Course: Educational Assessment and Evaluation (8602)
Level: B.Ed. (2.5 & 1.5 Years)
Semester: Autumn, 2019
ASSIGNMENT No. 1

Q.1 Highlight the important characteristics of classroom assessment. Also define the terms measurement, evaluation and test.

Important characteristics of classroom assessment:

Learner-Centered

The primary attention of teachers is focused on observing and improving learning.

Teacher-Directed

Individual teachers decide what to assess, how to assess, and how to respond to the information gained through the assessment

Teachers do not need to share results with anyone outside of the class.

Mutually Beneficial

Students are active participants.

Students are motivated by the increased interest of faculty in their success as learners.

Teachers improve their teaching skills and gain new insights.

Formative

Assessments are almost never "graded".

Assessments are almost always anonymous in the classroom and often anonymous online.

Assessments do not provide evidence for evaluating or grading students.

Context-Specific

Assessments respond to the particular needs and characteristics of the teachers, students, and disciplines to which they are applied.

Customize to meet the needs of your students and course.

Ongoing

Classroom assessment is a continuous process.

Part of the process is creating and maintaining a classroom "feedback loop"

Each classroom assessment event is of short duration.

Rooted in Good Teaching Practice

Classroom assessment builds on good practices by making feedback on students' learning more systematic, more flexible, and more effective.

Definition of assessment

The verb assess comes from the French 'assessor', but the origin is from the Medieval Latin 'assess are' meaning "fix a tax upon," Another derivation of the Latin term is 'assidere' or 'adsidere' meaning "to sit beside" (a judge). Reference is made to the assistant of the judge whose job was to fix the amount of a fine or tax by estimating the value of a property.

Assessment is thus the process of collecting information about students from diverse sources so that educators can form an idea of what they know and can do with this knowledge. While evaluation is concerned with making judgments about instruction, a curriculum, or an educational system, assessment is concerned with the students' performance. In other words, one assesses an individual but evaluates a program, a curriculum, an educational system, etc.

The verb 'assess' often collocates with:

- skills,
- abilities,
- performance,
- aptitude,
- Competence.

Assessment occurs when judgments are made about a learner's performance, and entails gathering and organizing information about learners in order to make decisions and judgments about their learning."

Assessment is thus the process of collecting information about learners using different methods or tools (e.g. tests, quizzes, portfolios, etc.).

Educators assess their students for a variety of purposes:

- To evaluate learners' educational needs,
- To diagnose students' academic readiness,
- To measure their progress in a course,
- To measure skill acquisition.
-

There are different types of assessment:

- **Formative assessment:**
  It is process-oriented and is also referred to as 'assessment for Learning'. It is an ongoing process to monitor learning, the aim of which is to provide feedback to improve teacher's instruction methods and improve students learning.
- **Summative assessment:**
  It is product-oriented and is often referred to as 'Assessment of Learning'. It is used to measure student learning progress and achievement at the end of a specific instructional period.
- **Alternative assessment:**
  It is also referred to as authentic or performance assessment. It is an alternative to traditional assessment that relies only on standardized tests and exams. It requires students to do tasks such as presentations, case studies, portfolios, simulations, reports, etc. Instead of measuring what students know, alternative assessment focuses on what students can do with this knowledge.
-

Definition of testing

Simply put, a test refers to a tool, technique or a method that is intended to measure students' knowledge or their ability to complete a particular task. In this sense, testing can be considered as a form of assessment. Tests should meet some basic requirements, such as validity and reliability.

- Validity refers to the extent to which a test measures what it is supposed to measure.
- Reliability refers to the consistency of test scores when administered on different occasions.
-

There are different types of tests:

- **Placement tests:** It is designed to help educators place a student into a particular level or section of a language curriculum or school
- **Diagnostic tests:** they help teachers and learners to identify strengths and weaknesses.
- **Proficiency tests:** they measure a learner's level of language.
- **Achievement tests:** they are intended to measure the skills and knowledge learned after some kind of instruction.

Q.2
a) Explain the cognitive domain of Bloom's Taxonomy of education objective.

