UNIT: 4

TYPES OF TESTS

Written By:
Dr. Naveed Sultana

Reviewed By:
Dr. Muhammad Tanveer Afzal

CONTENT

INTRODUCTION

Classroom tests play a central role in the assessment of student learning. Tests provide relevant measures of many important learning outcomes and indirect evidence concerning others. They make expected learning outcomes explicit to students and parents and show what types of performance are valued. The validity of the information they provide, however, depends on the care that goes into the planning and preparation of tests. The main goal of classroom testing is to obtain valid, reliable and useful information concerning assessment. This requires determining what is to be measured and then defining it precisely so that tasks that evoke the desired performance can be constructed. In a standard based approach to education and training, informed by Constructivist theory, assessment informed instruction is the expectation as is continuous improvement. One of the most widely used tools in assessment and evaluation is the traditional or classic classroom achievement test, whether the classroom is on- or offline. These measures are often fraught with reliability and validity problems as the process for constructing such tests is often not followed or misunderstood, thereby introducing significant measurement error into the measurement process. Poor measurement frequently leads to inaccurate data-based inferences, which in turn leads to bad decision-making. Moreover classroom tests and assessment can be used for a variety of instructional purposes such examining the quality of teaching learning process, students achievement individually and success of institution overall. So in this unit we will examine the test item type and item format, writing select response items (multiple-choice, true/false, matching, completion and short-answer) and supply response items (brief and extended response). Each type of test item has its own unique characteristics, uses, advantages, limitations and rules for construction, which will be elaborated in this unit.

OBJECTIVES

After reading this unit, you will be able to:

- define the nature of selection and supply type time items.

- examine the role, advantages and disadvantages of different types of objective and subjective type tests for measuring the students' achievement.

- describe the learning outcomes that are best measured with selection and supply test items.

- differentiate the characteristics of all types of selection and supply categories of items concentrating to measure the higher level of thinking of students.

4.1   Selection Type Items (objective type)

There are four types of test items in selection category of test which are in common use today. They are multiple-choice, matching, true-false, and completion items.

4.1.1   Multiple Choice Questions

Multiple-choice test items consist of a stem or a question and three or more alternative answers (options) with the correct answer sometimes called the keyed response and the incorrect answers called distracters. This form is generally better than the incomplete stem because it is simpler and more natural.

Grounlund (1995) writes that the multiple choice question is probably the most popular as well as the most widely applicable and effective type of objective test. Student selects a single response from a list of options. It can be used effectively for any level of course outcome. It consists of two parts: the stem, which states the problem and a list of three to five alternatives, one of which is the correct (key) answer and the others are distracters (incorrect options that draw the less knowledgeable pupil away from the correct response). Multiple choice questions consist of three obligatory parts:

1.   The question ("body of the question")
2.   The correct answer ("the key of the question")
3.   Several incorrect alternatives (the so called "distracters")
         and optional (and especially valuable in self-assessment)
4.   Feedback comment on the student's answer.

The stem may be stated as a direct question or as an incomplete statement. For example:

Direct question

Which is the capital city of Pakistan?   ----------------   (Stem)

A.   Paris.   --------------------------------------   (Distracter)
B.   Lisbon. ------------------------------------   (Distracter)
C.   Islamabad. ---------------------------------   **(Key)**
D.   Rome. --------------------------------------   (Distracter)

Incomplete Statement

The capital city of Pakistan is

A.   Paris.
B.   Lisbon.
C.   Islamabad.
D.   Rome.

Multiple choice questions are composed of one question with multiple possible answers (options), including the correct answer and several incorrect answers (distracters). Typically, students select the correct answer by circling the associated number or letter, or filling in the associated circle on the machine-readable response sheet. Students can generally respond to these types of questions quite quickly. As a result, they are often used to test student's knowledge of a broad range of content. Creating these questions can be time consuming because it is often difficult to generate several plausible distracters. However, they can be marked very quickly.

Multiple Choice Questions Good for:

- Application, synthesis, analysis, and evaluation levels

RULES FOR WRITING MULTIPLE-CHOICE QUESTIONS

There are several rules we can follow to improve the quality of this type of written examination.

