoring 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. Scores 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".

Topic 126: Scoring Rubric for Essay Type Questions

In this session student will learn:

1. Scoring Rubric for Essay Type Questions Scoring rubric

Scoring Rubric for Essay Type Questions

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 items but by two independent raters as well.

Judgments concerning the quality of a given writing sample may vary depending upon the criteria established by the individual evaluator.

© Copyright Virtual University of Pakistan

Writing samples are just one example of performance that may be evaluated using scoring rubric.

Scoring rubric have also been used to evaluate group activities, extended projects and oral presentations.

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.

.

Topic 127: Elements of Rubric

In this session student will learn:

1. Elements of Rubric

Elements of Rubric

A rubric includes:

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

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.

© Copyright Virtual University of Pakistan

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.

Topic 128: Holistic Scoring Rubric

In this session student will learn:

2. Holistic Rubric
3. Holistic rubric for essay type questions

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.

© Copyright Virtual University of Pakistan

Holistic Rubric for Essay Type Questions

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 pre-specified.

Topic 129: Approaches to Apply Holistic Scoring Rubric

In this session student will learn:

1. Approaches to Apply Holistic Scoring Rubric

Approaches to Apply Holistic Scoring Rubric

One way to implement holistic rubric is to decide beforehand on the number of quality categories into which you will sort the student's answers.

A second better way of using the holistic method is to craft a holistic rubric, which defines the qualities of paper that belong in each category. For example, defining what an "A" paper is a "B" paper is and so on.

A third refinement is to select specimen papers, which are good examples of each scoring category. Then you can compare the student's paper with the pre-specified specimens that define each category level.

A fourth way of implementing holistic rubric is to read the answer completely and one with an-other to decide which are the best. The next best and so on.

This will result in the rough ranking of all the papers this approach of holistic rubric cannot be applicable to a large number of papers.

Among these four approaches first three are consistent with a grading philosophy of criterion referenced or absolute quality standards

While the fourth one is consistent with the norm reference or relative standard grading philosophy

© Copyright Virtual University of Pakistan

Holistic Scoring Rubric

	Levels of Rubric
Q1. Critically evaluate the various approaches to research by Identifying sound and unsound reasoning in scientific and lay contexts.	**4. Outstanding** Discerning in judging the validity of findings as warranted or not by evidence and research design. Can articulate the basic implications of identified strengths and weaknesses of methods. **3. Effective** Can differentiate sound from flawed research methods and evaluate the validity of inferences based on available evidence. **2. Adequate** Recognizes major flaws in research. Critical judgment exercised only when pressed, elicited, or when prior (closely held) assumptions are challenged. **1. Ineffective** Unable to recognize inappropriate research methods or invalid inferences from evidence. Likely to accept results more on basis of preconceived notions, prejudice or style of presentation than on the basis of a critical assessment of the evidence, concepts, and methods.

Score	Description
5	Demonstrates complete understanding of the problem. All requirements of task are included in response.
4	Demonstrates considerable understanding of the problem. All requirements of task are included.
3	Demonstrates partial understanding of the problem. Most requirements of task are included.
2	Demonstrates little understanding of the problem. Many requirements of task are missing.
1	Demonstrates no understanding of the problem.
0	No response/task not attempted.

© Copyright Virtual University of Pakistan

Advantages

1. Quick scoring
2. Provides overview of student achievement

Disadvantages

1. Does not provide detailed information.
2. May be difficult to provide one overall score.

Topic 130: Analytic Scoring Rubric

In this session student will learn:

1. Analytic Scoring Rubric
2. Analytic rubric for essay type questions

Analytic Scoring 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.

Analytic Rubric for Essay Type Questions

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

© Copyright Virtual University of Pakistan

Next, to decide the number of points award to students when they include each element.

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.

In developing analytic rubric identify the certain elements of the answer which are more appropriate to the learning objectives of a course.

When assigning number to each element, be sure the total points should match with the essay's total value in relation to the overall number of points on test.

Student will get a partial credit award gives on a partially correct answer but partial credits can increase the inconsistency in scoring and decrease the reliability of scoring process.

Crafting a partial credit scoring may be difficult after reading a few papers. The pattern of students' errors and misconceptions are emerged then craft a partial credit scoring rubric and use it to score all papers.

	Elements of questions	Scoring key
Q1.	(a) What is meiosis and mitoses?	2 points
	(b) Compare these two.	
	Write four similarities.	4 points
	Write four differences.	4points

Criterion	Beginning 1	Developing 2	Accomplished 3	Exemplary 4	Score
#1	**Description reflecting beginning level of**	**Description reflecting movement towards mastery level of**	**Description reflecting achievement of mastery level**	**Description reflecting highest level of**	

© Copyright Virtual University of Pakistan

	performance	**performance**	**of performance**	**performance**	
#2	**Description reflecting beginning level of performance**	**Description reflecting movement towards mastery level of performance**	**Description reflecting achievement of mastery level of performance**	**Description reflecting highest level of performance**	
#3	**Description reflecting beginning level of performance**	**Description reflecting movement towards mastery level of performance**	**Description reflecting achievement of mastery level of performance**	**Description reflecting highest level of performance**	

Topic 131: Advantages and disadvantages of Analytic Rubric

In this session student will learn:

1. Advantages and disadvantages of Analytic Rubric
2. Suggestion

Advantages

1. More detailed feedback.
2. Scoring more consistent across students and grades.

Disadvantages

1. Time consuming to score.

Suggestions for scoring constructed response questions

© Copyright Virtual University of Pakistan

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.

© Copyright Virtual University of Pakistan

Lecture 15: Standardized Testing

Topic 132: Standardized Achievement Test

In this session student will learn:

1. Standardized Achievement Test

Achievement Test

A standardized achievement test a fixed set if items to measure defined achievement domain, specific direction for administrating and scoring test, and norms based on representative groups of individuals.

Most published achievement test are called standardized achievement tests. These typically are norm referenced tests. Quite a few criterion referenced achievement tests are also been published.

Achievement tests are used as part if broader assessment system or alone. They provide relatively inexpensive means of measuring broad achievement goals.

Standardized achievement tests are often customized to include characteristics of both norm and criterion referenced test.

Standard content and procedure makes it possible to give an identical test to individuals in different places at different times.

Equivalent forms are included in many standardized tests, which make it possible to repeat the test without fear that the test takers will remember the answers from first testing.

Topic 133: Characteristics of Standardized Achievement Test

In this session student will learn:

1. Characteristics of Standardized Achievement Test

© Copyright Virtual University of Pakistan

Characteristics of Standardized Achievement Test

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

Topic 134: Standardized Test versus Informal Classroom Test

In this session student will learn:

1. Standardized Test versus Informal Classroom Test

Standardized Test versus Informal Classroom Test

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

1. The nature of the learning outcomes and the content measured.
2. Quality of test items
3. Reliability of the tests
4. Procedure for administering and scoring
5. Interpretation of scores

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

1. Evaluating the learning outcomes and content unique to particular class or school

© Copyright Virtual University of Pakistan

2. Evaluating student's day to day progress and their achievement on the work units or varying sizes.
3. Evaluating knowledge of current developments in the rapidly changing content areas such as science and social studies.

Topic 135: Standardized test batteries and guidelines for SAT batteries

In this session student will learn:

1. Standardized test batteries and guidelines for SAT batteries

Standardized test batteries

Standardized achievement tests are frequently used in the form of survey test batteries. A battery consists of a series of individual tests all standardized on same national sample of students. Test batteries include subjects according to educational level of students.

Guidelines for SAT batteries

Achievement test batteries focus on the basic skills measuring important outcomes of the program.

Content oriented test in basic achievement batteries have broad coverage but limited sampling in each content area and may tend to become outdated more quickly.

The selection of battery should be based on its relevance to the schools objectives. Diagnostic batteries should contain a sufficient number of test items for each type of interpretation to be made.

Topic 136: SAT in specific area, separate content oriented test, reading test

In this session student will learn:

1. SAT in specific area, separate content oriented test, reading test

© Copyright Virtual University of Pakistan

SAT in specific area

There are separate tests designed to measure achievement in specific areas. This includes tests of course content, reading tests.

Separate content oriented tests

Attention should be directed on appropriateness for particular course in which it is to be used.

Standardized tests of specific knowledge are seldom as relevant and useful as well constructed teacher made test in same area.

Reading test

Such tests commonly measure

1. Vocabulary
2. Reading comprehension
3. Rate of reading

No two reading tests are exactly alike. They differ in the material that the reader is expected to comprehend, in the specific reading skills tested and in the adequacy with which each skill is measured.

Reading survey test measure only some of the outcomes of reading instruction, the mechanics of reading is measured by diagnostic reading tests.

In addition to matching the objectives of instruction, test selection should also take into account all the possible uses to be made of the results.

Topic 137: Concept of Interpreting Test Scores

In this session student will learn:

1. Concept of Interpreting Test Scores

© Copyright Virtual University of Pakistan

Concept of Interpreting Test Scores

Test scores can be interpreted in the terms of:

1. Types of tasks that can be performed (Criterion referenced or standard based)
2. Relative position held in some reference group (norm reference)

Interpreting test scores

The properties of physical measuring scales lacks in educational measurement. A student who receives a score of zero does not have zero knowledge of that subject.

A true zero point in achievement cannot be usually established.

60 correct items on a simple vocabulary test does not have the same meaning as 60 items correct on more difficult one or any other subject or study skills. However, this arbitrary starting point prevents us from claiming that a zero indicates no achievement at all or 100 represents twice the achievement of a score of 50.

Topic 138: Method of Interpreting Test Scores

In this session student will learn:

1. Method of Interpreting Test Scores

Raw scores

A raw score is a numerical summary of a student's test performance, it is not very meaningful without further information.

1. A raw score is simply the number of points received on a test.
2. "0" in raw scores does not mean absence of trait.

If a student in our class answered 35 items correctly on arithmetic test, and therefore has a raw score of 35. This statement creates a lot of questions in our mind.

1. What is 35 mean?
2. Is that a good score?

© Copyright Virtual University of Pakistan

3. How many items were there on the test?
4. What kind of arithmetic problems were presented?
5. How difficult was test?
6. What is student's position in his class?

Answer to these or similar questions are needed to make any raw score meaningful.

Raw score can be converted into:

1. A description of specific tasks that the student can perform (Criterion reference interpretation)
2. Some type of derived scores to indicate the student's relative position in a clearly defined reference group.

Topic 139: Criterion-referenced Interpretation

In this session student will learn:

1. Criterion-referenced Interpretation

Criterion-referenced Interpretation

In case of standardized test, interpretation can only be made with reference to the constructions on which test was based by the developer.

This is primarily useful in mastery testing where a clearly defined and delimited domain of learning tasks can be most readily obtained.

Such interpretations must be made with caution because these tests were typically designed to discriminate among individuals rather than describe the specific tasks they can perform.

Criterion-referenced interpretations of test results are most meaningful when the test has been specifically designed for this purpose e.g. designing a test that measures a set of clearly stated learning tasks.