Answer:

The Three Domains of Learning Bloom's Taxonomy was created in 1956 under the leadership of educational psychologist Dry Benjamin Bloom in order to promote higher forms of thinking in education, such as analyzing and evaluating concepts, processes, procedures, and principles, rather than just remembering facts (rote learning). It is most often used when designing educational, training, and learning processes.

The committee identified three domains of educational activities or learning (Bloom, et al.

1956): o **Cognitive**: mental skills (knowledge) o **Affective**: growth in

feelings or emotional areas (attitude or self) o **Psychomotor**: manual

or physical skills (skills)

Since the work was produced by higher education, the words tend to be a little bigger than we normally use. Domains may be thought of as categories. Instructional designers, trainers, and educators often refer to these three categories as KSA (Knowledge [cognitive], Skills [psychomotor], and Attitudes [affective]). This taxonomy of learning behaviors may be thought of as "the goals of the learning process." That is, after a learning episode, the learner should have acquired a new skill, knowledge, and/or attitude.

While the committee produced an elaborate compilation for the cognitive and affective domains, they omitted the psychomotor domain. Their explanation for this oversight was that they have little experience in teaching manual skills within the college level. However, there have been at least three psychomotor models. Their compilation divides the three domains into subdivisions, starting from the simplest cognitive process or behavior to the most complex. The divisions outlined are not absolutes and there are other systems or hierarchies that have been devised, such as the Structure of Observed Learning Outcome (SOLO). However, Bloom's taxonomy is easily understood and is probably the most widely applied one in use today.

Cognitive Domain

The cognitive domain involves knowledge and the development of intellectual skills (Bloom, 1956). This includes the recall or recognition of specific facts, procedural patterns, and concepts that serve in the development of intellectual abilities and skills. There are six major categories of cognitive and processes, starting from the simplest to the most complex (see the table below for an in-depth coverage of each category):
- Knowledge
- Comprehension
- Application
- Analysis
- Synthesis
- Evaluation

The categories can be thought of as degrees of difficulties. That is, the first ones must normally be mastered before the next one can take place.

Blooms Revised Taxonomy

Loran Anderson, a former student of Bloom, and David Krathwohl revisited the cognitive domain in the mid-nineties and made some changes, with perhaps the three most prominent ones being (Anderson, Krathwohl, Airasian, Cruikshank, Mayer, Pintrich, Raths, Wittrock, 2000):

o changing the names in the six categories from noun to verb forms o

rearranging them as shown in the chart below o creating a processes and

levels of knowledge matrix

b) **Describe the importance of cognitive domain in the development of achievement test.**

importance of cognitive domain in the development of achievement

Following are the steps for constructing an achievement test.

1. **Planning the test**
a. Designing the test
b. Preparation of blue print

2. **Preparing Preliminary draft**
a. Item writing
b. Item editing
c. Pre try out

3. **The try out**

4. **Item analysis**

5. **Preparing the final draft**

6. **Establishment of**
a. Reliability
b. Validity
c. Norms
   Description of all above steps is described below briefly.
1. **Planning the test**: Planning of achievement test will be carried out with the help of two steps.

a. **Designing of test**
   Designing is most important step in the building test. Designer should be careful about planning and making test successful. He should keep this in mind that how and what types of questions are used. He has to decide weightage for different objectives and units in the course. Following points helps designer to design test in systematic way.
i. Identification of objectives
ii. Measurement of content
iii. Allotment of time
iv. Allotment of marks
v. Include multiple choice question

b.   Preparation of blueprint

After designing preparation of blueprint is the last stage of the planning of test. Here test constructor put various type of question in blueprint and allots them marks depending on the time. The tester writes down his decisions in the form of a blueprint.

Following figure is the example of blueprint.

	Blue Print of Question Paper: Standard X : English															
	Type of Questions			Mental Process										Score	Discourse	Unit
Area	Obj	SA	Essay	1	2	3	4	5	6	7	8	9	10			
Reading	6	5			✓	✓		✓		✓	✓	✓		10	Narrative Essay	2,5
Writing			3		✓	✓	✓			✓		✓		12	Description, Narration	3,4,5,2,1
Literary Skills		8						✓		✓				8	Poems	2,6
Language Elements	4	3					✓		✓	✓				20	Dialogue, Report Description	1,2,4,7

With the completion of blueprint the remaining decisions of the design will become the basis for writing the items.

