

UNIT-7

Planning and Administering Classroom Tests

Written By:
Muhammad Idrees
Reviewed By:
Dr. Naveed Sultana

CONTENTS

| <i>Sr. No</i> | <i>Topic</i> | <i>Page No</i> |
|---------------|---|----------------|
| | Introduction..... | |
| | Objectives | |
| 7.1 | Planning a Test..... | |
| 7.2 | General Consideration in Constructing Objective Test Items | |
| 7.3 | General Consideration in Constructing Essay type Test Items..... | |
| 7.4 | Administering the Test..... | |
| 7.5 | Scoring the Test | |
| 7.6 | Summary of the Unit..... | |
| 7.7 | Activities | |
| 7.8 | Self Assessment Questions | |
| 7.9 | References/Suggested Readings | |

INTRODUCTION

With the approach for increased accountability in the educational system, it is vital that educators are able to apply a wide range of psychometric skills appropriate to the assessment of the students with different pace of learning and backgrounds. It is equally critical that educators have a comprehensive understanding of current measurement and evaluative trends such as competency testing, performance assessment, curriculum-based assessment and standardized assessment.

There are six major steps in planning and conduction an assessment: defining instructional objectives, outlining course contents, developing a test specification, construction of test items, administration of assessment and interpreting test scores.

There are two types of assessment i.e. assessment of learning & assessment for learning. Assessment of learning is used to assess students learning achievement at terminal stages whereas assessment for learning enhances teaching–learning process. Tests and assessments are an essential part of the instructional process. When properly done, they can not only effectively evaluate but also enhance students' learning and teachers' instruction. When poorly done, they can confuse and alienate students, distort the curriculum, and hinder good instruction. Test scores and grades sometimes affect "high-stakes" decisions about students, prompting intense concern that they be accurate and fair.

This course is designed to provide the students/prospective teachers with the principles and techniques necessary to develop sound student assessment strategies. The primary focus of the course will be on writing instructional objectives, developing different types of test items (selected response & constructed response), utilizing performance based and alternative assessment techniques, administering classroom evaluation procedures and interpreting test score for different purposes. There are eight major steps in planning and conducting an assessment:

- defining instructional objectives,
- outlining course contents,
- developing a test specification,
- selection of appropriate assessment tasks,
- preparation of relevant assessment tasks or construction of test items,
- assembly of assessment tasks,
- administration of assessment
- Interpreting test scores.

These steps will be discussed in detail in the unit. This unit will be of great value for the teachers/prospective teachers in developing and assembling suitable test to assess students learning achievements.

OBJECTIVES

After intensive study of this unit, the students will be capable to:

- appreciate qualities needed to determine the quality of classroom tests.
- develop different types of test items for selected response test items
- develop different types of test items for constructed response test items
- efficiently administer classroom test
- utilize the techniques of objectively score and grade tests.

7.1 Planning a Test

The main objective of classroom assessment is to obtain valid, reliable and useful data regarding student learning achievement. This requires determining what is to be measured and then defining it precisely so that assessments tasks to measure desired performance can be developed. Classroom tests and assessments can be used for the following instructional objectives:

i. Pre-testing

Tests and assessments can be given at the beginning of an instructional unit or course to determine:-

- weather the students have the prerequisite skills needed for the instruction (readiness, motivation etc)
 - to what extent the students have already achieved the objectives of planned instruction (to determine placement or modification of instruction)
- ii. During the Instruction Testing
- provides bases for formative assessment
 - monitor learning progress
 - detect learning errors
 - provide feedback for students and teachers
- iii. End of Instruction Testing
- measure intended learning outcomes
 - used for formative assessment
 - provides bases for grades, promotion etc

Prior to developing an effective test, one needs to determine whether or not a test is the appropriate type of assessment. If the learning objectives are of primarily types of procedural knowledge (how to perform a task) then a written test may not be the best approach. Assessment of procedural knowledge generally calls for a performance demonstration assessed using a rubric. Where demonstration of a procedure is not appropriate, a test can be an effective assessment tool.

The first stage of developing a test is planning the test content and length. Planning the test begins with development of a blueprint or test specifications for the test structured on the learning outcomes or instructional objectives to be assessed by the test instrument. For each learning outcome, a weight should be assigned based on the relative importance of that outcome in the test. The weight will be used to determine the number of items related to each of the learning outcomes.

