

Course Code 8602

Course Name Educational Assessment and Evaluation

Assignment No 1

Question No 1

Explain classroom Assessment. Write a note on principles of classroom assessment.

Ans.

Classroom assessment is something all teachers do as part of their responsibility for enhancing student learning and motivation, documenting student performance, and reporting results for accountability purposes. There are many different types of classroom assessments that can be used, each with advantages and disadvantages. To the extent that teachers base their assessments on known indicators of high quality, such as reliability and validity, align their assessments with clear and appropriate learning objectives and targets, use formative assessment with meaningful feedback, and grade student work to provide meaningful summary indications of performance, student learning and motivation will be enhanced, as well as documented.

Classroom Assessment Techniques, also referred to as CATs, are strategies educators use to gauge how well students are comprehending key points during a lesson or a course. The techniques are meant to be a type of formative assessment that also allow teachers to make adjustments to a lesson based on students' needs. CATs are most commonly ungraded, unanimous, and are conducted during class time.

Principals of classroom assessment:

A good strategy for using CATs is the following.

Decide what you want to assess about your students' learning from a CAT. Choose a CAT that provides this feedback, is consistent with your teaching style, and can be implemented easily in your class. Explain the purpose of the activity to students, and then conduct it.

After class, review the results, determine what they tell you about your students' learning, and decide what changes to make, if any. Let your students know what you learned from the CAT and how you will use this information.

Examples of CATs include the following..

The Background knowledge simple questionnaire given to students at the start of a course, or before the on or topic. It is designed to uncover students' pre-conceptions.

The Minute Paper tests how students are knowledge, or not. The instructor ends class by asking students to write brief response following questions: "What was the most important thing you learned during at class at important question remains unanswered?"

The Muddiest Point the s' CATs to help assess where students are having difficulties. The technique asking students to note down a quick response to one question: "What was the muddiest" The term "muddiest" most or "most confusing."

The What's principle. is useful in courses requiring problem-solving. After students figure out what type of problem they are dealing with, they often must decide what principle(s) to apply in order to solve the problem. This CAT provides students with a few problems and asks them to state the principle that best applies to each problem.

Question No 2

How contact outline is prepared while developing classroom tests.

Ans.

Basic Steps in Classroom Assessment

1. Determining the purpose of the assessment (pre-test, formative, or summative)
2. Developing the test specifications (this is the table you are creating)
3. Selecting the appropriate assessment tasks (form and type)
4. Prepare the relevant assessment tasks
5. Assemble the assessment
6. Provide instruction
7. Evaluate the assessment
8. Use the assessment results .

1. Determining the purpose of the assessment

Pre-testing

- 1) Whether students have the prerequisite skills needed for the instruction
- 2) To what extent students have already achieved the objectives of the planned instruction -- are confined to a limited domain - low level of difficulty - serve as a basis for remedial work or for adaptation of instructional plans - not usually different from post test (an equivalent form)

3 During instruction assessment

1. This is called diagnostic or formative assessment; done about midway through a unit or chapter

- 1) To monitor learning progress
 - 2) Provide feedback to students and teachers
 - 3) Detect learning errors, diagnostic - practice tests, quizzes - predefined segment of instruction - limited sample of learning outcomes
2. End of instruction assessment

This is called summative assessment and measures the extent to which the intended learning outcomes have been achieved; can serve the same purposes as pre-testing (for the following unit) and formative assessment

2. Developing the specifications for tests and assessments (this is the table you are creating)

Steps:

- 1) Prepare a list of instructional objectives
- 2) Outline course content
- 3) Prepare a two-way table / chart; table is limited to those objectives that are measurable

3. Selecting the appropriate assessment tasks [two forms: objective and performance]

First Form = Objective

Objective items – highly structured; single right answer; limits type of response student can make; scoring is quick, easy, and accurate

Supply types

- 1) Short answer
- 2) Completion

Selection types: (1) alternate choice (2) matching (3) multiple choice (4) keyed response (5) interpretive exercise

3. **Second Form = Performance**

Performance items -- less structure (problem can be redefined and the answer organized and presented in their own words); scoring is more difficult and less reliable

Essay questions:

- 1) Extended-response
- 2) Restricted response

Active (evaluates process):

- 1) Construction of graphs, diagrams, models
- 2) Use of equipment or playing an instrument.

Question No 3

what are the types of achievements tests? For what purposes achievement tests are used?

Ans.

Meaning of Achievement Test:

Achievement test is most probably the very important area of appraisal for a guidance programme for the benefit of the individual. Scores on achievement test are excellent means for evaluating educational (academic) attainments and for the individual in the concerned area of the subject covered by the test.

Types of Achievement Test:

Achievement tests can be of various categories basing on form, purpose, time, method and subject area. Achievement test can be of different forms like oral test, written test and practical test. Items of achievement test can be essay type questions or short answer questions or objective type of questions or combination of all these types.

Achievement test may be of different types on the basis of the purpose for which it is administered. They are diagnostic tests, prognostic test, accuracy test, power test, spit test etc. Achievement tests can be administered in different period of time. When it is based on time or period factor, the test is summative test, daily test, weekly test, fortnightly test, monthly test, quarterly test, half yearly test, annual test or final examination at the end of course of study of an academic year.