© Copyright Virtual University of Pakistan

Topic 140: Guidelines for Criterion-referenced Interpretation

In this session student will learn:

1. Guidelines for Criterion-referenced Interpretation

Guidelines for Criterion-referenced Interpretation

Are the achievement domains (objective or content clusters) homogenous, delimited, and clearly specified? If not, avoid specific descriptive statements.

Are there enough items for each type of interpretation? If not, make tentative judgment and or combine items into larger content clusters for interpretation.

In constructing the test, were the easy items omitted to increase the discrimination among individual? If so, remember that descriptions of what low achievers can do will be severely limited.

Does the test use selection type items only? If so, keep in mind that a proportion of correct answers may be based on guessing.

Do the test items provide a directly relevant measure of the objectives? If not, base the interpretation on hat the items actually measured.

Example: "ability to identify misspelled words" rather than "ability to spell". They are related but not the same process.

Topic 141: Norm-referenced Interpretation

In this session student will learn:

1. Norm-referenced Interpretation

Norm-referenced Interpretation

This interpretation tells us how an individual compares with other persons who have taken the same test e.g. ranking of scores from highest to lowest and to note where an individual's score falls.

© Copyright Virtual University of Pakistan

Standardized tests typically have been designed for norm-referenced interpretations which involves converting the raw scores to derived scores by means of table of norms.

Derived Scores

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

Criteria Most Desired in Norms

Test norms should be

1. Normal.
2. Representative.
3. Up to date.
4. Comparable.
5. Adequately described

Cautions in Interpreting Test Scores

A test score should be interpreted

1. In terms of the specific test form which it was derived.
2. In light of all of the student's relevant characteristics.
3. According to the type of decision to be made.
4. As a band of scores rather than a specific value.

A test score must be verified by supplementary evidence.

© Copyright Virtual University of Pakistan

Lecture 16: High Stake Testing

Topic 142: High Stake Testing

In this session student will learn:

1. High Stake Testing

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

Topic 143: High Stake Testing in Pakistan

In this session student will learn:

1. High Stake Testing in Pakistan

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

© Copyright Virtual University of Pakistan

3. 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.

Topic 144: Recommendations for Effective HST

In this session student will learn:

1. Recommendations for effective High Stake Testing

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

Topic 145: Preparation for Effective HST

In this session student will learn:

1. Preparation for effective High Stake Testing

Preparation for effective High Stake Testing

1. Focus on task, not on your feelings towards it
2. Inform presents and students about importance of the test.

© Copyright Virtual University of Pakistan

3. Teach test taking skills as part of regular instruction
4. As the test day approaches respond to students questions openly and directly
5. Take advantage of whatever preparation material is available

Topic 146: Institutions involve in assessment

In this session student will learn:

1. Institutions involve in assessment

Institutions involve in assessment

1. Examination commission
2. Board of intermediate and Secondary education (BISEs)
3. Board of Technical Education

Scope of Examination Bodies

1. Examination commission conduct examination at grade 5 and 8 level
2. BISEs hold SSC and HSSC annual examinations
3. Boards of Technical Education conduct examination of various diplomas and certificates etc.

Inter board committee of chairman (IBCC) is a forum to discuss matters relating to development and promotion of intermediate and secondary education and technical education in Pakistan.

For high stake examination at primary and elementary level Baluchistan and Punjab has established examination commissions

National Education Assessment System (NEAS) provides a countrywide picture of situation of education and report to federal policy makers.

Topic 147: Institutions involve in assessment

In this session student will learn:

© Copyright Virtual University of Pakistan

1. Institutions involve in assessment

Following institutions are involved in assessment.

1. National Education Assessment System (NEAS)
2. Provincial Education Assessment Centre (PEAC) in KPK and Sindh
3. Examination Commission (Punjab and Baluchistan)

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

Objectives of NEAS

1. Informing policy
2. Monitoring standards
3. Identifying correlation of achievement
4. Directing teachers' efforts and raising students' achievement

The areas centers of NEAS were established in all provinces and areas. They are still working except in Punjab which was merged in PEC.

Topic 148: National Education Assessment System

In this session student will learn:

1. National Education Assessment system

National Education Assessment system

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

NEAS was established as five years development project with the financial assistance of the World Bank and Development for International Development (DfID) in year 2003.

© Copyright Virtual University of Pakistan

NEAS is subordinate office under the ministry of Federal Education & Professional Training.

Objectives of NEAS

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

Working of NEAS

Every year NEAS conduct large scale assessment at primary and elementary level. The content areas it usually cover are reading and writing of language, mathematics and science.

Completion of four cycles of Assessment on a large sale i.e. in 2005, 2006, 2007 and 2008. The assessment results of cycle in 2016 is about come.

© Copyright Virtual University of Pakistan

EDU431 Quiz# 2 and Quiz 3

"Types 40 words per minute without error" is an example of: Classical test theory

Which of the following is helpful for more consistent scoring across students and grades? Analytic Rubric

Guessing parameter can be measured in: IRT

Zero in raw score indicates:

A tool used to interpret and grade students' work against criteria and standards. Rubric

Identify the excellent item from given values of item discrimination. 0.50

Identify the excellent item from given values of item discrimination. 0.48

Psychometric properties of items are measured for: High Stake test

Test item is very difficult when value of difficulty level is less than? 0.3

Which of the following term used to describe the performance levels in scoring rubric? Criteria

"Types 40 words per minute without error" is an example of Classical test theory

Identify the value of item difficulty for most difficult test item. 0.1

Identify the value of item difficulty for easiest test item. 0.9 or above

Identify the acceptable item from given values of item discrimination. 0.26 or 0.29

Which of the following is not a basic step to design rubric? Give hints to a student directly

Descriptors are _____. Details description for each level of performance

Extended response items are helpful to determine: Length and complexity of response

MCQs cannot be used to assess the knowledge of: Illusion

A system of number or values assigned to work is known as: Score

Item analysis focuses to find out?

A scoring tool that teachers use to assess student learning. test

A set of criteria that provides an interpretation of the stated objectives (performance, behavior, quality).

Which of the following are major threats to the validity of test score interpretations_____?

Which of the following is appropriate to measure student's abilities to synthesize when no definite answer can be pre-specified? Holistic Rubric

A set of criteria that provides an interpretation of the stated objectives (performance, behavior, quality).

When you are ready to package and reproduce the test _____. Objectives and test blueprint

Multiple test forms should be assembled so that _____. Students don't cheat

Test item cannot discriminate low achievers and high achievers when its value is lower than? 0.3

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 Rubric

Identify the good item from given values of item discrimination. 0.3 - 0.39

A rubric is _____? used for assessing a particular type of work or performance

Identify the poor item from given values of item discrimination. 0.15,0.17

There are _____ main types of rubric. Two

Scoring of essay items is a _____. Time Consuming

Holistic rubric are good for evaluating _____. overall performance on a task

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

A test form is statistically equated to another test form to make the resulting test scores _____. Comparable

Difficulty level of an item determines

The detail of each level of performance in scoring rubric is known as: Descriptors

Following are the decisions made on the basis of high stake testing EXCEPT: Measure learning process during session

Advantages of holistic rubric involves _____. Quick scoring

Test that measure learning outcomes of students is called: Achievement test

Analytical rubric works best for: Restricted Response items

Which of the following should be school's professional security practices? Materials must be kept in a locked storage

Identifying relationship between two things is in: Matching Items

Feature, trait or dimension to be measured is decided in the _____ element of rubric. Criteria

To give an answer or make a judgment about something without being sure of all the facts is called: Guess

Which of the following option includes in reproducing the test? Knowing the photocopying machine

Which type of test typically designed for norm-referenced interpretations? Standardized

Discrimination power of an item is acceptable when its value ranges from? : 0.20 or higher

Which of the following are elements of scoring rubric? 1. Score 2. Criteria 3. Level of performance 4. Descriptors

The most commonly used guessing correction formula to predict and control is_____. S=R-W/N-1

Zero in raw score indicates?

When the test is ready, the next step is to _____ it. administer

Which type of test is mostly used in survey test batteries? Standardized achievement

Objective type question has advantage over essay type because such questions?
A. Are easy to prepare
B. Are easy to solve
C. Are easy to mark
D. None
Answer is = C

Discrimination value of more than 0.4 means
A. Item is good
B. Item is acceptable
C. Item is weak
D. None
Answer is = B

Test involving the construction of certain patterns are called?
A. Intelligence test
B. Performance tests
C. Scholastic test
D. None

Answer is = B

In multiple choice items the stem of the items should be?
A. Large
B. Small
C. Meaningful
D. None
Answer is = C

Which appropriate verb will you use to make an objective behavioral?
A. To know
B. To appreciate
C. To understand
D. To construct
Answer is = D

Objectives representing the purposes of instruction of a teacher are called?
A. Performance
B. Instructional
C. Attainment
D. None
Answer is = B

Running description of active behavior of a student as observed by the teacher is?
A. Anecdotal record
B. Autobiography
C. Interview
D. None
Answer is = A

True-False Items are also known as: Alternatives form items

A test very popular with class room teacher is?
A. True false test
B. Multiple choices
C. Matching
D. Completion test
Answer is = B

Frequently used tools of summative evaluation are?
A. Test
B. Teacher observation
C. Daily assignment
D. None
Answer is = A

The most commonly used guessing correction formula to predict and control is?
A. S=R-W
B. S=R-W/N-1
C. S=R-w/2-1
D. None
Answer is = B

The summative evaluation is?
A. Diagnostic
B. Certifying judgment
C. Continuous
D. None
Answer is = B

The difference between maximum and minimum values is?
A. Mean
B. Mode
C. Range
D. None
Answer is = C

The number of scores lying in a class interval is?
A. Mid-point
B. Quartiles
C. Class
D. Frequencies
Answer is = D

A multiple-choice question is composed of question referred as?
A. Stem
B. Distracter
C. Foil
D. Response
Answer is = A

In a norm referenced test which item is best?
A. Item difficulty is near zero
B. Item difficulty is near 100
C. Item difficulty is near 70
D. Item difficulty is near 50
Answer is = D

Which question has increasing objectivity of marking?
A. Unstructured essay
B. Structured essay
C. Short answer

D. Multiple type questions
Answer is = D

The most widely used format on standardized test in USA is?
A. Unstructured essay
B. Structured essay
C. Short answer
D. Multiple type questions
Answer is = D

Which questions are difficult to mark with reliability?
A. Unstructured essay
B. Structured essay
C. Short answer
D. Multiple type questions
Answer is = A

Projective techniques are used to measure?
A. Aptitude
B. Intelligence
C. Knowledge
D. Personality
Answer is = D

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 _**criteria**_ element of rubric.

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 ____M**CQ 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 __**two**__ main types of rubric
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.**

Bad distractor is that: **Attracts high achievers more than lower achievers**.