7.1.1 Test Specifications

When an engineer prepares a design to construct a building and choose the materials, he intends to use in construction, he usually know what a building is going to be used for, and therefore designs it to meet the requirements of its planned inhabitants. Similarly, in testing, table of specification is the blueprint of the assessment which specifies percentages and weightage of test items and measuring constructs. It includes constructs

and concepts to be measured, tentative weightage of each construct, specify number of items for each concept, and description of item types to be constructed. It is not surprising that specifications are also referred to as ‘blueprints’, for they are literally architectural drawings for test construction. Fulcher & Davidson (2009) divided test specifications into the following four elements:

- **Item specifications:** Item specifications describe the items, prompts or tasks, and any other material such as texts, diagrams, and charts which are used as stimuli. Typically, a specification at this sub-level contains two key elements: samples of the tasks to be produced, and guiding language that details all information necessary to produce the task.
- **Presentation Model:** Presentation model provides information how the items and tasks are presented to the test takers.
- **Assembly Model:** Assembly model helps the test developer to combine test items and tasks to develop a test format.
- **Delivery Model:** Delivery Model tells how the actual test is delivered. It includes information regarding test administration, test security/confidentiality and time constraint.

Table 7.1: Table of Specifications for Social Studies Class VI

| Objectives/ Contents | Knowledge | | | Understanding | | | Application | | | Percentage |
|-------------------------|-----------|----------|-----------|---------------|----------|-----------|-------------|----------|-----------|-------------|
| | LA | SA | MCQ | LA | SA | MCQ | LA | SA | MCQ | |
| Climate | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 | 25% |
| Resources | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 | 25% |
| Population | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 | 25% |
| Society | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 | 25% |
| Total | 4 | 8 | 12 | 4 | 8 | 12 | 4 | 8 | 12 | 100% |

LA: Long Answer, SA: Short Answers, MCQ: Multiple Choice Questions

Note: Number of items/questions and percentage may be changed according to the objectives/contents and hierarchy of learning.

7.2 General Consideration in Constructing Objective Test Items

The second step in test planning is determining the format and length of the test. The format is based on the different types of items to be included in the test. The construction of valid and good test items is a skill just like effective teaching. Some rules are to be followed and some techniques are to be used to construct good test items. Test items can be used to assess student’s ability to recognize concepts or to recall concepts. Generally there are two types of objective test items:-

- i. Select type.
- ii. Supply type.

7.2.1 Select Type Items

A. Matching Items

According to W. Wiersma and S.G. Jurs (1990), 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 |
| Hague | Netherlands |
| | Turkey |
| | West Germany |

According to W. Wiersma and S.G. Jurs (1990) in some matching exercises the number of premises and responses are the same, termed a balanced or perfect matching exercise. In others, the number and responses may be different.

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.

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.

Suggestions for Writing Matching Items

Here are some suggestions for writing matching items:

- i. 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.
- ii. Make sure that all the options are plausible distracters for each description to ensure homogeneity of lists.
- iii. 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.
- iv. Each description in the list should be numbered (each is an item), and the list of options should be identified by letter.
- v. 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.
- vi. In the directions, specify the basis for matching and whether options can be used more than once.

B. Multiple Choice Questions (MCQ’s)

Norman E. Grounlund (1990) 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 (“foils” or incorrect options that draw the less knowledgeable pupil away from the correct response).

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. Lahore. -----(Distracter)
B. Karachi. -----(Distracter)
C. Islamabad. ----- **(Key)**
D. Peshawar. ----- (Distracter)

Incomplete Statement

The capital city of Pakistan is

- A. Lahore.
- B. Karachi.
- C. Islamabad.
- D. Peshawar.

RULES FOR WRITING MULTIPLE-CHOICE QUESTIONS

1. Use Plausible Distracters (wrong-response options)

- Only list plausible distracters, even if the number of options per question changes
- Write the options so they are homogeneous in content
- Use answers given in previous open-ended exams to provide realistic distracters

2. Use a Question Format

- Experts encourage multiple-choice items to be prepared as questions (rather than incomplete statements)

Incomplete Statement Format:

The capital of AJK is in-----.

