

- Use only one central idea in each item.
- Don't emphasize the trivial.
- Use exact quantitative language
- Don't lift items straight from the book.
- Make more false than true (60/40). (Students are more likely to answer true.)
- The desired method of marking true or false should be clearly explained before students begin the test.
- Construct statements that are definitely true or definitely false, without additional qualifications. If opinion is used, attribute it to some source.

**Avoid the following:**

- a. verbal clauses, absolutes, and complex sentences;
- b. broad general statements that are usually not true or false without further qualifications;
- c. terms denoting indefinite degree (e.g., large, long time, or regularly) or absolutes (e.g., never, only, or always).
- d. placing items in a systematic order (e.g., TTFF, TFTF, and so on);
- e. taking statements directly from the text and presenting them out of context.

**Activity 4.2:** Enlist five items by indicating them T/F (True & False)

### 4.1.3 Matching items

According to Cunningham (1998), the matching items consist of two parallel columns. The column on the left contains the questions to be answered, termed premises; the column on the right, the answers, termed responses. The student is asked to associate each premise with a response to form a matching pair.

**For example;**

<b>Column “A” Capital City</b>	<b>Column “B” Country</b>
Islamabad	Iran
Tehran	Spain
Istanbul	Portugal
Madrid	Pakistan
Jaddah	Netherlands
	Turkey
	West Germany

Matching test items are used to test a student's ability to recognize relationships and to make associations between terms, parts, words, phrases, clauses, or symbols in one column with related alternatives in another column. When using this form of test item, it is a good practice to provide alternatives in the response column that are used more than once, or not at all, to preclude guessing by elimination. Matching test items may have either an equal or unequal number of selections in each column.

Matching-Equal Columns. When using this form, providing for some items in the response column to be used more than once, or not at all, can preclude guessing by elimination.

**Good for:**

- Knowledge level
- Some comprehension level, if appropriately constructed

**Types:**

- Terms with definitions
- Phrases with other phrases
- Causes with effects
- Parts with larger units
- Problems with solutions

**Advantages:**

The chief advantage of matching exercises is that a good deal of factual information can be tested in minimal time, making the tests compact and efficient. They are especially well suited to who, what, when and where types of subject matter. Further students frequently find the tests fun to take because they have puzzle qualities to them.

- Maximum coverage at knowledge level in a minimum amount of space/prep time
- Valuable in content areas that have a lot of facts

**Disadvantages:**

The principal difficulty with matching exercises is that teachers often find that the subject matter is insufficient in quantity or not well suited for matching terms. An exercise should be confined to homogeneous items containing one type of subject matter (for instance, authors-novels; inventions inventors; major events-dates terms – definitions; rules examples and the like). Where unlike clusters of questions are used to adopt but poorly informed student can often recognize the ill-fitting items by their irrelevant and extraneous nature (for instance, in a list of authors the inclusion of the names of capital cities).

Student identifies connected items from two lists. It is useful for assessing the ability to discriminate, categorize, and association amongst similar concepts.

- Time consuming for students
- Not good for higher levels of learning

**Tips for Writing Good Matching items:**

Here are some suggestions for writing matching items:

- Keep both the list of descriptions and the list of options fairly short and homogeneous – they should both fit on the same page. Title the lists to ensure homogeneity and arrange the descriptions and options in some logical order. If this is impossible you're probably including too wide a variety in the exercise. Try constructing two or more exercises.
- Make sure that all the options are plausible distracters for each description to ensure homogeneity of lists.
- The list of descriptions on the left side should contain the longer phrases or statements, whereas the options on the right side should consist of short phrases, words or symbols.
- Each description in the list should be numbered (each is an item), and the list of options should be identified by letter.
- Include more options than descriptions. If the option list is longer than the description list, it is harder for students to eliminate options. If the option list is shorter, some options must be used more than once. Always include some

options that do not match any of the descriptions, or some that match more than one, or both.

- In the directions, specify the basis for matching and whether options can be used more than once.
- Need 15 items or less.
- Give good directions on basis for matching.
- Use items in response column more than once (reduces the effects of guessing).
- Make all responses plausible.
- Put all items on a single page.
- Put response in some logical order (chronological, alphabetical, etc.).

**Activity 4.3:** Keeping in view the nature of matching items, construct at least five items of matching case about any topic.

#### 4.1.4 Completion Items

Like true-false items, completion items are relatively easy to write. Perhaps the first tests classroom teachers' construct and students take completion tests. Like items of all other formats, though, there are good and poor completion items. Student fills in one or more blanks in a statement. These are also known as "Gap-Fillers." Most effective for assessing knowledge and comprehension learning outcomes but can be written for higher level outcomes. e.g.

The capital city of Pakistan is -----.

#### Suggestions for Writing Completion or Supply Items

Here are our suggestions for writing completion or supply items:

- I. If at all possible, items should require a single-word answer or a brief and definite statement. Avoid statements that are so indefinite that they may be logically answered by several terms.
  - a. **Poor item:**  
World War II ended in \_\_\_\_\_.
  - b. **Better item:**  
World War II ended in the year \_\_\_\_\_.
- II. Be sure the question or statement poses a problem to the examinee. A direct question is often more desirable than an incomplete statement because it provides more structure.