MCQs cannot be used to assess the knowledge of

Illusions

Objective type question have advantage over essay because

Easy to mark

The problem stated in the form of question or incomplete statement in MCQs is

Stem

Identify the poor item from given values of item discrimination

0.15

Which of the following are major threats to the validity of test score interpretations

Teacher's behavior

Identify the item difficulty **0.3**

Identify the excellent item from given values of item discrimination

0.48'

Difficulty level of an item determines

All above

A MCQ is composed of statement referred as
Stem

Holistic rubric are good for evaluating

Overall performance

Good item

0.31''

Most difficult test item

Lowest 0.1

An item in column for which a match is sought

Premise

Most difficult

0.2

Limited number of content area is needed to be tested in

Restricted response items

A steeper item characteristics curve shows that item is

Good discriminator

Test item cannot discriminate low achievers and high achievers when value lower than

0.3

Example of supply response item

Short answer

Not basic step to design rubric

Give hints to a student directly

Schools professional practices

Material kept in a locked storage

Identifying relationship between two thing

Matching items

True-false items are also known as

Alternative form items

Advantages of holistic rubric involves

All above

MCq cannot be used to measure the knowledge of

Illusions

To give an answer or make a judgement about something without being sure of all facts

Guess

Extended response items are helpful to determine

Length and complexity of response

1. Assessment that monitors Learning Progress is ----------
   - ✓ Formative Assessment
2. ---------- is used when we want to assess student ‘s prior knowledge
   - ✓ Placement assessment
3. Measurement and Evaluation can be used to ----------
   - ✓ All of Above
4. The Summative evaluation is ----------
   - ✓ Certifying Judgment
5. Persistent Difficulties in Learning are Investigated in ----------
   - ✓ Diagnostic Assessment
6. Who was the founder of Pakistan is the question of --------- in Bloom taxonomy of cognitive domain?
   - ✓ Knowledge
7. “What is your name?” is the example of which level of Solo taxonomy ----------
   - ✓ Pre-structural
8. ---------- determines learning progress, provides feedback to reinforce learning, and correct learning errors
   - ✓ Formative Assessment
9. The purpose of formative learning is ----------
   - ✓ Monitoring progress of student
10. The purpose of measurement in education is to ----------
    - ✓ Collect information for Decision Making???
11. The purpose of the evaluation is to make ----------
    - ✓ Judgment
12. “What individual will do under natural condition “is the ----------.
    - ✓ Typical Performance Assessment
13. Multiple choice question, matching exercise, fill in the blanks and true false is the example of ------
    - ✓ Fixed Choice Assessment
14. Which of the following is the type of assessment by the use of classroom instruction? ---------
    - ✓ Placement Assessment
15. Curriculum Wing is also known as ----------
    - ✓ National Bureau of Curriculum
16. List, enlist, name, define keywords are used in -------- level of Depth of Knowledge
    - ✓ Recall
17. Frequently used tools of summative evaluation are ----------.
    - ✓ Test
18. Explain the process of water cycle is the question of -------- level in bloom taxonomy of cognitive domain
    - ✓ Comprehension
19. Student’s Performance is compared with other students in ----------.
    - ✓ Norm referenced test
20. The test made to compare the performance of student with the other students is called -----.
    - ✓ Norm reference
21. To Assess achievement at the end of instruction is ----------.
    - ✓ Summative Assessment
22. Permanent difficulties in learning are investigated in----------
    - ✓ Diagnostic Assessment

23. The purpose of evaluation is to make judgment about educational----------.
    - ✓ Quality
24. In norm referenced test the comparison is between ---------.
    - ✓ Individuals
25. How many domains of Bloom Taxonomy ……..?
    - ✓ 3
26. Procedures used to determine person best abilities are…….?
    - ✓ Maximum performance test
27. The type of essay item in which content are limited is----------?
    - ✓ Restricted response question
28. Identifying relationship between two things is done in --------?
    - ✓ Matching item
29. How many column matchings items have...?
    - ✓ Two
30. Cognitive domain of Blooms taxonomy are -------
    - ✓ 6
31. 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
32. SOLO stands for:
    - ✓ Structure of Observed Learning Outcomes
33. The process of obtaining a numerical description of the degree to which an individual processes particular characteristic is called:
    - ✓ Measurement
34. Measure of student's performance against a certain criteria is :
    - ✓ Criterion-referenced
35. The way of judging how well students are doing by looking at their work during educational process is called
    - ✓ Evaluation??
36. Hina Exceeded 85% of the sixth graders on the mathematics test. The interpretation of this statement is:
    - ✓ Norm Referenced
37. The Assessment determines what individual can do when performing at their best
    - ✓ Maximum Performance Assessment
38. In DOK taxonomy, create, synthesize, reflect are key words for learning level
    - ✓ Extended thinking
39. Which type of assessment determines the student's prerequisite skills to begin instruction
    - ✓ Placement
40. Identify the assessment used to find the root cause of continuous or constant problem in students learning
    - ✓ Diagnostic
41. Cognitive domain was divided into sub subgroups by bloom in
    - ✓ 1956
42. Degree to which a test measures an intended content area is called
    - ✓ construct validity
43. Which type of assessment measures entry level performance of student?
    - ✓ Placement

44. The spearman-brown formula is used to estimate:
    - ✓ Split-half
45. To judge the worth or value of material for a given purpose is known as
    - ✓ Evaluation
46. to use the previous learned material in new situation is called
    - ✓ Application
47. Which of the following decision determines prerequisite skills, degree of mastery of course goals and modes of learning?
    - ✓ Placement
48. Test designed to measure the number of items and individual can attempt correctly in a given time is referred as:
    - ✓ Speed
49. The degree to which an instrument measures what it is supposed to measured is called
    - ✓ Validity
50. Monitoring the outcomes with reference to the objectives is called:
    - ✓ Evaluation
51. Which type of Assessment is designed to measure whether the students have mastered the skills presented in instruction?
    - ✓ Summative Assessment
52. 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
53. How well the sample of assessment tasks represents the domain of the tasks to be measured is known as
    - ✓ Content validity
54. Reliability over different forms of assessment is known as
    - ✓ Equivalence
55. Which of the following provides a framework of categories with different hierarchical levels of outcomes?
    - ✓ Taxonomy
56. The procedure used to gain information about student learning and the formation of value judgement concerning learning progress is called:
    - ✓ Evaluation
57. Reliability within an assessment itself is known as:
    - ✓ Internal Consistency
58. ---------- refers to the assessment of participants, and summarize their development at the particular time
    - ✓ Summative assessment
59. In SOLO taxonomy; identify, memorize, are keywords for learning level:
    - ✓ Uni-Structural
60. Student creatively applies knowledge to construct an overall theory called:
    - ✓ Application
61. 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
62. “What is your country name?” is example of which level of SOLO taxonomy.
    - ✓ Pre-structural

63. The domain which deals with physical abilities and coordination objectives is
    - ✓ Psychomotor
64. The Degree to which a test measures an intended hypothetical construct is called:
    - ✓ Construct validity
65. What kind of assessment, permanent difficulties in learning are investigated ?
    - ✓ Diagnostic
66. The alternative name of the “table of specification” is ?
    - ✓ Test Blue Print
67. Which of the following are three domains of Bloom’s Taxonomy?
    - ✓ Cognitive, affective, psychomotor
68. Measure of performance which interprets an individual’s relative standing in some known group is:
    - ✓ Norm referenced
69. Which of following learning level of DOK taxonomy demands extended use of higher order thinking processes such as synthesis, reflection, assessment
    - ✓ Extended thinking
70. Identify the lowest learning level of SOLO taxonomy.
    - ✓ Pre-Structural
71. The purpose of the evaluation is to make judgment about educational ----------.
    - ✓ Quality
72. The quality of test to give same scores when administered at different occasions is called
    - ✓ Reliability
73. To assess achievement at the end of instruction is ----------
    - ✓ Summative assessment
74. Which of the following test with time limit strict that no one is supposed to complete all items?
    - ✓ Speed
75. In what kind of assessment, recurrent difficulties in learning are investigated?
    - ✓ Diagnostic Assessment
76. In SOLO taxonomy: classify, list, combine are keywords for learning level:
    - ✓ Multi- Structural
77. The Scores of students in papers are
    - ✓ Assessment??
78. Aptitude test is example of which type of Assessment
    - ✓ Placement Assessment
79. In DOK taxonomy; argue, critique, formulate are key words for learning level:
    - ✓ Strategic Thinking

Edu 431 mid MCQ

1. ----------is used when we want to assess student's prior knowledge

A) Placement assessment

b) Formative Assessment

c) Summative assessment

d) Diagnostic Assessment

2. Cognitive domain of Bloom Taxonomy are

a) 3

b) 4

c) 6

d) 8

3. Frequently used tools of summative evaluation are---------?

a) Teacher observation

b) Daily assignment

c) Test

d) None

4. What individual will do under natural conditions is the

a) Maximum Performance Assessment
b) Norm-referenced Assessment
c)Criterion-referenced Assessment
d) Typical Performance Assessment

03047838705 (M.Javeed)

5. To assess achievement at the end of the instruction is……..

a) Summative assessment

b) Placement assessment

c) Formative Assessment

d) Diagnostic Assessment

6. Student's performance is compared with other student's in_____

a) Standardized Test

b) Essay type Test

c) Objective type test

d) Norm referenced test

7. What is your name? is the example of which level of SOLO taxonomy

a) Pre-structural
b) Uni Structural
c) Multi- Structural
d) Relational

8. The purpose of the evaluation is to make ------
a) Decision
b) Prediction
c) Judgment
d) Opinion

9. Explain the process of water cycle is the question of -------- level in Blooms taxonomy of cognitive domain.
a) Knowledge
b) Comprehension
c) Analysis
d) Application

03047838705 (M.Javeed)

10. Permanent difficulties in learning are investigated in----------

a) Summative assessment

b) Placement assessment

c) Formative Assessment

d) Diagnostic Assessment

11. ----------determines learning progress, provides feedback to reinforce learning, and correct learning errors

A) Placement assessment

b) Formative Assessment

c) Summative assessment

d) Diagnostic Assessment

12. The purpose of formative Evaluation is----------

a) Decision of what to measure
b) Development of the test
c) Administering the test
d) Monitoring progress of the students

13. Curriculum Wing also known as

a) **National Bureau of Curriculum**
b) Curriculum authority development
c) Development of curriculum development
d) None of the above

14. Assessment that monitors learning progress ----------

A) Placement assessment

b) Formative Assessment

c) Diagnostic Assessment

d) Summative assessment

03078838705 (M.Javeed)

15. The purpose of measurement in education is to ----------

a) Make decision about students

b) Determine if objectives are being met

c) Assess teacher skills

d) Collect information for decision making

16. The summative evaluation is……

a) Diagnostic

b) Certifying judgment

c) Continuous

d) None

17. The purpose of evaluation is to make judgment about educational----------

a) Quantity

b) Quality

c) Time period

d) Age

18. Multiple choice question, matching exercise, fill in the blanks and true false is the example of

a) Typical performance assessment

b) Fixed choice assessment

c) Maximum performance assessment

d) Formative assessment

03047838705 (M.Javeed)