1. Examine only the Important Facts!

Make sure that every question examines only the important knowledge. Avoid detailed questions - each question has to be relevant for the previously set instructional goals of the course.

2. Use Simple Language!

Use simple language, taking care of spelling and grammar. Spelling and grammar mistakes (unless you are testing spelling or grammar) only confuse students. Remember that you are examining knowledge about your subject and not language skills.

3. Make the Questions Brief and Clear!

Clear the text of the body of the question from all superfluous words and irrelevant content. It helps students to understand exactly what is expected of them. It is desirable to formulate a question in such way that the main part of the text is in the body of the question, without being repeated in the answers.

4. Form the Questions Correctly!

Be careful that the formulation of the question does not (indirectly) hide the key to the correct answer. Student (adept at solving tests) will be able to recognize it easily and will find the right answer because of the word combination, grammar etc, and not because of their real knowledge.

**5.	Take into Consideration the Independence of Questions!**

Be careful not to repeat content and terms related to the same theme, since the answer to one question can become the key to solve another.

**6.	Offer Uniform Answers!**

All offered answers should be unified, clear and realistic. For example, unlikely realisation of an answer or uneven text quantity of different answers can point to the right answer. Such a question does not test real knowledge. The position of the key should be random. If the answers are numbers, they should be listed in an ascending order.

**7.	Avoid Asking Negative Questions!**

If you use negative questions, negation must be emphasized by using CAPITAL letters, e.g. "Which of the following IS NOT correct..." or "All of the following statements are true, EXCEPT...".

**8.	Avoid Distracters in the Form of "All the answers are correct" or "None of the Answers is Correct"!**

Teachers use these statements most frequently when they run out of ideas for distracters. Students, knowing what is behind such questions, are rarely misled by it. Therefore, if you do use such statements, sometimes use them as the key answer. Furthermore, if a student recognizes that there are two correct answers (out of 5 options), they will be able to conclude that the key answer is the statement "all the answers are correct", without knowing the accuracy of the other distracters.

**9.	Distracters must be Significantly Different from the Right Answer (key)!**

Distracters which only slightly differ from the key answer are bad distracters. Good or strong distracters are statements which themselves seem correct, but are not the correct answer to a particular question.

**10.	Offer an Appropriate Numbers of Distracters.**

The greater the number of distracters, the lesser the possibility that a student could guess the right answer (key). In higher education tests questions with 5 answers are used most often (1 key + 4 distracters). That means that a student is 20% likely to guess the right answer.

Advantages:

Multiple-choice test items are not a panacea. They have advantages and advantages just as any other type of test item. Teachers need to be aware of these characteristics in order to use multiple-choice items effectively.

Advantages

Versatility

Multiple-choice test items are appropriate for use in many different subject-matter areas, and can be used to measure a great variety of educational objectives. They are adaptable to various levels of learning outcomes, from simple recall of knowledge to more complex levels, such as the student's ability to:

•   Analyze phenomena

•   Apply principles to new situations

•   Comprehend concepts and principles

•   Discriminate between fact and opinion

•   Interpret cause-and-effect relationships

•   Interpret charts and graphs

•   Judge the relevance of information

•   Make inferences from given data

•   Solve problems

The difficulty of multiple-choice items can be controlled by changing the alternatives, since the more homogeneous the alternatives, the finer the distinction the students must make in order to identify the correct answer. Multiple-choice items are amenable to item analysis, which enables the teacher to improve the item by replacing distracters that are not functioning properly. In addition, the distracters chosen by the student may be used to diagnose misconceptions of the student or weaknesses in the teacher's instruction.

Validity

In general, it takes much longer to respond to an essay test question than it does to respond to a multiple-choice test item, since the composing and recording of an essay answer is such a slow process. A student is therefore able to answer many multiple-choice items in time it would take to answer a single essay question. This feature enables the teacher using multiple-choice items to test a broader sample of course contents in a given amount of testing time. Consequently, the test scores will likely be more representative of the students' overall achievement in the course.

Reliability

Well-written multiple-choice test items compare favourably with other test item types on the issue of reliability. They are less susceptible to guessing than are true-false test items, and therefore capable of producing more reliable scores. Their scoring is more clear-cut than short answer test item scoring because there are no misspelled or partial answers to deal with. Since multiple-choice items are objectively scored, they are not affected by scorer inconsistencies as are essay questions, and they are essentially immune to the influence of bluffing and writing ability factors, both of which can lower the reliability of essay test scores.