- III. Be sure the answer that the student is required to produce is factually correct. Be sure the language used in the question is precise and accurate in relation to the subject matter area being tested.
- IV. Omit only key words; don't eliminate so many elements that the sense of the content is impaired.
- a. **Poor item:**  
The \_\_\_\_\_ type of test item is usually more \_\_\_\_\_ than the \_\_\_\_\_ type.
- b. **Better item:**  
The supply type of test item is usually graded less objectively than the \_\_\_\_\_ type.
- I. Word the statement such that the blank is near the end of the sentence rather than near the beginning. This will prevent awkward sentences.
- II. If the problem requires a numerical answer, indicate the units in which it is to be expressed.

<p><b>Activity 4.3:</b> Construct five fill in the blanks about Pakistan.</p>
---

## 4.2 Supply Type Items

The aviation instructor is able to determine the students' level of generalized knowledge of a subject through the use of supply-type questions. There are four types of test items in supply type category of test. Commonly these are completion items, short answers, restricted response and extended response (essay type comprises the restricted and extended responses).

### 4.2.1 Short Answer

Student supplies a response to a question that might consist of a single word or phrase. Most effective for assessing knowledge and comprehension learning outcomes but can be written for higher level outcomes. Short answer items are of two types.

- Simple direct questions  
Who was the first president of the Pakistan?
- Completion items

The name of the first president of Pakistan is \_\_\_\_\_.

The items can be answered by a word, phrase, number or symbol. Short-answer tests are a cross between essay and objective tests. The student must supply the answer as with an essay question but in a highly abbreviated form as with an objective question.

**Good for:**

- Application, synthesis, analysis, and evaluation levels

**Advantages:**

- Easy to construct
- Good for "who," "what," "where," "when" content
- Minimizes guessing
- Encourages more intensive study-student must know the answer vs. recognizing the answer.

Gronlund (1995) writes that short-answer items have a number of advantages.

- They reduce the likelihood that a student will guess the correct answer
- They are relatively easy for a teacher to construct.
- They are well adapted to mathematics, the sciences, and foreign languages where specific types of knowledge are tested (The formula for ordinary table salt is \_\_\_\_\_).
- They are consistent with the Socratic question and answer format frequently employed in the elementary grades in teaching basic skills.

**Disadvantages:**

- May overemphasize memorization of facts
- Take care - questions may have more than one correct answer
- Scoring is laborious

According to Gronlund (1995) there are also a number of disadvantages with short-answer items.

- They are limited to content areas in which a student's knowledge can be adequately portrayed by one or two words.
- They are more difficult to score than other types of objective-item tests since students invariably come up with unanticipated answers that are totally or partially correct.
- Short answer items usually provide little opportunity for students to synthesize, evaluate and apply information.

**Tips for Writing Good Short Answer Items:**

- When using with definitions: supply term, not the definition-for a better judge of student knowledge.
- For numbers, indicate the degree of precision/units expected.

- Use direct questions, not an incomplete statement.
- If you do use incomplete statements, don't use more than 2 blanks within an item.
- Arrange blanks to make scoring easy.
- Try to phrase question so there is only one answer possible.

<p><b>Activity 4.5:</b> Develop a test of short answers on democracy in Pakistan.</p>
---

### 4.2.3 Essay

Essay questions are supply or constructed response type questions and 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. The student has to formulate and write a response, which may be detailed and lengthy. The accuracy and quality of the response are judged by the teacher.

Essay questions provide a complex prompt that requires written responses, which can vary in length from a couple of paragraphs to many pages. Like short answer questions, they provide students with an opportunity to explain their understanding and demonstrate creativity, but make it hard for students to arrive at an acceptable answer by bluffing. They can be constructed reasonably quickly and easily but marking these questions can be time-consuming and grade agreement can be difficult.

Essay questions differ from short answer questions in that the essay questions are less structured. This openness allows students to demonstrate that they can integrate the course material in creative ways. As a result, essays are a favoured approach to test higher levels of cognition including analysis, synthesis and evaluation. However, the requirement that the students provide most of the structure increases the amount of work required to respond effectively. Students often take longer time to compose a five paragraph essay than they would take to compose paragraph answer to short answer questions.

Essay items can vary from very lengthy, open ended end of semester term papers or take home tests that have flexible page limits (e.g. 10-12 pages, no more than 30 pages etc.) to essays with responses limited or restricted to one page or less. Essay questions are used both as formative assessments (in classrooms) and summative assessments (on standardized tests). There are 2 major categories of essay questions -- short response (also referred to as restricted or brief) and extended response.

- Restricted Response: more consistent scoring, outlines parameters of responses
- Extended Response Essay Items: synthesis and evaluation levels; a lot of freedom in answers

### **A. Restricted Response Essay Items**

An essay item that poses a specific problem for which a student must recall proper information, organize it in a suitable manner, derive a defensible conclusion, and express it within the limits of posed problem, or within a page or time limit, is called a restricted response essay type item. The statement of the problem specifies response limitations that guide the student in responding and provide evaluation criteria for scoring.