19. The test made to compare of student with other students is called……..

a) Criterion reference

b) Norm reference

c) Achievement

d) None

20. List, enlist, name, define keywords are used in …… level of Depth of knowledge

a) Recall

b) skill/concept

c) Strategic thinking

d) Extended thinking

21. In norm referenced test the comparison is between ---------

a) Groups

b) Individuals

c) Areas

d) Interest

22. Who was the founder of Pakistan is the question of----------in Blooms taxonomy of cognitive

a) Knowledge

b) Comprehension

c) Application

d) Synthesis

(03047838705 (M.Javeed))

23. Measurement and evaluation can be used to ……..

a) Motivate students

b) Help develop the scope and sequence of teaching

c) Assess the effectiveness of learning activities

d) all of the above

24. How many domains of Bloom Taxonomy are……..

a)1

b) 2

c) 3

d) 4

25. Procedures used to determine person best abilities are…….

a) Maximum performance test

b) Typical performance test

c) Norm reference test

d) Criterion reference test

26. Which of following is the type of assessment by the use of classroom instruction……..

a) Complex performance assessment

b) Placement assessment

c) Fixed choice assessment

d) Norm reference assessment

03047838705 (M.Javeed)

27. In …….. Federal government appointed university grant commission

a) 1960

b) 1964

c) 1976

d) 1980

28. The type of essay item in which content are limited is:

a) Restricted response question

b) Extended response question

c) Matching item

d) M.C.Q items

29. Identifying relationship between two things is done in

a) True/false items

b) Matching item

c) M.C.Q items

d) Completion item

30. How many column matchings items have...?

a) One

b) Two

c) three

d) Four

31. The incorrect option in MCQ are called

a) Answer

b) Premise

03047838705 (M.Javeed)

c) Response

d) Distractor

32. The item in the column for which a match is sought is called

a) Premise

b) Response

c) Distractor

d) Answer key

33. The correct option in MCQ is called

a) Answer

b) Premise

c) Response

d) Distractor

34. Alternative response item is a type of ……..

a) Matching List

b) Completion List

c) Correct/Incorrect

d) All

35. In which type of question marking will be more reliable?

a) Completion

b) Short answer

c) Multiple choice question

d) Essay

36. In multiple choice items the stem of the items should be

03047838705 (M.Javeed)

a) Large

b) Small

c) Meaningful

d) None of the above

37. A multiple choice question is composed of question referred as

a) Stem

b) Distractor

c) Foil

d) Response

1. Instrument used for measuring sample of behavior is?

A. Test

B. Measurement

C. Assessment

D. Evaluation

2. Limited to quantitative description of pupil's performance is?

A. Test

B. Measurement

C. Assessment

D. Evaluation

3. The purpose of the evaluation is to make?

A. Decision

B. Prediction

C. Judgment

D. Opinion

4. The purpose of evaluation is to make judgment about educational?

A. Quantity

B. Quality

C. Time period

03047838705 (M.Javeed)

D. Age

5. Evaluation that monitors learning progress is?

A. Placement evaluation

B. Formative evaluation

C. Diagnostic evaluation

D. Summative evaluation

6. A formal and systematic procedure of getting information is?

A. Test

B. Measurement

C. Assessment

D. Evaluation

7. The process of obtaining numerical value is?

A. Test

B. Measurement

C. Assessment

D. Evaluation

Answer is = B

8. A sum of questions is?

A. Test

B. Measurement

C. Assessment

D. Evaluation

9. The first step in measurement is?

A. Decision of what to measure

B. Development of the test

C. Administering the test

D. None

10. The purpose of formative evaluation is?

A. Decision of what to measure

B. Development of the test

C. Administering the test

D. Monitoring progress of students

03047838705 (M.Javeed)

11. To assess achievement at the end of instruction is?

A. Placement Assessment

B. Formative Assessment

C. Summative Assessment

D. Diagnostic Assessment

12. Vast of all in scope?

A. Test

B. Measurement

C. Assessment

D. Evaluation

13. The least in scope is?

A. Test

B. Measurement

C. Assessment

D. Evaluation

14. Permanent difficulties in learning are investigated in?

A. Placement Assessment

B. Formative Assessment

C. Summative Assessment

D. Diagnostic Assessment

15. Broader in meaning is?

A. Aims

B. Objectives

C. Instructional objectives

D. Specific Objectives

16. Procedures used to determine person abilities is?

A. Maximum performance test

B. Typical performance test

C. Norm performance test

D. Criterion performance test

17. In norm referenced test the comparison is between?

A. Groups

B. Individuals

03047838705 (M.Javeed)

C. Areas
D. Interest
18. In which question marking will be more reliable?
A. Completion
B. Short answer
C. Multiple choice question
D. Essay
19. Facility value of less than 0.20 means?
A. Item is too easy
B. Item is difficult
C. Item is acceptable
D. Item is easy
20. Objective type question have advantage over essay type because such questions?
A. Are easy to prepare
B. Are easy to solve
C. Are easy to mark
D. None
21. Discrimination value of more than 0.4 means
A. Item is good
B. Item is acceptable
C. Item is weak
D. None
22. Test involving the construction of certain patterns are called?
A. Intelligence test
B. Performance tests
C. Scholastic test
D. None
23. In multiple choice items the stem of the items should be?
A. Large
B. Small
C. Meaningful
D. None

03047838705 (M.Javeed)

24. Which appropriate verb will you use to make an objective behavioral?

A. To know

B. To appreciate

C. To understand

D. To construct

25. Objectives representing the purposes of instruction of a teacher are called?

A. Performance

B. Instructional

C. Attainment

D. None

26. Running description of active behavior of a student as observed by the teacher is?

A. Anecdotal record

B. Autobiography

C. Interview

D. None

27. A test very popular with class room teacher is?

A. True false test

B. Multiple choices

C. Matching

D. Completion test

28. Frequently used tools of summative evaluation are?

A. Test

B. Teacher observation

C. Daily assignment

D. None

29. The most commonly used guessing correction formula to predict and control is?

A. S=R-W

B. S=R-W/N-1

C. S=R-w/2-1

D. None

03047838705 (M.Javeed)

30. The summative evaluation is?

A. Diagnostic

B. Certifying judgment

C. Continuous

D. None

31. The difference between maximum and minimum values is?

A. Mean

B. Mode

C. Range

D. None

32. The number of score lying in a class interval is?

A. Mid-point

B. Quartiles

C. Class

D. Frequencies

33. A multiple choice question is composed of question referred as?

A. Stem

B. Distracter

C. Foil

D. Response

34. In a norm referenced test which item is best?

A. Item difficulty is near zero

B. Item difficulty is near 100

C. Item difficulty is near 70

D. Item difficulty is near 50

35. Which question has increasing objectivity of marking?

A. Unstructured essay

B. Structured essay

C. Short answer

D. Multiple type questions

03047838705 (M.Javeed)

36. The most widely used format on standardized test in USA is?

A. Unstructured essay

B. Structured essay

C. Short answer

D. Multiple type questions

37. Which questions are difficult to mark with reliability?

A. Unstructured essay

B. Structured essay

C. Short answer

D. Multiple type questions

38. Projective techniques are used to measure?

A. Aptitude

B. Intelligence

C. Knowledge

D. Personality

39. Test meant for prediction on a certain criterion are called?

A. Aptitude test

B. Intelligence

C. Knowledge

D. Personality

40. Kuder Richardson method is used to estimate?

A. Reliability

B. Validity

C. Objectivity

D. Usability

41. Value that divides the data into two equal parts is?

A. Mean

B. Median

C. Mode

D. None

03047838705 (M.Javeed)

42. The test measures what we intend to measure. This quality of the test is called?

A. Reliability

B. Validity

C. Objectivity

D. Usability

43 The length of a test is an important factor in obtaining a representative?

A. Mean

B. Median

C. Mode

D. Sample

44. Median of 1,2,4,5,2,3, is ?

A. 2

B. 5

C. 3.5

D. None

45. The test made to compare the performance of student with the other students is called?

A. Criterion reference

B. Norm reference

C. Achievement

D. None

46. The summative evaluation is used?

A. At the end of the program

B. At the middle of the program

C. At the start of the program

D. None

47.The appearance of normal curve resembles with?

A. U

B. Bell

C. V

D. None

03047838705 (M.Javeed)

48. The alternative name of the “table of spectification” is?

A. Test Blue Print

B. Test Construction

C. Test Administration

D. Test Scoring

49. ”table of specification” helps in?

A. Test development

B. Test Construction

C. Test Administration

D. Test Scoring

50. The supply type test item is?

A. True / False items

B. Matching items

C. M.C.Q items

D. Completion items

51. Alternative response item is?

A. True / False items

B. Right / wrong

C. Correct / incorrect

D. All above

52. How many columns matching items have?

A. One

B. Two

C. Four

D. Five

53 The item in the column for which a match is sought is?

A. Premise

B. Response

C. Destructor

D. None

03047838705 (M.Javeed)

54. Identifying relationship between two things is demonstrated by?

A. True / False items

B. Matching items

C. M.C.Q items

D. Completion items

55. The statement of problem in M.C .Qs is?

A. Premise

B. Response

C. Stem

D. None

56. The correct option in M.C.Q is?

A. Answer

B. Premise

C. Response

D. Destructor

57. The incorrect options in M.C.Q are?

A. Answer

B. Premise

C. Response

D. Destructor

58. The most widely applicable test item is?

A. True / False items

B. Matching items

C. M.C.Q items

D. Completion items

59. The type of essay item in which contents are limited is?

A. Restricted Response Questions

B. Extended Response Questions

C. Matching items

D. M.C.Q items

03047838705 (M.Javeed)

60. The ability to select organize, integrate and evaluate ideas is demonstrated by?

A. Restricted Response Questions

B. Extended Response Questions

C. Matching items

D. M.C.Q items

61. The Analysis of items is necessary in?

A. Standardized Test

B. Essay Type Test

C. Objective type test

D. Norm referenced test

62. Which one is not the type of test of test by purpose?

A. Standardized Test

B. Essay Type Test

C. Objective type test

D. Norm referenced test

63. The type of the test by method is?

A. Standardized Test

B. Essay Type Test

C. Objective type test

D. Norm referenced test

64. Student's performance is compared with other students in?

A. Standardized Test

B. Essay Type Test

C. Objective type test

D. Norm referenced test

65. Student performance is compared with clearly defined learning tasks in?

A. Standardized Test

B. Essay Type Test

C. Criterion reverenced test

D. Norm referenced test

03047838705 (M.Javeed)

66. Test that measure learning outcome of students is

A. Achievement test

B. Aptitude test

C. Criterion reverenced test

D. Norm referenced test

66.The tests designed to predict future performance is?