2.   Preliminary draft

Preliminary draft is prepared by following three stages.

a.   **Item writing:** It is important step in preliminary draft and here blue print is used as a guide writing this draft. Test conductor should have following points in mind when he writes items for preliminary draft.

I.   Each item contains single idea.
ii.   Questions should be clear.
iii.   Simple and easy to understand.
iv.   Double barreled question should be avoided.
v.   Arrangement of items should be from simple to complex.
vi.   Subjective question should also be avoided.

b.   **Item editing**: Then the item should be edited and reviewed by language teacher and also by experts of measurement. Language teacher will check the errors in language and defect in words. By removing defects it is submitting to experts. Experts of measurement measure the level of achievement of items.

c.   **Pre try out:** The draft is modified and remove the shortcomings of preliminary draft. In this stage constructer is confident for the usability of test.

3. **The try out**

   To ensure the proper operating of test items and to remove shortcoming, it is essential to have try out of items. This helps in predict how students will work in actual practices.

   At the try out stage, the test should be so timed that nearly 90% of the sample should be able to attempt all the items of the test. The test sheets along with answer sheets are collected and answer sheets are used for scoring with the help of scoring keys. Keep following points in mind while testing in sample.

   I.   Proper sitting arrangements.
   ii.   Time of administrating the test.
   iii.   Total time required for test.
   iv.   Proper motivation with the pupil.

4. **Item analysis**

   In this stage test constructer examining the responses of respondents in the sample to each test item. It can also be define as it is a statistical procedure by which the appropriate items are selected for the final draft and poor items are rejected. Item analysis is an analysis of responses made to 'teacher made tests' by the pupil in the class.

5. **Final draft**

   In the final draft questions should be arranged in such way that easy, average and difficult and again starts from easy and so on. By doing this student at least go through whole test because he know that some of remaining questions are easy. When the test is arranged properly with the help of principles, a clear identity of test is appeared.

6. **Establishment of Reliability, Validity and Norms**

   With the selection of good items final draft has been prepared and the final step is the establishment of reliability, validity and norms. These test items administered to a larger sample.

   a. **Reliability**: The reliability of a test measures whatever it measure. There are many methods for computing the reliability of a test and the most appropriate method for computing reliability of achievement test is Split-half method.

   b. **Validity**: It is very important aspect of test and it can be determined as the degree which is capable of measuring achievements and it is design to do so. Valid test is highly reliable. Validity is measured by four methods.

   I.   Face validity
   ii.   Content validity
   iii.   Concurrent validity
   iv.   Predictive validity

   c. **Norms**: It is average score of sample population. These are the level obtained by a particular group of persons on a test. There are many types of norms like age norms, grade norms, percentile norms and standard scores.

Q.3 Compare and contrast the characteristics of Criterion and Norm referenced Tests. Also highlight its utilization in teaching learning process.

Criterion-referenced tests vs. norm-referenced tests

Criterion-referenced vs. norm-referenced

To understand what happened, we need to understand the difference between criterion-referenced tests and norm-referenced tests.

The first thing to understand is that even an assessment expert couldn't tell the difference between criterion-referenced test and a norm-referenced test just by looking at them. The difference is actually in the scores—and some tests can provide both criterion-referenced results and norm-referenced results!

Criterion-referenced tests

Criterion-referenced tests compare a person's knowledge or skills against a predetermined standard, learning goal, performance level, or other criterion. With criterion-referenced tests, each person's performance is compared directly to the standard, without considering how other students perform on the test. Criterion-referenced tests often use "cut scores" to place students into categories such as "basic," "proficient," and

Criterion-referenced tests compare a student's knowledge and skills against a predetermined standard, cut score, or other criterion.

In criterion-referenced tests, the performance of other students does not affect a student's score.

"advanced."

If you've ever been to a carnival or amusement park, think about the signs that read "You must be this tall to ride this ride!" with an arrow pointing to a specific line on a height chart. The line indicated by the arrow functions as the criterion; the ride operator compares each person's height against it before allowing them to get on the ride.