Efficiency

Multiple-choice items are amenable to rapid scoring, which is often done by scoring machines. This expedites the reporting of test results to the student so that any follow-up clarification of instruction may be done before the course has proceeded much further. Essay questions, on the other hand, must be graded manually, one at a time. Overall multiple choice tests are:

- Very effective
- Versatile at all levels
- Minimum of writing for student
- Guessing reduced
- Can cover broad range of content

Disadvantages

Versatility

Since the student selects a response from a list of alternatives rather than supplying or constructing a response, multiple-choice test items are not adaptable to measuring certain learning outcomes, such as the student's ability to:

- Articulate explanations
- Display thought processes
- Furnish information
- Organize personal thoughts.
- Perform a specific task
- Produce original ideas
- Provide examples

Such learning outcomes are better measured by short answer or essay questions, or by performance tests.

Reliability

Although they are less susceptible to guessing than are true false-test items, multiple-choice items are still affected to a certain extent. This guessing factor reduces the reliability of multiple-choice item scores somewhat, but increasing the number of items on the test offsets this reduction in reliability.

Difficulty of Construction

Good multiple-choice test items are generally more difficult and time-consuming to write than other types of test items. Coming up with plausible distracters requires a certain amount of skill. This skill, however, may be increased through study, practice, and experience.

Gronlund (1995) writes that multiple-choice items are difficult to construct. Suitable distracters are often hard to come by and the teacher is tempted to fill the void with a "junk" response. The effect of narrowing the range of options will available to the test wise student. They are also exceedingly time consuming to fashion, one hour per question being by no means the exception. Finally multiple-choice items generally take student longer to complete (especially items containing fine discrimination) than do other types of objective question.

- Difficult to construct good test items.
- Difficult to come up with plausible distracters/alternative responses.

Activity 4.1:	Construct two items of direct question and two items of incomplete statement while following the rules of multiple items.

4.1.2   True/False Questions

A True-False test item requires the student to determine whether a statement is true or false. The chief disadvantage of this type is the opportunity for successful guessing.

According to Gronlund (1995) the alternative response test items that consists of a declaration statement that the pupil is asked to mark true or false, right or wrong, correct or incorrect, yes or no, fact or opinion, agree or disagree and the like. In each case there are only two possible answers. Because the true-false option is the most common, this type is mostly refers to true-false type. Students make a designation about the validity of the statement. Also known as a "binary-choice" item because there are only two options to select from. These types of items are more effective for assessing knowledge, comprehension, and application outcomes as defined in the cognitive domain of Blooms' Taxonomy of educational objectives.

Example

Directions: Circle the correct response to the following statements.

1. Allama Iqbal is the founder of Pakistan. T/F
2. Democracy system is for the people. T/F
3. Quaid-e-Azam was the first Prime Minister of Pakistan. T/F

Good for:

- Knowledge level content
- Evaluating student understanding of popular misconceptions
- Concepts with two logical responses

Advantages:

- Easily assess verbal knowledge
- Each item contains only two possible answers
- Easy to construct for the teacher
- Easy to score for the examiner
- Helpful for poor students
- Can test large amounts of content
- Students can answer 3-4 questions per minute

Disadvantages:

- They are easy to construct.
- It is difficult to discriminate between students that know the material and students who don't know.
- Students have a 50-50 chance of getting the right answer by guessing.
- Need a large number of items for high reliability.
- Fifty percent guessing factor.
- Assess lower order thinking skills.
- Poor representative of students learning achievement.

Tips for Writing Good True/False items:

- Avoid double negatives.
- Avoid long/complex sentences.
- Use specific determinants with caution: never, only, all, none, always, could, might, can, may, sometimes, generally, some, few.

- 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;

Column "A" Capital City	Column "B" Country
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.

Activity 4.3: Construct five fill in the blanks about Pakistan.

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 consistent 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 work, 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 will 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 Grounlund (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.

Activity 4.5: Develop a test of short answers on democracy in Pakistan.

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.

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