A. Achievement test

B. Aptitude test

C. Criterion reverenced test

D. Norm referenced test

67. The founder of modern intelligent tests was?

A. Alfred Binet

B. Stern

C. Gulford

D. None

68. The formula to determine I.Q was presented by?

A. Alfred Binet

B. Stern

C. Gulford

D. None

69. I.Q of a student having same physical and mental age will be?

A. 100

B. 120

C. 50

D. 140

70 The I.Q of a student having twelve years mental age and tem years physical age will be?

A. 100

B. 120

C. 50

D. 140

03047838705 (M.Javeed)

71. The quality of test that measures “what it claims to measure” is?

A. Validity

B. Differentiability

C. Objectivity

D. Reliability

72. The characteristic of a test to discriminate between high achievers and low achievers is?

A. Validity

B. Differentiability

C. Objectivity

D. Reliability

73. If the scoring of the test is not effected by any factor, quality of test is called?

A. Validity

B. Differentiability

C. Objectivity

D. Reliability

74. The quality of test to give same scores when administered at different occasions is?

A. Validity

B. Differentiability

C. Objectivity

D. Reliability

75. If the sample of the question in the test is sufficiently large enough, the quality of test is?

A. Adequacy

B. Differentiability

C. Objectivity

D. Reliability

03047838705 (M.Javeed)

76. The quality of test showing ease of time, cost, administration and interpretation is called?

A. Usability

B. Differentiability

C. Objectivity

D. Reliability

77. Facility index of an item determines?

A. Ease or difficulty

B. Discrimination power

C. Objectivity

D. Reliability

78. High and low achievers are sorted out by?

A. Ease or difficulty

B. Discrimination power

C. Objectivity

D. Reliability

79. Test item is acceptable which its faculty index /difficulty level ranges from?

A. 30-70 %

B. 70 %

C. 30%

D. None

80. Test item is very easy when value of faculty index/ difficulty level is higher than?

A. 30-70 %

B. 70 %

C. 30%

D. None

81. Test item is very difficult when value of facility index/ difficulty level is less than?

A. 30-70 %

B. 70 %

C. 30%

D. None

03047838705 (M.Javeed)

82. Discrimination power of an item is acceptable when its value ranges from?

A. 0.30 – 1

B. 1

C. 0.30

D. None

83. Test item discriminates 100% when its value for discrimination is?

A. 0.30 – 1

B. 1

C. 0.30

D. None

84. Test item cannot discriminate low achievers and high achievers when its value is lower than?

A. 0.30 – 1

B. 1

C. 0.30

D. None

03047838705 (M.Javeed)

EDU431 Quiz 3 File by Tanveer Online Academy

Which test requires the students to match a series of responses with corresponding items in stimulus list? Matching items

True-False Items are also known as: Alternatives form items
The item in the column for which a match is sought is called _____. Premise

The supply type test item is? Fill in the blanks

Which of the following types of items is not a selection item: Short answer

Identifying relationship between two things is done in_____. Matching items

Which of the following are supply type items? Fill in the blanks

Identify the types of objective type items. Selection and supply

A multiple choice question is composed of question referred as _____: Stem

A list of three or more choices from which an examinee is required to choose the correct one is given in: M.C.Q items

A multiple choice question is composed of question or statement referred as: Stem

Good distractor is that: Attracts equally low and high achievers

In which type of question marking will be more reliable_____. Multiple choice questions

The correct option in MCQ is called: Answer

The ability to select organize, integrate and evaluate ideas is demonstrated by_____. Extended Response Questions

Which of the following is an example of supply response items? Short answer

Following are the usefulness of completion questions EXCEPT: Measure complex learning outcomes.

How many columns matching items have _____? Two

Alternative response item is a type of _____. All

A test very popular with class room teacher is_____. Multiple choices test

The incorrect option in MCQ are called: Distractor

Which of the following is an example of supply response items? Short answer

Objective type question have advantage over essay type because such questions are _____. Easy to mark

"Types 40 words per minute without error" is an example of: Classical test theory

Which of the following is helpful for more consistent scoring across students and grades? Analytic Rubric

Guessing parameter can be measured in: IRT

Zero in raw score indicates:

A tool used to interpret and grade students' work against criteria and standards. Rubric

The list of suggested solutions in MCQs in the form of word, symbol, phrase etc. are called: Alternatives

The problem stated in the form of question or incomplete statement in MCQs is called: Stem

Identify the excellent item from given values of item discrimination. 0.50

Identify the excellent item from given values of item discrimination. 0.48

Psychometric properties of items are measured for: High Stake test

Test item is very difficult when value of difficulty level is less than? 0.3

Which of the following term used to describe the performance levels in scoring rubric? Criteria

"Types 40 words per minute without error" is an example of Classical test theory

Identify the value of item difficulty for most difficult test item. 0.1

Identify the value of item difficulty for easiest test item. 0.9 or above

Identify the acceptable item from given values of item discrimination. 0.26 or 0.29

Which of the following is not a basic step to design rubric? Give hints to a student directly

Descriptors are _____. Details description for each level of performance

Extended response items are helpful to determine: Length and complexity of response

MCQs cannot be used to assess the knowledge of: Illusion

A system of number or values assigned to work is known as: Score

Item analysis focuses to find out?

A scoring tool that teachers use to assess student learning. test

A set of criteria that provides an interpretation of the stated objectives (performance, behavior, quality).

Which of the following are major threats to the validity of test score interpretations___?

Which of the following is appropriate to measure student's abilities to synthesize when no definite answer can be pre-specified? Holistic Rubric

A description of specific tasks that the student can perform is called: Criterion referenced interpretation

A flatter item characteristics curve shows that item is _____. Poor Discriminator

A set of criteria and standard, educators can assess each student's performance on a wide variety of work. Rubric

A steeper item characteristics curve shows that item is _____. Good Discriminator

Under Item Response Theory, the difficulty of an item describes where the item functions along the _____. ability scale

A set of criteria that provides an interpretation of the stated objectives (performance, behavior, quality).

When you are ready to package and reproduce the test _____. Objectives and test blueprint

Multiple test forms should be assembled so that _____. Students don't cheat

Test item cannot discriminate low achievers and high achievers when its value is lower than? 0.3

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 Rubric

Identify the good item from given values of item discrimination. 0.3 - 0.39

A rubric is _____? used for assessing a particular type of work or performance

Identify the poor item from given values of item discrimination. 0.15,0.17

There are _____ main types of rubric. Two

Scoring of essay items is a _____. Time Consuming

Holistic rubric are good for evaluating _____. overall performance on a task

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

A test form is statistically equated to another test form to make the resulting test scores _____. Comparable

Difficulty level of an item determines

The detail of each level of performance in scoring rubric is known as: Descriptors

Following are the decisions made on the basis of high stake testing EXCEPT: Measure learning process during session

Advantages of holistic rubric involves _____. Quick scoring

Test that measure learning outcomes of students is called: Achievement test

Analytical rubric works best for: Restricted Response items

Which of the following should be school's professional security practices? Materials must be kept in a locked storage

Identifying relationship between two things is in: Matching Items

Feature, trait or dimension to be measured is decided in the _____ element of rubric. Criteria

To give an answer or make a judgment about something without being sure of all the facts is called: Guess

Which of the following option includes in reproducing the test? Knowing the photocopying machine

Which type of test typically designed for norm-referenced interpretations? Standardized

Discrimination power of an item is acceptable when its value ranges from? : 0.20 or higher

Which of the following are elements of scoring rubric? 1. Score 2. Criteria 3. Level of performance 4. Descriptors

The most commonly used guessing correction formula to predict and control is_____. S=R-W/N-1

Zero in raw score indicates?

When the test is ready, the next step is to _____ it. administer

Which type of test is mostly used in survey test batteries? Standardized achievement

Objective type question has advantage over essay type because such questions?
A. Are easy to prepare
B. Are easy to solve
C. Are easy to mark
D. None
Answer is = C

Discrimination value of more than 0.4 means
A. Item is good
B. Item is acceptable
C. Item is weak
D. None
Answer is = B

Test involving the construction of certain patterns are called?
A. Intelligence test
B. Performance tests
C. Scholastic test
D. None
Answer is = B

In multiple choice items the stem of the items should be?
A. Large
B. Small
C. Meaningful
D. None
Answer is = C

Which appropriate verb will you use to make an objective behavioral?
A. To know
B. To appreciate
C. To understand
D. To construct
Answer is = D

Objectives representing the purposes of instruction of a teacher are called?
A. Performance
B. Instructional
C. Attainment
D. None
Answer is = B

Running description of active behavior of a student as observed by the teacher is?
A. Anecdotal record
B. Autobiography
C. Interview

D. None
Answer is = A

A test very popular with class room teacher is?
A. True false test
B. Multiple choices
C. Matching
D. Completion test
Answer is = B

Frequently used tools of summative evaluation are?
A. Test
B. Teacher observation
C. Daily assignment
D. None
Answer is = A

The most commonly used guessing correction formula to predict and control is?
A. S=R-W
B. S=R-W/N-1
C. S=R-w/2-1
D. None
Answer is = B

The summative evaluation is?
A. Diagnostic
B. Certifying judgment
C. Continuous
D. None
Answer is = B

The difference between maximum and minimum values is?
A. Mean
B. Mode
C. Range
D. None
Answer is = C

The number of scores lying in a class interval is?
A. Mid-point
B. Quartiles
C. Class
D. Frequencies
Answer is = D

A multiple-choice question is composed of question referred as?

A. Stem
B. Distracter
C. Foil
D. Response
Answer is = A

In a norm referenced test which item is best?
A. Item difficulty is near zero
B. Item difficulty is near 100
C. Item difficulty is near 70
D. Item difficulty is near 50
Answer is = D

Which question has increasing objectivity of marking?
A. Unstructured essay
B. Structured essay
C. Short answer
D. Multiple type questions
Answer is = D

The most widely used format on standardized test in USA is?
A. Unstructured essay
B. Structured essay
C. Short answer
D. Multiple type questions
Answer is = D

Which questions are difficult to mark with reliability?
A. Unstructured essay
B. Structured essay
C. Short answer
D. Multiple type questions
Answer is = A

Projective techniques are used to measure?
A. Aptitude
B. Intelligence
C. Knowledge
D. Personality
Answer is = D