Note that it doesn't matter how many other people are in line or how tall or short they are; whether or not you're allowed to get on the ride is determined solely by your height. Even if you're the tallest person in line, if the top of your head doesn't reach the line on the height chart, you can't ride.

Criterion-referenced assessments work similarly: An individual's score, and how that score is categorized, is not affected by the performance of other students. In the charts below, you can see the student's score and performance category ("below proficient") do not change, regardless of whether they are a top-performing student, in the middle, or a low-performing student.
Norm-referenced tests

Norm-referenced measures compare a person's knowledge or skills to the knowledge or skills of the norm group. The composition of the norm group depends on the assessment. For student assessments, the norm group is often a nationally representative sample of several thousand students in the same grade (and sometimes, at the same point in the school year). Norm groups may also be further narrowed by age, English Language Learner (ELL) status, socioeconomic level, race/ethnicity, or many other characteristics.

One norm-referenced measure that many families are familiar with is the baby weight growth charts in the pediatrician's office, which show which percentile a child's weight falls in. A child in the 50th percentile has an average weight; a child in the 75th percentile weighs more than 75% of the babies in the norm group and the same as or less than the heaviest 25% of babies in the norm group; and a child in the 25th percentile weighs more than 25% of the babies in the norm group and the same as or less than 75% of them. It's important to note that these norm-referenced measures do not say whether a baby's birth weight is "healthy" or "unhealthy," only how it compares with the norm group.

For example, a baby who weighed 2,600 grams at birth would be in the 7th percentile, weighing the same as or less than 93% of the babies in the norm group. However, despite the very low percentile, 2,600 grams is classified as a normal or healthy weight for babies born in the United States—a birth weight of 2,500 grams is the cut-off, or criterion, for a child to be considered low weight or at risk. (For the curious, 2,600 grams is about 5 pounds and 12 ounces.) Thus, knowing a baby's percentile rank for weight can tell you how they compare with their peers, but not if the baby's weight is "healthy" or "unhealthy."

Norm-referenced assessments work similarly: An individual student's percentile rank describes their performance in comparison to the performance of students in the norm group, but does not indicate whether or not they met or exceed a specific standard or criterion.

In the charts below, you can see that, while the student's score doesn't change, their percentile rank does change depending on how well the students in the norm group performed. When the individual is a top-performing student, they have a high percentile rank; when they are a low-performing student, they have a low percentile rank. What we can't tell from these charts is whether or not the student should be categorized as proficient or below proficient.

Utilization in teaching learning process:

Many college entrance exams and nationally used school tests use norm-referenced tests. The SAT, Graduate Record Examination (GRE), and Wechsler Intelligence Scale for Children (WISC) compare individual student performance to the performance of a normative sample. Test takers cannot "fail" a norm-referenced test, as each test taker receives a score that compares the individual to others that have taken the test, usually given by a percentile. This is useful when there is a wide range of acceptable scores, and the goal is to find out who performs better.

IQ tests are norm-referenced tests, because their goal is to see which test taker is more intelligent than the other test takers. The median IQ is set to 100, and all test takers are ranked up or down in comparison to that level.

Theater auditions and job interviews are norm-referenced tests, because their goal is to identify the best candidate compared to the other candidates, not to determine how many of the candidates meet a fixed list of standards.

Q.4 Elaborate the different techniques for the measurement of aptitude of the learners by providing examples. Why aptitude measurement is important for the teachers in teaching learning process?

Profile Your Aptitudes

WHAT ARE YOUR APTITUDES? Everyone is born with certain natural abilities or aptitudes that can be developed to make you special. Aptitudes are natural talents or inclinations for certain types of activities.

For example, you may have a very high aptitude for artistic projects but little aptitude for numerical activities. In this case you probably will enjoy your art class more than your math class and find yourself drawing pictures instead of planning a budget.

Some examples of aptitudes include the following:

* **Verbal or non-verbal communications** - written or spoken words or actions that communicate ideas, emotions, or information. Giving a speech in class, writing a letter, acting in a school play, and hugging a friend are all methods of communication.

* **Verbal comprehension -** understanding the meanings of ideas or emotions in verbal or nonverbal communications. Examples include listening and reacting to a friend's problem, discussing current events with the class, and trying to sell your neighbor on your abilities to mow their lawn.