Direct Question Format:

In which of the following cities is the capital of AJK?

3. Emphasize Higher-Level Thinking

- Use memory-plus application questions. These questions require students to recall principles, rules or facts in a real life context.
- The key to prepare memory-plus application questions is to place the concept in a life situation or context that requires the student to first recall the facts and then apply or transfer the application of those facts into a situation.
- Seek support from others who have experience writing higher-level thinking multiple-choice questions.

EXAMPLES:

Memory Only Example (Less Effective)

Which description best characterizes whole foods?

- a. orange juice

- b. toast
- c. bran cereal
- d. grapefruit

Memory-Plus Application Example (More Effective)

Sana's breakfast this morning included one glass of orange juice (from Concentrate), one slice of toast, a small bowl of bran cereal and a grapefruit. What "whole food" did Sana eat for breakfast?

- a. orange juice
- b. toast
- c. bran cereal
- d. grapefruit

Memory-Plus Application Example

Ability to Interpret Cause-and-Effect Relationships Example

Why does investing money in common stock protect against loss of assets during inflation?

- a. It pays higher rates of interest during inflation.
- b. It provides a steady but dependable income despite economic conditions.
- c. It is protected by the Federal Reserve System.
- d. It increases in value as the value of a business increases.

Ability to Justify Methods and Procedures Example

Why is adequate lighting necessary in a balanced aquarium?

- a. Fish need light to see their food.
- b. Fish take in oxygen in the dark.
- c. Plants expel carbon dioxide in the dark.
- d. Plants grow too rapidly in the dark.

4. Keep Option Lengths Similar

- Avoid making your correct answer the long or short answer

- 5. Balance the Placement of the Correct Answer**
 - Correct answers are usually the second and third option

- 6. Be Grammatically Correct**
 - Use simple, precise and unambiguous wording
 - Students will be more likely to select the correct answer by finding the grammatically correct option

- 7. Avoid Clues to the Correct Answer**
 - Avoid answering one question in the test by giving the answer somewhere else in the test
 - Have the test reviewed by someone who can find mistakes, clues, grammar and punctuation problems before you administer the exam to students
 - Avoid extremes – never, always, only
 - Avoid nonsense words and unreasonable statements

- 8. Avoid Negative Questions**
 - 31 of 35 testing experts recommend avoiding negative questions
 - Students may be able to find an incorrect answer without knowing the correct answer

- 9. Use Only One Correct Option (Or be sure the best option is clearly the best option)**
 - The item should include one and only one correct or clearly best answer
 - With one correct answer, alternatives should be mutually exclusive and not overlapping
 - Using MC with questions containing more than one right answer lowers discrimination between students

- 10. Give Clear Instructions**

Such as:

 - Questions 1 - 10 are multiple-choice questions designed to assess your ability to remember or recall basic and foundational pieces of knowledge related to this course.
 - Please read each question carefully before reading the answer options. When you have a clear idea of the question, find your answer and mark your selection on the answer sheet. Please do not make any marks on this exam.

- Questions 11 – 20 are multiple-choice questions designed to assess your ability to think critically about the subject.
 - Please read each question carefully before reading the answer options.
 - Be aware that some questions may seem to have more than one right answer, but you are to look for the one that makes the most sense and is the most correct.
 - When you have a clear idea of the question, find your answer and mark your selection on the answer sheet.
 - You may justify any answer you choose by writing your justification on the blank paper provided.
- 11. Use Only a Single, Clearly-Defined Problem and Include the Main Idea in the Question**
- Students must know what the problem is without having to read the response options
- 12. Avoid “All the Above” Option**
- Students merely need to recognize two correct options to get the answer correct
- 13. Avoid the “None of the Above” Option**
- You will never know if students know the correct answer
- 14. Don’t Use MCQ When Other Item Types Are More Appropriate**
- Limited distracters or assessing problem-solving and creativity

Advantages

The chief advantage of the multiple-choice question according to N.E. Gronlund (1990) is its versatility. For instance, it is capable of being applied to a wide range of subject areas. In contrast to short answer items limit the writer to those content areas that are capable of being stated in one or two words, multiple choice item necessary bound to homogeneous items containing one type of subject matter as are matching items, and a multiple choice question greatly reduces the opportunity for a student to guess the correct answer from one choice in two with a true – false items to one in four or five, there by increasing the reliability of the test. Further, since a multiple – choice item contains plausible incorrect or less correct alternative, it permits the test constructor to fine tune the discriminations (the degree or homogeneity of the responses) and control the difficulty level of the test.