1. Assessment that monitors Learning Progress is ----------
   - ✓ Formative Assessment
2. ---------- is used when we want to assess student ‘s prior knowledge
   - ✓ Placement assessment
3. Measurement and Evaluation can be used to ----------
   - ✓ All of Above
4. The Summative evaluation is ----------
   - ✓ Certifying Judgment
5. Persistent Difficulties in Learning are Investigated in ----------
   - ✓ Diagnostic Assessment
6. Who was the founder of Pakistan is the question of --------- in Bloom taxonomy of cognitive domain?
   - ✓ Knowledge
7. “What is your name?” is the example of which level of Solo taxonomy ----------
   - ✓ Pre-structural
8. ---------- determines learning progress, provides feedback to reinforce learning, and correct learning errors
   - ✓ Formative Assessment
9. The purpose of formative learning is ----------
   - ✓ Monitoring progress of student
10. The purpose of measurement in education is to ----------
    - ✓ Collect information for Decision Making???
11. The purpose of the evaluation is to make ----------
    - ✓ Judgment
12. “What individual will do under natural condition “is the ----------.
    - ✓ Typical Performance Assessment
13. Multiple choice question, matching exercise, fill in the blanks and true false is the example of ------
    - ✓ Fixed Choice Assessment
14. Which of the following is the type of assessment by the use of classroom instruction? ---------
    - ✓ Placement Assessment
15. Curriculum Wing is also known as ----------
    - ✓ National Bureau of Curriculum
16. List, enlist, name, define keywords are used in -------- level of Depth of Knowledge
    - ✓ Recall
17. Frequently used tools of summative evaluation are ----------.
    - ✓ Test
18. Explain the process of water cycle is the question of -------- level in bloom taxonomy of cognitive domain
    - ✓ Comprehension
19. Student’s Performance is compared with other students in ----------.
    - ✓ Norm referenced test
20. The test made to compare the performance of student with the other students is called -----.
    - ✓ Norm reference
21. To Assess achievement at the end of instruction is ----------.
    - ✓ Summative Assessment
22. Permanent difficulties in learning are investigated in----------
    - ✓ Diagnostic Assessment

23. The purpose of evaluation is to make judgment about educational----------.
    - ✓ Quality
24. In norm referenced test the comparison is between ---------.
    - ✓ Individuals
25. How many domains of Bloom Taxonomy ……..?
    - ✓ 3
26. Procedures used to determine person best abilities are…….?
    - ✓ Maximum performance test
27. The type of essay item in which content are limited is----------?
    - ✓ Restricted response question
28. Identifying relationship between two things is done in --------?
    - ✓ Matching item
29. How many column matchings items have...?
    - ✓ Two
30. Cognitive domain of Blooms taxonomy are -------
    - ✓ 6
31. 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
32. SOLO stands for:
    - ✓ Structure of Observed Learning Outcomes
33. The process of obtaining a numerical description of the degree to which an individual processes particular characteristic is called:
    - ✓ Measurement
34. Measure of student's performance against a certain criteria is :
    - ✓ Criterion-referenced
35. The way of judging how well students are doing by looking at their work during educational process is called
    - ✓ Evaluation??
36. Hina Exceeded 85% of the sixth graders on the mathematics test. The interpretation of this statement is:
    - ✓ Norm Referenced
37. The Assessment determines what individual can do when performing at their best
    - ✓ Maximum Performance Assessment
38. In DOK taxonomy, create, synthesize, reflect are key words for learning level
    - ✓ Extended thinking
39. Which type of assessment determines the student's prerequisite skills to begin instruction
    - ✓ Placement
40. Identify the assessment used to find the root cause of continuous or constant problem in students learning
    - ✓ Diagnostic
41. Cognitive domain was divided into sub subgroups by bloom in
    - ✓ 1956
42. Degree to which a test measures an intended content area is called
    - ✓ construct validity
43. Which type of assessment measures entry level performance of student?
    - ✓ Placement

44. The spearman-brown formula is used to estimate:
    - ✓ Split-half
45. To judge the worth or value of material for a given purpose is known as
    - ✓ Evaluation
46. to use the previous learned material in new situation is called
    - ✓ Application
47. Which of the following decision determines prerequisite skills, degree of mastery of course goals and modes of learning?
    - ✓ Placement
48. Test designed to measure the number of items and individual can attempt correctly in a given time is referred as:
    - ✓ Speed
49. The degree to which an instrument measures what it is supposed to measured is called
    - ✓ Validity
50. Monitoring the outcomes with reference to the objectives is called:
    - ✓ Evaluation
51. Which type of Assessment is designed to measure whether the students have mastered the skills presented in instruction?
    - ✓ Summative Assessment
52. 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
53. How well the sample of assessment tasks represents the domain of the tasks to be measured is known as
    - ✓ Content validity
54. Reliability over different forms of assessment is known as
    - ✓ Equivalence
55. Which of the following provides a framework of categories with different hierarchical levels of outcomes?
    - ✓ Taxonomy
56. The procedure used to gain information about student learning and the formation of value judgement concerning learning progress is called:
    - ✓ Evaluation
57. Reliability within an assessment itself is known as:
    - ✓ Internal Consistency
58. ---------- refers to the assessment of participants, and summarize their development at the particular time
    - ✓ Summative assessment
59. In SOLO taxonomy; identify, memorize, are keywords for learning level:
    - ✓ Uni-Structural
60. Student creatively applies knowledge to construct an overall theory called:
    - ✓ Application
61. 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
62. “What is your country name?” is example of which level of SOLO taxonomy.
    - ✓ Pre-structural

63. The domain which deals with physical abilities and coordination objectives is
    - ✓ Psychomotor
64. The Degree to which a test measures an intended hypothetical construct is called:
    - ✓ Construct validity
65. What kind of assessment, permanent difficulties in learning are investigated ?
    - ✓ Diagnostic
66. The alternative name of the “table of specification” is ?
    - ✓ Test Blue Print
67. Which of the following are three domains of Bloom’s Taxonomy?
    - ✓ Cognitive, affective, psychomotor
68. Measure of performance which interprets an individual’s relative standing in some known group is:
    - ✓ Norm referenced
69. Which of following learning level of DOK taxonomy demands extended use of higher order thinking processes such as synthesis, reflection, assessment
    - ✓ Extended thinking
70. Identify the lowest learning level of SOLO taxonomy.
    - ✓ Pre-Structural
71. The purpose of the evaluation is to make judgment about educational ----------.
    - ✓ Quality
72. The quality of test to give same scores when administered at different occasions is called
    - ✓ Reliability
73. To assess achievement at the end of instruction is ----------
    - ✓ Summative assessment
74. Which of the following test with time limit strict that no one is supposed to complete all items?
    - ✓ Speed
75. In what kind of assessment, recurrent difficulties in learning are investigated?
    - ✓ Diagnostic Assessment
76. In SOLO taxonomy: classify, list, combine are keywords for learning level:
    - ✓ Multi- Structural
77. The Scores of students in papers are
    - ✓ Assessment??
78. Aptitude test is example of which type of Assessment
    - ✓ Placement Assessment
79. In DOK taxonomy; argue, critique, formulate are key words for learning level:
    - ✓ Strategic Thinking

Edu431 Grand quiz 2020

1. **Seven levels** in psychomotor domain the lowest level in psychomotor domain is **Origination**
2. The highest level of psychomotor domain is **Perception**
3. **Diagnostic Assessment** determines causes (intellectual physical, emotional environmental) of persistent learning difficulties
4. There are five levels in Affective Domain
5. The highest level of Affective Domain is **Receiving**
6. The lowest level of Affective Domain is **Characterization.**
7. Identify the highest learning level of DOK taxonomy: **Recall**
8. Identify the lowest learning level of DOK taxonomy: **Extended Thinking**
9. The cognitive level where components of a whole are broken into parts to know it‟s organizational structure is called: **Synthesis**
10. In revised Bloom taxonomy; assess, defend, appraise are key words for level:

11. . Which type of assessment is designed to measure whether the students have mastered the skills presented in instruction? **Fixed Choice Assessment**
12. Type of assessment includes attitude, Interest and personality inventories, observational techniques, peer appraisal :**typical performance test**
13. **Fixed Choice Assessment** is used for efficient measurement of knowledge and skills
14. **Placement test** is used to measure student,,s entry level performance to know either student have a sufficient knowledge required for a particular course or not
15. Readiness test: It is a test used to determine the students,, knowledge or concept about a particular course of instruction or what is the level of students
16. Aptitude test: It is used for the admission in a particular program
17. Pretest: It is made according to the course objectives and determines the student present knowledge about them
18. Self- report inventories: Determines the student level by interviewing or discussion

19. who was the founder of Pakistan is the question of **Knowledge** in Blooms taxonomy of cognitive domain.
20. **Comprehension** defined as the ability to grasp the meaning of the material
21. The assessment determines what individual can do when performing at their best is called: Maximum Performance Assessment
22. The classification of cognitive domain was presented by **Benjamin Bloom** in the 1956

Which of the following method is a measure of internal consistency? : Split Half

Which of the following method is a measure of stability **: Test-Retest**

25. Which of the following method is a measure of equivalence: **Equivalent Forms**
26. Which of the following method is a measure of stability and equivalence :**Test-Retest with Equivalent Forms**
27. Which of the following method is a measure of Internal Consistency :**Kuder-Richardson**
28. Which of the following method is a measure of Internal Consistency :**Cronbach Alpha**
29. Which of the following method is a measure of consistency of rating :**Inter-rater Reliability**
30. Reliability over different forms of assessment is known as: **Equivalence**
31. Consistency over period of time is known as: **Stability**
32. Reliability Within the assessment itself is known as**: Internal consistency**
33. . Achievement tests and aptitude tests are example of: **Maximum performance assessment**
34. **Assessment:** It means appraisal of something to improve quality of teaching and learning process for
35. deciding what more can be done to improve the teaching, learning and outcomes.
36. **Evaluation** is process of making a value judgment against intended learning outcomes and behavior, to decide quality and extent of learning.
37. In revised Bloom taxonomy; define, name, select are key words for level**:**

Remembering

38. Define, Find, How, Label, List, Match, Name, Omit, Recall, Relate, Select, Show, Spell,
39. Tell, What, When, Where, Which, Who, Why}
40. In revised Bloom taxonomy: Classify, Compare, Contrast, Demonstrate, Explain, Extend, Illustrate, Infer, Interpret, Outline, Relate, Rephrase, Show, Summarize,

49. In revised Bloom taxonomy: Analyze, Assume, Categorize, Classify, Compare, Conclusion, Contrast, Discover, Dissect, Distinguish, Divide, Examine, Function, Inference, Inspect are key words for level : **analyzing**
50. In revised Bloom taxonomy: Agree, Appraise, Assess, Award, Choose, Compare, Conclude, Criteria, Criticize, Decide, Deduct, Defend, Determine, Disprove, Estimate are key words for level: **Evaluating**
51. Students‟ performance is compared with clearly defined learning tasks in: **Criterion-referenced**
52. The alternative name of the “table of specification” is **blue print of test** .
53. In what kind of assessment a teacher uses information on learners‟ progress during a course to adapt their teaching or to give learners feedback on their learning?
54. **Measurement** is the process by which the attributes or dimensions of some object
55. (both physical and abstract) are quantified.
56. There **are six levels** in the cognitive domai
57. Bloom‟s Taxonomy cognitive domain highest level: **Knowledge**
58. Bloom‟s Taxonomy cognitive domain lowest level: **Evaluation**
59. **Diagnostic assessment** is the type of the assessment that determines the causes intellectual, emotional, and environmental of persistent learning difficulties.
60. One of the tools used by the teachers to develop a blueprint for the test is called :**table of specification TOC**
61. A student is asked to critique a painting based on explicit criteria. What is the cognitive level of this outcome according to Bloom‟s taxonomy? **Evaluating**
62. **Criterion validity** refers how much test can measure some performance with reference to
63. some defined standard.
64. **Predictive validity** refers that how much test can predict future performance

65. with reference to some defined standard.
66. Identify the lowest learning level of SOLO taxonomy: **Extended Abstract**
67. Identify the highest learning level of Solo taxonomy: **Pre-structural**
68. There are **three** main domains of learnings namely Cognitive Domain,Affective Domain and Psychomotor Domain
69. Student judges information based upon standards, criteria, values and opinions is known as:
70. **Competency :** It is a key learning area. For example algebra, arithmetic, geometry etc. in mathematics and vocabulary, grammar, composition etc. in English.
71. In national curriculum, key learning areas are defined as: **Competency**
72. **Standards**: These define the competency by specifying broadly, the knowledge, skills and attitudes that
73. students will acquire.
74. **Benchmarks:** indicating what the students will accomplish at the end of each of the five developmental levels in order to meet the standard. Elaborate the development level further,
75. **Student learning outcomes:** These are built on the descriptions of the benchmarks and describe what students will accomplish at the end of each grade. It is the lowest level of hierarchy.