On the basis of content or subject matter, achievement tests are categorized as language test, reading test, spelling test, history test, geography test, mathematic test, science test etc. Broadly speaking, all these achievement tests can be divided into two on the basis of quality that is standardized test and teacher made test. Here let us have a discussion on the objective type of achievement test.

Use of achievement tests:

Achievement tests are often used in educational and training settings. In schools, for example, achievements tests are frequently used to determine the level of education for which students might be prepared. Students might take such a test to determine if they are ready to enter into a particular grade level or if they are ready to pass of a particular subject or grade level and move on to the next.

Standardized achievement tests are also used extensively in educational settings to determine if students have met specific learning goals. Each grade level has certain educational expectations, and testing is used to determine if schools, teachers, and students are meeting those standards.

So how exactly are achievement tests created? In many instances, subject matter experts help determine what content standards should exist for a certain subject. These standard represent the things that an individual at a certain skill or grade level should know about a particular subject. Test designers can then use this information to develop exams that accurately reflect the most important things that a person should know about that topic.

Achievement tests play an important role in education, but they have also been the subject of criticism at times. Some feel that excessive testing interferes with the educational process and places too much emphasis on passing a test while ignoring more important abilities such as

critical and creative thinking. However, such tests do provide a fairly efficient way to get an idea of how well students are performing.

Question No 4

Advantage and disadvantages of multiple type of questions.

Ans.

A multiple choice question is a question type where the respondent is asked to choose one or more items from a limited list of choices. A multiple choice question consists of a stem, the correct answer, and distractors.

The stem is the beginning part of the item that presents the item as a problem to be solved, a question, or an incomplete statement to be completed. The options are the possible answers you can choose from, with the correct answer called the key and the incorrect answers called distractors.

Multiple choice, objective response, or MCQ is a form of an objective assessment in which respondents are asked to select only correct answers from the choices offered as a list.

Advantages of multiple choice questions:

- They have fast processing times
- There's no room for subjectivity
- You can ask more questions, it takes less time to complete a multiple choice question compared to an open question
- Respondents don't have to formulate an answer but can focus on the content.

Disadvantages of multiple choice questions:

- While they are fast processed, they are **time-consuming to create**: they require time to draw up effective stem questions and corresponding choices
- They don't produce any qualitative data, solely **quantitative**

- They **limit** the respondent in his **answers**, that's why it's important to provide an "other" option with a textbox.

Question No 5.

Write a detailed note on different types of reliability of test.

Ans.

Reliability tells you how consistently a method measures something. When you apply the same method to the same sample under the same conditions, you should get the same results. If not, the method of measurement may be unreliable.

There are four main types of reliability. Each can be estimated by comparing different sets of results produced by the same the same method.

- **Test-retest.**
- **Interrater.**
- **Parallel forms.**
- **Internal consistency.**

Test-retest reliability:

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.

Why it's important:

Many factors can influence your results at different points in time: for example, respondents might experience different moods, or external conditions might affect their ability to respond accurately.

Test-retest reliability can be used to assess how well a method resists these factors over time. The smaller the difference between the two sets of results, the higher the test-retest reliability.

How to measure it:

To measure test-retest reliability, you conduct the same test on the same group of people at two different points in time. Then you calculate the correlation between the two sets of results.

Interrater reliability

Interrater reliability measures the degree of agreement between different people observing or assessing the same thing. You use it when data is collected by researchers assigning ratings, scores or categories to one or more variables.

Why it's important:

People are subjective, so different observers' perceptions of situations and phenomena naturally differ. Reliable research aims to minimize subjectivity as much as possible so that a different researcher could replicate the same results.

When designing the scale and criteria for data collection, it's important to make sure that different people will rate the same variable consistently with minimal bias. This is especially important when there are multiple researchers involved in data collection or analysis.

How to measure it:

To measure interrater reliability, different researchers conduct the same measurement or observation on the same sample. Then you calculate the correlation between their different sets of results. If all the researchers give similar ratings, the test has high interrater reliability.

Parallel forms reliability:

Parallel forms reliability measures the correlation between two equivalent versions of a test. You use it when you have two different assessment tools or sets of questions designed to measure the same thing.

Why it's important:

If you want to use multiple different versions of a test (for example, to avoid respondents repeating the same answers from memory), you first need to make sure that all the sets of questions or measurements give reliable results.

How to measure it:

The most common way to measure parallel forms reliability is to produce a large set of questions to evaluate the same thing, then divide these randomly into two question sets.

The same group of respondents answers both sets, and you calculate the correlation between the results. High correlation between the two indicates high parallel forms reliability.

Internal consistency:

Internal consistency assesses the correlation between multiple items in a test that are intended to measure the same construct.

You can calculate internal consistency without repeating the test or involving other researchers, so it's a good way of assessing reliability when you only have one data set.

Why it's important:

When you devise a set of questions or ratings that will be combined into an overall score, you have to make sure that all of the items really do reflect the same thing. If responses to different items contradict one another, the test might be unreliable.

How to measure it:

Two common methods are used to measure internal consistency.

Average inter-item correlation: For a set of measures designed to assess the same construct, you calculate the correlation between the results of all possible pairs of items and then calculate the average.

Split-half reliability: You randomly split a set of measures into two sets. After testing the entire set on the respondents, you calculate the correlation between the two sets of responses.

-----End-----

Tajassus.com