Disadvantages

N.E. Gronlund (1990) 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 they generally take student longer to complete (especially items containing fine discrimination) than do other types of objective question.

Suggestions for Writing MCQ's Items

Here are some guidelines for writing multiple-choice tests:

- I. The stem of the item should clearly formulate a problem. Include as much of the item as possible, keeping the response options as short as possible. However, include only the material needed to make the problem clear and specific. Be concise – don't add extraneous information.
- II. Be sure that there is one and only one correct or clearly best answer.
- III. Be sure wrong answer choices (distracters) are plausible. Eliminate unintentional grammatical clues, and keep the length and form of all the answer choices equal. Rotate the position of the correct answer from item to item randomly.
- IV. Use negation questions or statements only if the knowledge being tested requires it. In most cases it is more important for the student to know what a specific item of information is rather than what it is not.
- V. Include from three to five options (two to four distracters plus one correct answer) to optimize testing for knowledge rather than encouraging guessing. It is not necessary to provide addition distracters from an item simply to maintain the same number of distracters for each item. This usually leads to poorly constructed distracters that add nothing to test validity and reliability.
- VI. To increase the difficulty of a multiple-choice item, increase the similarity of content among the options.
- VII. Use the option “none of the above” sparingly and only when the keyed answer can be classified unequivocally as right or wrong.
- VII. Avoid using “all of the above”. It is usually the correct answer and makes the item too easy for students with partial information.

II. Supply Type Items

A. 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:**
Motorway (M1) opened for traffic in _____.
 - b. **Better item:**
Motorway (M1) opened for traffic 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.
 - c. **Poor item:**
The _____ type of test item is usually more _____ than the _____ type.
 - d. **Better item:**
The supply type of test item is usually graded less objectively than the _____ type.
- V. Word the statement such that the blank is near the end of the sentence rather than near the beginning. This will prevent awkward sentences.
- VI. If the problem requires a numerical answer, indicate the units in which it is to be expressed.

B. 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.

Advantages

Norman E. Gronlund (1990) 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

According to Norman E. Gronlund (1990) 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.

7.5 General Consideration in Constructing Essay type Test Items

Robert L. Ebel and David A. Frisbie (1991) in their book, write that "teachers are often as concerned with measuring the ability of students to think about and use knowledge as they are with measuring the knowledge their students possess. In these instances, tests are needed that permit students some degree of latitude in their responses. Essay tests are adapted to this purpose. Student writes a response to a question that is several paragraphs to several pages long. Essays can be used for higher learning outcomes such as synthesis or evaluation as well as lower level outcomes. They provide items in which students supply rather than select the appropriate answer, usually the students compose a response in one or more sentences. Essay tests allow students to demonstrate their ability to recall, organize, synthesize, relate, analyze and evaluate ideas.

Types of Essay Tests

Essay tests may be divided into many types. W.S. Monree and R.I. Cater (1993) divide essay tests into the many categories like Selective recall-basis given, Evaluation recall-basis given, Comparison of two things on a single designated basis, Comparison of two things in general, Decisions – For or against, cause and effect, explanation of the use or exact meaning of some word, phrase or statement, summary of some unit of the text book or article, Analysis, statement of relationships, Illustration or examples, classification, application of rules, laws, or principles to new situation, discussion, statement of an author's purpose in the selection or organization of material, Criticism – as to the adequacy, correctness or relevance of a printed statement or to a class mate's answer to a question on the lesson, reorganization of facts, formulation of new question – problems and question raised, new methods of procedure etc.