→ Types of Written Tests Verbal, Non-verbal, Objective , Subjective, Teacher

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

77. **Verbal tests** are emphasizing on reading, writing, or speaking.
78. Most tests in education are **verbal tests.**
79. **Standardized** Test designed to measure broad national objectives and have a uniform set of instructions .
80. **Teacher made** test are designed according to need and issues related to specific class.
81. When it is difficult for two scorers to agree on whether an item is correct or incorrect, the test is a subjective one such as essay tests: **Subjective**

82. **Non-verbal:** Does not require reading, writing or speaking ability, tests composed of numerals or drawings is example.
83. **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, MCqs.
84. Tests with time limits so strict that no one is expected to complete all items. Items tend to be Easy: **Speed**
85. Tests with liberal time limits that allow each student to attempt each item. Item tend to be Difficult: **Power**
86. A test designed to measure **depression** must only measure that particular **construct.**
87. **Formative** assessment focuses on the process toward completing the product.
88. **Formative** assessment focus of measurement in formative assessment is predefined segment of instruction.
89. In **Formative** assessment , Limited sample of learning tasks are addressed.
90. **Formative assessment** is conducted periodically during the instructional process.
91. Results of formative assessment are used to improve and direct learning through **ongoing feedback**.
92. An **arithmetic test** have high degree of validity for computational skill and low degree for arithmetical reasoning
93. Observations during in-class activities: **Formative assessment**
94. Homework exercises as review for exams and class discussions: **Formative assessment**

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

    Reflections journals that are reviewed periodically during the semester: **Formative assessment**
96. Question and answer sessions, both formal (planned) and informal (spontaneous): **Formative assessment**
97. Conferences between the instructor and student at various points in the semester: **Formative assessment**
98. In-class activities where students informally present their results **Formative assessment**
99. Student feedback collected by periodically answering specific question about the instruction and their self-evaluation of performance and progress: **Formative assessment**

100. **Validity** is matter of degree

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

102. **summative** assessment focsues on course or unit objectives

103. **Summative** assessment is done at the end of the unit or the course.

104. Most important functions of **summative assessment** is to assign grade, certification of accomplishment and evaluation of teaching.

105. **Summative assessment** takes place after the learning has been completed.

106. **Summative** assessment is product-oriented and assesses the final product

107. **Types of Summative included** 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) and Performances

108. an assessment procedure must meet certain practical requirement which includes feasibility, administration environment and availability of results for decision makers: Usability

109. Reliability of measurement is needed to obtain the valid results, but we can have reliability without **validity**

110. **Reliability** refers to the consistency of assessment results.

111. Validity is an evaluation of adequacy and appropriateness of the interpretation and uses of results.

112. Characteristics of good test are Validity , Reliability , Usability.

113. Standards for selecting appropriate test First define the purpose for testing and the population to be tested and select the test accordingly.

114. Published tests commonly used by provincial or national testing programs are: Aptitude tests, Readiness tests, Placement tests(all)

115. Published test are designed and conducted in such a manner that each and every

116. characteristic is pre planned and known.

117. Time in class spent on objective (min) / total time for the instruction being examined (min) Percentage of instruction time is equal to **instruction time.**

118. Published test are also called supplement and complement informal classroom tests, and aid in many Instructional decisions.

119. Carey (1988) listed **six** major elements that should be attended to in developing a Table of

120. Specifications

121. The blueprint is meant to insure **content** validity.

122. Content validity is the most important factor in constructing an **achievement test.**

123. 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 .

124. In classrooms most used published tests are two Achievement tests and Reading test

125. In SOLO taxonomy; identify, memorize, are key words for learning level: **Uni Structural**

126. In SOLO taxonomy; Students are simply able to acquire bits of unconnected information and respond to a question in meaningless way such as What is your name? : Pre-structural

127. "What is your name?" is the example of which level of SOLO taxonomy: **Pre-structural**

128. In SOLO taxonomy; enumerate, classify, describe, list, combine, do algorithms are key words for learning level: **Multi- Structural**

129. Student can understand several components but the understanding of each remains discreet. **Multi- Structural**

130. Student can indicate connection between facts and theory, action and purpose.: **Rational**

131. In SOLO taxonomy ,compare/contrast, explain causes, integrate, analyze, relate, and apply are key words for learning level: **Rational**

132. Student at this level is able to think hypothetically and can synthesize a material logically : **Extended Abstract**

133. Student make connections not only with in the given subject area but understanding is

134. transferable and generalizable to different areas: **Extended Abstract**

135. In SOLO taxonomy ,theorize, generalize, hypothesize, reflect, generate are key words for learning level: **Extended Abstract**

136. Highest Levels of SOLO**: Pre-structural**

137. Lowest Level of SOLO: **Pre-structural**

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

139. A collection of the constituents or parts of a concept so as to make a whole: **Synthesis**

140. Design a classification Scheme for writing **educational objectives** that combines the cognitive, affective, and Psychomotor domains: **Synthesis**

141. The breakdown of a concept into its constituents parts such that the relative hierarchy of the concept is made easy to understand: **analysis**

142. Compare and contrast the cognitive and affective domains: **analysis**

143. How far the different BISEs and universities are developing papers using Bloom's taxonomy? Support your answer with arguments: **Evaluation**

144. In SOLO taxonomy; classify, list, combine are key words for learning level: **Multi- Structural**

Edu 431 mid MCQ

1. ----------is used when we want to assess student's prior knowledge

A) Placement assessment

b) Formative Assessment

c) Summative assessment

d) Diagnostic Assessment

2. Cognitive domain of Bloom Taxonomy are

a) 3

9. 4

10.6

11.8

16. Frequently used tools of summative evaluation are---------

? a) Teacher observation

b) Daily assignment

20.Test

21.None

28. What individual will do under natural conditions is the

Maximum Performance Assessment
Norm-referenced Assessment
c)Criterion-referenced Assessment
d) Typical Performance Assessment

5. To assess achievement at the end of the instruction is……..

33.Summative assessment

34.Placement assessment

35.Formative Assessment

36.Diagnostic Assessment

35.Student's performance is compared with other student's in_____

a) Standardized Test

b) Essay type Test

c) Objective type test

d) Norm referenced test

38. What is your name? is the example of which level of SOLO taxonomy

Pre-structural
Uni Structural
Multi- Structural
Relational

46. The purpose of the evaluation is to make ------
a) Decision
b) Prediction
c) Judgment
d) Opinion

47. Explain the process of water cycle is the question of -------- level in Blooms taxonomy of cognitive domain.
a) Knowledge
b) Comprehension
c) Analysis
d) Application

50. Permanent difficulties in learning are investigated in----------
Summative assessment

Placement assessment

Formative Assessment

Diagnostic Assessment

52. ----------determines learning progress, provides feedback to reinforce learning, and correct learning errors

A) Placement assessment

Formative Assessment

Summative assessment d)

Diagnostic Assessment

67. The purpose of formative Evaluation is----------
Decision of what to measure
Development of the test
Administering the test
Monitoring progress of the students

68. Curriculum Wing also known as
National Bureau of Curriculum
75. Curriculum authority development
76. Development of curriculum development
77. None of the above