#### **Example 1:**

List the major similarities and differences in the lives of people living in Islamabad and Faisalabad.

#### **Example 2:**

Compare advantages and disadvantages of lecture teaching method and demonstration teaching method.

### **When Should Restricted Response Essay Items be used?**

Restricted Response Essay Items are usually used to:-

- Analyze relationship
- Compare and contrast positions
- State necessary assumptions
- Identify appropriate conclusions
- Explain cause-effect relationship
- Organize data to support a viewpoint
- Evaluate the quality and worth of an item or action
- Integrate data from several sources

### **B. Extended Response Essay Type Items**

An essay type item that allows the student to determine the length and complexity of response is called an extended-response essay item. This type of essay is most useful at the synthesis or evaluation levels of cognitive domain. We are interested in determining whether students can organize, integrate, express, and evaluate information, ideas, or pieces of knowledge the extended response items are used.

#### **Example:**

Identify as many different ways to generate electricity in Pakistan as you can? Give advantages and disadvantages of each. Your response will be graded on its accuracy, comprehension and practical ability. Your response should be 8-10 pages in length and it will be evaluated according to the RUBRIC (scoring criteria) already provided.

Over all Essay type items (both types restricted response and extended response) are



**Good for:**

- Application, synthesis and evaluation levels

**Types:**

- Extended response: synthesis and evaluation levels; a lot of freedom in answers
- Restricted response: more consistent scoring, outlines parameters of responses

**Advantages:**

- Students less likely to guess
- Easy to construct
- Stimulates more study
- Allows students to demonstrate ability to organize knowledge, express opinions, show originality.

**Disadvantages:**

- Can limit amount of material tested, therefore has decreased validity.
- Subjective, potentially unreliable scoring.
- Time consuming to score.

**Tips for Writing Good Essay Items:**

- Provide reasonable time limits for thinking and writing.
- Avoid letting them to answer a choice of questions (You won't get a good idea of the broadness of student achievement when they only answer a set of questions.)
- Give definitive task to student-compare, analyze, evaluate, etc.
- Use checklist point system to score with a model answer: write outline, determine how many points to assign to each part
- Score one question at a time-all at the same time.

<b>Activity 4.6:</b> Develop an essay type test on this unit while covering the levels of knowledge, application and analysis.
--

### **4.3 Self Assessment Questions:**

1. In an area in which you are teaching or plan to teach, identify several learning outcomes that can be best measured with objective and subjective types questions.
2. Criticize the different types of selection and supply categories. In your opinion which type is more appropriate for measuring the achievement level of elementary students?
3. What factors should be considered in deciding whether subjective or objective type questions should be included in a classroom tests?
4. Compare the functions of selection and supply types items.

#### 4.4 References/Suggested Readings

- Airasian, P. (1994) "Classroom Assessment," Second Edition, NY" McGraw-Hill.
- American Psychological Association. (1985). Standards for Educational and Psychological Testing. Washington, DC: American Psychological Association.
- Anastasi, A. (1988). Psychological Testing (6<sup>th</sup> ed.). New York, NY: MacMillan Publishing Company.
- Cangelosi, J. (1990) "Designing Tests for Evaluating Student Achievement." NY: Addison-Wesley.
- Cunningham, G.K. (1998). Assessment in the Classroom. Bristol, PA: Falmer Press.
- Ward, A.W., & Murray-Ward, M. (1999). Assessment in the Classroom. Belmont, CA: Wadsworth Publishing Co.
- Gronlund, N. (1993) "How to Make Achievement Tests and Assessments," 5th Edition, NY: Allyn and Bacon.
- Gronlund, N. E. & Linn, R. L. (1995). Measurement and Assessment in Teaching. New Delhi: Baba Barkha Nath Printers.
- Haladyna, T.M. & Downing, S.M. (1989) Validity of a Taxonomy of Multiple-Choice Item-Writing Rules. "Applied Measurement in Education," 2(1), 51-78.
- Monahan, T. (1998) The Rise of Standardized Educational Testing in the U.S. – A Bibliographic Overview.
- Ravitch, Diane, "The Uses and Misuses of Tests", in The Schools We Deserve (New York: Basic Books, 1985), pp. 172–181.
- Thissen, D., & Wainer, H. (2001). Test Scoring. Mahwah, NJ: Erlbaum.
- Wilson, N. (1997) Educational standards and the problem of error. Education Policy Analysis Archives, Vol 6 No 10

## **UNIT-5**

# **RELIABILITY OF THE ASSESSMENT TOOLS**

**Written by:**

**Dr. Muhammad Tanveer Afzal**

**Reviewed by:**

**Prof. Dr. Rehana Masrur**

## CONTENT

<i>Sr. No</i>	<i>Topic</i>	<i>Page No</i>
	Introduction.....	
	Objectives .....	
5.1	Reliability .....	
5.2	Types of Reliability .....	
5.3	Factor Affecting Rehabilitee.....	
5.4	Usability of Assessment Tools .....	
5.5	Summary .....	
5.6	Self Assessment Questions .....	
5.7	References/Suggested Readings .....	