Types of Constructed Response Items

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. Thus essay type items are of two types:-

- Restricted Response Essay Items
- Extended Response Essay Items

I. 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

II. 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.

Scoring Essay Type Items

A rubric or scoring criteria is developed to evaluate/score an essay type item. A rubric is a scoring guide for subjective assessments. It is a set of criteria and standards linked to learning objectives that are used to assess a student's performance on papers, projects, essays, and other assignments. Rubrics allow for standardized evaluation according to specified criteria, making grading simpler and more transparent. A rubric may vary from simple checklists to elaborate combinations of checklist and rating scales. How elaborate your rubric is depends on what you are trying to measure. If your essay item is

A restricted-response item simply assessing mastery of factual content, a fairly simple listing of essential points would be sufficient. An example of the rubric of restricted response item is given below.

Test Item:

Name and describe five of the most important factors of unemployment in Pakistan. (10 points)

Rubric/Scoring Criteria:

- (i) 1 point for each of the factors named, to a maximum of 5 points
- (ii) One point for each appropriate description of the factors named, to a maximum of 5 points
- (iii) No penalty for spelling, punctuation, or grammatical error
- (iv) No extra credit for more than five factors named or described.
- (v) Extraneous information will be ignored.

However, when essay items are measuring higher order thinking skills of cognitive domain, more complex rubrics are mandatory. An example of Rubric for writing test in language is given below.

Table 7.2: Scoring Criteria (Rubrics) for Essay Type Item for 8th grade

| Sr. No. | Criteria | Unsatisfactory | Proficient | Advance |
|---------|-------------------------------------|--|---|---|
| 1 | Length | Length of Text will be according to the Prompt | Length of Text will be according to the Prompt | Length of Text will be according to the Prompt |
| 2 | Layout | Writing is not according to the provided format | Writing is according to the provided format to some extent | Writing is completely according to the provided format |
| 3 | Vocabulary | Expected KEY WORDS* are not used | Expected KEY WORDS* are used to some extent | Expected KEY WORDS* are used mostly |
| 4 | Spelling | Spellings of most words are incorrect | Spellings of some words are incorrect | Spellings of all words are correct |
| 5 | Selection and Organization of Ideas | Few ideas are relevant to the task and the given task organization | Some ideas are relevant to the task and the given task organization | Almost all ideas are relevant to the task and the given task organization |
| 6 | Punctuation | Very few Punctuation Marks are used | Some Punctuation Marks are used | Almost all Punctuation Marks are used |
| 7 | Grammar | Use of some basic GRAMMAR RULES** | Occasional use of basic GRAMMAR RULES** | Use of some basic GRAMMAR RULES** |

* KEY WORDS: Expected Key Words will be provided for each Writing Prompt

Advantages of Essay Type Items

The main advantages of essay type tests are as follows:

- (i) They can measure complex learning outcomes which cannot be measured by other means.
- (ii) They emphasize integration and application of thinking and problem solving skills.
- (iii) They can be easily constructed.
- (iv) They give examinees freedom to respond within broad limits.
- (v) The students cannot guess the answer because they have to supply it rather than select it.
- (vi) Practically it is more economical to use essay type tests if number of students is small.
- (vii) They require less time for typing, duplicating or printing. They can be written on the blackboard also if number of students is not large.
- (viii) They can measure divergent thinking.
- (ix) They can be used as a device for measuring and improving language and expression skill of examinees.
- (x) They are more helpful in evaluating the quality of the teaching process.
- (xi) Studies have supported that when students know that the essay type questions will be asked, they focus on learning broad concepts and articulating relationships, contrasting and comparing.
- (xii) They set better standards of professional ethics to the teachers because they expect more time in assessing and scoring from the teachers.

Limitations of Essay Type Items

The essay type tests have the following serious limitations as a measuring instrument:

- (i) A major problem is the lack of consistency in judgments even among competent examiners.
- (ii) They have halo effects. If the examiner is measuring one characteristic, he can be influenced in scoring by another characteristic. For example, a well behaved student may score more marks on account of his good behaviour also.
- (iii) They have question to question carry effect. If the examinee has answered satisfactorily in the beginning of the question or questions he is likely to score more than the one who did not do well in the beginning but did well later on.
- (iv) They have examinee to examinee carry effect. A particular examinee gets marks not only on the basis of what he has written but also on the basis that whether the previous examinee whose answer book was examined by the examiner was good or bad.