76. Assessment that monitors learning progress ----------
Placement assessment

Formative Assessment

Diagnostic Assessment

Summative assessment

 The purpose of measurement in education is to ----------

a) Make decision about students

Determine if objectives are being met

Assess teacher skills

d) Collect information for decision making

16. The summative evaluation is……

a) Diagnostic

b) Certifying judgment

c) Continuous

d) None

78. The purpose of evaluation is to make judgment about educational----------

a) Quantity

82. Quality

83. Time period

84. Age

95. Multiple choice question, matching exercise, fill in the blanks and true false is the example of

Typical performance assessment

Fixed choice assessment

108. Maximum performance assessment

109. Formative assessment

110. The test made to compare of student with other students is called……..

Criterion reference

Norm reference

112. Achievement

113. None

116. List, enlist, name, define keywords are used in …… level of Depth of knowledge

a) Recall

b) skill/concept

119. Strategic thinking

120. Extended thinking

123. In norm referenced test the comparison is between ---------

Groups

126. Individuals

127. Areas

128. Interest

127. Who was the founder of Pakistan is the question of----------in Blooms taxonomy of cognitive

Knowledge

Comprehension

128. Application

129. Synthesis

131. Measurement and evaluation can be used to ……..

Motivate students

Help develop the scope and sequence of teaching

Assess the effectiveness of learning activities

d) all of the above

137. How many domains of Bloom Taxonomy are……..

a)1

b) 2

c) 3

d) 4

25. Procedures used to determine person best abilities are…….

139. Maximum performance test

140. Typical performance test

141. Norm reference test

142. Criterion reference test

140. Which of following is the type of assessment by the use of classroom instruction……..

142. Complex performance assessment

143. Placement assessment

144. Fixed choice assessment

145. Norm reference assessment

27. In …….. Federal government appointed university grant commission

145. 1960

146. 1964

147. 1976

148. 1980

28. The type of essay item in which content are limited is:

a) Restricted response question

b) Extended response question

c) Matching item

d) M.C.Q items

29. Identifying relationship between two things is done

in a) True/false items

b) Matching item

c) M.C.Q items

d) Completion item

30. How many column matchings items have...?

a) One

b) Two

c) three

d) Four

31. The incorrect option in MCQ are called

a) Answer

b) Premise

c) Response

d) Distractor

32. The item in the column for which a match is sought is called

a) Premise

b) Response

c) Distractor

d) Answer key

33. The correct option in MCQ is called

a) Answer

b) Premise

c) Response

d) Distractor

34. Alternative response item is a type of ……..

a) Matching List

b) Completion List

c) Correct/Incorrect

d) All

35. In which type of question marking will be more

reliable? a) Completion

b) Short answer

c) Multiple choice question

d) Essay

36. In multiple choice items the stem of the items should be

a) Large

b) Small

c) Meaningful

d) None of the above

37. A multiple choice question is composed of question referred as

a) Stem

b) Distractor

c) Foil

d) Response

1. Instrument used for measuring sample of behavior is?

A. Test

B. Measurement

C. Assessment

D. Evaluation

2. Limited to quantitative description of pupil's performance is?

A. Test

B. Measurement

C. Assessment D.
Evaluation

3. The purpose of the evaluation is to make?

A. Decision

B. Prediction

C. Judgment

D. Opinion

4. The purpose of evaluation is to make judgment about educational?

A. Quantity

B. Quality

C. Time period

D. Age

5. Evaluation that monitors learning progress is?

A. Placement evaluation

B. Formative evaluation

C. Diagnostic evaluation

D. Summative evaluation

6. A formal and systematic procedure of getting information is?

A. Test

B. Measurement

C. Assessment

D. Evaluation

7. The process of obtaining numerical value is?

A. Test

B. Measurement

C. Assessment

D. Evaluation

Answer is = B

8. A sum of questions is?

A. Test

B. Measurement

C. Assessment

D. Evaluation

9. The first step in measurement is?

A. Decision of what to measure

B. Development of the test

C. Administering the test

D. None

10. The purpose of formative evaluation is?

A. Decision of what to measure

B. Development of the test

C. Administering the test

D. Monitoring progress of students

11. To assess achievement at the end of instruction is? A. Placement Assessment

B. Formative Assessment

C. Summative Assessment

D. Diagnostic Assessment

12. Vast of all in scope?

A. Test

B. Measurement

C. Assessment

D. Evaluation

13. The least in scope is?

A. Test

B. Measurement

14. Permanent difficulties in learning are investigated in? A. Placement Assessment

B. Formative Assessment

C. Summative Assessment

D. Diagnostic Assessment

15. Broader in meaning is?

A. Aims

B. Objectives

C. Instructional objectives

D. Specific Objectives

16. Procedures used to determine person abilities is?

A. Maximum performance test

B. Typical performance test C. Norm performance test

D. Criterion performance test

17. In norm referenced test the comparison is between? A. Groups

B. Individuals

C. Areas

D. Interest

18. In which question marking will be more reliable? A. Completion

B. Short answer

C. Multiple choice question D. Essay

19. Facility value of less than 0.20 means?

A. Item is too easy

B. Item is difficult

C. Item is acceptable

D. Item is easy

20. Objective type question have advantage over essay type because such questions?

A. Are easy to prepare

B. Are easy to solve

C. Are easy to mark

D. None

21. Discrimination value of more than 0.4 means A. Item is good

B. Item is acceptable

C. Item is weak

D. None

22. Test involving the construction of certain patterns are called?
A. Intelligence test

B. Performance tests

C. Scholastic test

D. None

23. In multiple choice items the stem of the items should be?

A. Large

B. Small

C. Meaningful

D. None

24. Which appropriate verb will you use to make an objective behavioral?
A. To know

B. To appreciate

C. To understand

D. To construct

25. Objectives representing the purposes of instruction of a teacher are called? A. Performance

B. Instructional

C. Attainment

D. None

26. Running description of active behavior of a student as observed by the teacher is?

A. Anecdotal record

B. Autobiography

C. Interview

D. None

27. A test very popular with class room teacher is? A. True false test

B. Multiple choices

C. Matching

D. Completion test

28. Frequently used tools of summative evaluation are?

A. Test

B. Teacher observation

C. Daily assignment D. None

29. The most commonly used guessing correction formula to predict and control is?

A. S=R-W

B. S=R-W/N-1

C. S=R-w/2-1

D. None

30. The summative evaluation is?
A. Diagnostic

B. Certifying judgment

C. Continuous

D. None

31. The difference between maximum and minimum values is?
A. Mean

B. Mode

C. Range

D. None

32. The number of score lying in a class interval is?

A. Mid-point

B. Quartiles

C. Class

D. Frequencies

33. A multiple choice question is composed of question referred as?

A. Stem

B. Distracter
C. Foil

D. Response

34. In a norm referenced test which item is best?

A. Item difficulty is near zero

B. Item difficulty is near 100

C. Item difficulty is near 70

D. Item difficulty is near 50

35. Which question has increasing objectivity of
marking? A. Unstructured essay

B. Structured essay
C. Short answer

D. Multiple type questions

36. The most widely used format on standardized test in USA is? A. Unstructured essay

B. Structured essay

C. Short answer

D. Multiple type questions

37. Which questions are difficult to mark with reliability?

A. Unstructured essay

B. Structured essay

C. Short answer

D. Multiple type questions

38. Projective techniques are used to measure? A. Aptitude

B. Intelligence

C. Knowledge

D. Personality

39. Test meant for prediction on a certain criterion are called?

A. Aptitude test

B. Intelligence C.

Knowledge D.

Personality

40. Kuder Richardson method is used to estimate?

A. Reliability

B. Validity

C. Objectivity

D. Usability

41. Value that divides the data into two equal parts is?

A. Mean

B. Median

C. Mode

D. None

42. The test measures what we intend to measure. This quality of the test is called?

A. Reliability

B. Validity C.

Objectivity

D. Usability

43 The length of a test is an important factor in obtaining a representative?

A. Mean

B. Median

C. Mode

D. Sample

44. Median of 1,2,4,5,2,3, is

? A. 2

B. 5

C. 3.5

D. None

45. The test made to compare the performance of student with the other students is called?

A. Criterion reference

B. Norm reference

C. Achievement D.

None

46. The summative evaluation is used?

A. At the end of the program

B. At the middle of the program

C. At the start of the program

D. None

47.The appearance of normal curve resembles with?

A. U

B. Bell

C. V

D. None

48. The alternative name of the “table of spectification” is?

A. Test Blue Print B.

Test Construction C.

Test Administration

D. Test Scoring

49. ”table of specification” helps in?

A. Test development

B. Test Construction

C. Test Administration

D. Test Scoring

50. The supply type test item is? A. True / False items

B. Matching items

C. M.C.Q items

D. Completion items

51. Alternative response item is?

A. True / False items

B. Right / wrong

C. Correct / incorrect

D. All above

52. How many columns matching items have?

A. One

B. Two

C. Four

D. Five

53 The item in the column for which a match is sought is?

A. Premise B.

Response C.

Destructor D.

None

54. Identifying relationship between two things is demonstrated by?
A. True / False items

B. Matching items

C. M.C.Q items

D. Completion items

55. The statement of problem in M.C .Qs is?

A. Premise

B. Response

C. Stem

D. None

56. The correct option in M.C.Q is?

A. Answer

B. Premise

57. The incorrect options in M.C.Q are?
A. Answer

B. Premise C.

Response D.

Destructor

58. The most widely applicable test item is?
A. True / False items

B. Matching items

C. M.C.Q items

D. Completion items

59. The type of essay item in which contents are limited is?

A. Restricted Response

Questions B. Extended Response

Questions C. Matching items

D. M.C.Q items

60. The ability to select organize, integrate and evaluate ideas is demonstrated by?

A. Restricted Response Questions

B. Extended Response Questions C. Matching items

D. M.C.Q items

61. The Analysis of items is necessary in?

A. Standardized Test

B. Essay Type Test

C. Objective type test

D. Norm referenced test

62. Which one is not the type of test of test by purpose? A. Standardized Test

B. Essay Type Test

C. Objective type test

D. Norm referenced test

63. The type of the test by method is?

A. Standardized Test

B. Essay Type Test

C. Objective type test D. Norm referenced test

64. Student's performance is compared with other students in?

A. Standardized Test

B. Essay Type Test C. Objective type test

D. Norm referenced test

65. Student performance is compared with clearly defined learning tasks in?

A. Standardized Test

B. Essay Type Test

C. Criterion reverenced test

D. Norm referenced test

66. Test that measure learning outcome of students is

A. Achievement test

B. Aptitude test

C. Criterion reverenced test

D. Norm referenced test

66.The tests designed to predict future performance is?

A. Achievement test

B. Aptitude test

C. Criterion reverenced test

D. Norm referenced test

67. The founder of modern intelligent tests was?

A. Alfred Binet

B. Stern

C. Gulford

D. None

68. The formula to determine I.Q was presented by?

A. Alfred Binet

B. Stern

C. Gulford

D. None

69. I.Q of a student having same physical and mental age will be?

A. 100

B. 120

C. 50

D. 140

70 The I.Q of a student having twelve years mental age and tem years physical age will be?

A. 100

B. 120

C. 50

D. 140

71. The quality of test that measures “what it claims to measure” is?

A. Validity

B. Differentiability
C. Objectivity
D. Reliability

72. The characteristic of a test to discriminate between high achievers and low achievers is?

A. Validity

B. Differentiability

C. Objectivity
D. Reliability

73. If the scoring of the test is not effected by any factor, quality of test is called? A. Validity

B. Differentiability

C. Objectivity

D. Reliability

74. The quality of test to give same scores when administered at different occasions is?

A. Validity

B. Differentiability
C. Objectivity

D. Reliability

75. If the sample of the question in the test is sufficiently large enough, the quality of test is?

A. Adequacy

B. Differentiability
C. Objectivity
D. Reliability

76. The quality of test showing ease of time, cost, administration and interpretation is called?

A. Usability

B. Differentiability

C. Objectivity

D. Reliability

77. Facility index of an item determines?

A. Ease or difficulty

B. Discrimination power

C. Objectivity

D. Reliability

78. High and low achievers are sorted out by?

A. Ease or difficulty

B. Discrimination power

C. Objectivity

D. Reliability

79. Test item is acceptable which its faculty index /difficulty level ranges from?

A. 30-70

% B.70%

C. 30% D.

None

80. Test item is very easy when value of faculty index/ difficulty level is higher than?

A. 30-70 %

B.70%

C. 30%

D. None

81. Test item is very difficult when value of facility index/ difficulty level is less than?

A. 30-70 %

B.70%

C. 30%

D. None

82. Discrimination power of an item is acceptable when its value ranges from?

A. 0.30 – 1

B. 1

C. 0.30

D. None

83. Test item discriminates 100% when its value for discrimination is?

A. 0.30 – 1

B. 1

C. 0.30

D. None

84. Test item cannot discriminate low achievers and high achievers when its value is lower than?

A. 0.30 – 1

B. 1

C. 0.30

2. None

Achievement test , student learn by themselves or we teach them and at the end we want to see that how much student learn

3. **Achievement test** designed to indicate the degree of success in some past learning activity.

d) **Fixed Choice Assessment** is sued to measure the skills of people efficiently. (multiple choice questions, MCQs)

e) List, enlist, name, define keywords are used in **Recall** level of Depth of Knowledge

f) Graph, separate, relate, contrast, narrate, compare keywords are used in **skill or concept level** of Depth of Knowledge

g) Argue, critique, formulate keywords are used in **Strategic thinking level of** Depth of Knowledge

h) Create, Synthesize, Design and reflection keywords are used in **Extended Thinking level** of Depth of Knowledge

i) **Placement assessment** is used when we want to assess student,,s prior knowledge so that we can decide what the level of student is