- (v) They have limited content validity because of sample of questions can only be asked in essay type test.
- (vi) They are difficult to score objectively because the examinee has wide freedom of expression and he writes long answers.
- (vii) They are time consuming both for the examiner and the examinee.
- (viii) They generally emphasize the lengthy enumeration of memorized facts.

Suggestions for Writing Essay Type Items

- I. Ask questions or establish tasks that will require the student to demonstrate command of essential knowledge. This means that students should not be asked merely to reproduce material heard in a lecture or read in a textbook. To "demonstrate command" requires that the question be somewhat novel or new. The substance of the question should be essential knowledge rather than trivia that might be a good board game question.
- II. Ask questions that are determinate, in the sense that experts (colleagues in the field) could agree that one answer is better than another. Questions that contain phrases such as "What do you think..." or "What is your opinion about..." are indeterminate. They can be used as a medium for assessing skill in written expression, but because they have no clearly right or wrong answer, they are useless for measuring other aspects of achievement.
- III. Define the examinee's task as completely and specifically as possible without interfering with the measurement process itself. It is possible to word an essay item so precisely that there is one and only one very brief answer to it. The imposition of such rigid bounds on the response is more limiting than it is helpful. Examinees do need guidance, however, to judge how extensive their response must be considered complete and accurate.
- IV. Generally give preference to specific questions that can be answered briefly. The more questions used, the better the test constructor can sample the domain of knowledge covered by the test. And the more responses available for scoring, the more accurate the total test scores are likely to be. In addition, brief responses can be scored more quickly and more accurately than long, extended responses, even when there are fewer of the latter type.
- V. Use enough items to sample the relevant content domain adequately, but not so many that students do not have sufficient time to plan, develop, and review their responses. Some instructors use essay tests rather than one of the objective types because they want to encourage and provide practice in written expression. However, when time pressures become great, the essay test is one of the most unrealistic and negative writing experiences to which students can be exposed. Often there is no time for editing, for rereading, or for checking spelling. Planning time is short changed so that writing time will not be. There are few, if any, real writing tasks that require such conditions. And there are few writing experiences that discourage the use of good writing habits as much as essay testing does.

- VI. Avoid giving examinees a choice among optional questions unless special circumstances make such options necessary. The use of optional items destroys the strict comparability between student scores because not all students actually take the same test. Student A may have answered items 1-3 and Student B may have answered 3-5. In these circumstances the variability of scores is likely to be quite small because students were able to respond to items they knew more about and ignore items with which they were unfamiliar. This reduced variability contributes to reduced test score reliability. That is, we are less able to identify individual differences in achievement when the test scores form a very homogeneous distribution. In sum, optional items restrict score comparability between students and contribute to low score reliability due to reduced test score variability.
- VII. Test the question by writing an ideal answer to it. An ideal response is needed eventually to score the responses. If it is prepared early, it permits a check on the wording of the item, the level of completeness required for an ideal response, and the amount of time required to furnish a suitable response. It even allows the item writer to determine if there is any "correct" response to the question.
- VIII. Specify the time allotment for each item and/or specify the maximum number of points to be awarded for the "best" answer to the question. Both pieces of information provide guidance to the examinee about the depth of response expected by the item writer. They also represent legitimate pieces of information a student can use to decide which of several items should be omitted when time begins to run out. Often the number of points attached to the item reflects the number of essential parts to the ideal response. Of course if a definite number of essential parts can be determined, that number should be indicated as part of the question.
- IX. Divide a question into separate components when there are obvious multiple questions or pieces to the intended responses. The use of parts helps examinees organizationally and, hence, makes the process more efficient. It also makes the grading process easier because it encourages organization in the responses. Finally, if multiple questions are not identified, some examinees may inadvertently omit some parts, especially when time constraints are great.

7.3 Administering the Test

I. Test Assembly

We have discussed various aspects of test planning and construction. If you have written instructional objectives, constructed a test, and written items that match your objectives, then more than likely you will have a good test. All the "raw material" will be there. However, sometimes the raw material, as good as it may be, can be rendered useless because of poorly assembled and administered test. By now you know it requires a substantial amount of time to write objectives, put together a test blueprint, and write items. It is worth a little more time to properly assemble or package your test so that your efforts will not be wasted. Assembly of the test comprises the following steps:-