* **Logical -** applying reason or logic to problems. Solving a mystery, measuring the fabric available and deciding if there is enough to make a skirt, and determining how to design a spread sheet on your computer are all examples of logical abilities.

* **Artistic** - visual creativity, using artistic talents, musical abilities, or dramatic talents. Examples include being able to draw, write poetry, arrange flowers, play a musical instrument, take pleasing photographs, sing, or design an outfit.

* **Mechanical** - understanding relationships between parts of machines, how things are made, and making things work. Fixing the motor of an automobile, putting a radio back together, un-jamming a sewing machine, and even driving a car are examples of mechanical abilities.

* **Numerical** - working with numbers. Examples include solving any type of math problem, determining how many miles a car travels on a gallon of gas, doubling a recipe to feed eight instead of four people, and making a budget for the use of your allowance.

* Clerical - arranging and recording number and letter combinations. Examples include alphabetizing or putting items in a special order, filing reports, typing information, keeping records. *** Spatial -** understanding how parts of things fit together, multidimensional understanding. Being able to put together a jigsaw puzzle, rearranging furniture attractively in a room, and putting together a model airplane all use spatial understanding.

* Physical - bodily strength and coordination, manual dexterity. Lifting weights, enjoying aerobic exercise, moving furniture, and building furniture make use of physical skills.

* Organizational - planning, implementing and evaluating actions for yourself or others. Planning a party, organizing your vacation trip, and conducting a meeting all make use of organizational abilities.

* Intellectual - original thinking, seeking knowledge, thinking ahead, and developing concepts. Reading books, studying for tests, creating a new way to make a cake, and developing a business plan illustrate these abilities.

Entrepreneurs see their aptitudes as a special opportunity to develop a business. Often their aptitudes have lead them to experiences where these aptitudes helped them excel. Through the many experiences of our lives, people learn about their own natural abilities and seek additional experiences where they believe they can be successful because they can do it.

Personal Profile

1. Each student should now write a paragraph each about four experiences in their lives that they particularly enjoyed. After completing a short paragraph about each experience, ask them to reread their descriptions of each experience and decide which of the above list of aptitudes they showed in each experience.

2. Then ask them to think about the 5 aptitudes from the given list that they see in themselves. List the aptitudes on the left side of the paper, and opposite of it on the right, list the interests and experiences they have had that support their choice of aptitudes for their personal profile.

3. Finally, ask the students to think of the types of businesses they could start that would make use of their personal aptitudes. Explain why they believe this is true. Discuss what they might do in their future to use these aptitudes to improve their skills and prepare to be able to be an entrepreneur in their area of interest.

Aptitude tests explained

An aptitude test is a structured assessment that aims to evaluate job candidates' or even existing employees' talent and/or skill in completing certain tasks without any prior knowledge or training. There are lots of different types of aptitude test which can assess a series of core skills such as technical knowledge and comprehension.

An increasing number of employers are now using aptitude tests as a way to assess the specific capabilities of potential employees. They are often used as a method of screening to identify high caliber candidates. The content of an aptitude test, particularly in an industry such as innovation, needs careful thought and consideration. They are used to assess candidates in terms of their skills, knowledge, ability and personality. Aptitude tests are an accurate way of predicting the potential contribution of a candidate and the likelihood of them achieving success within the organization. It can sometimes be difficult from an interview alone to assess the candidate's skills, so an aptitude test provides the employer with a greater insight into their capabilities.

Typical aptitude tests include:

- **Numerical Reasoning**
  This type of test gauges how well the candidate can interpret graphs, data, statistics or charts
- **Verbal Reasoning**
  Verbal reasoning tests are designed to identify whether or not the candidate understands written information and passage of text.
- **Diagrammatic reasoning**
  this type of aptitude test evaluate candidates' ability to follow diagrams, identify patterns, synthesize data, and solve problems logically.
- **Logical Reasoning**
  Logical reasoning gauges candidates' ability to follow something through to a conclusion when given basic information.
- **Inductive reasoning**
  this test measures a candidate's ability to solve problems when given unfamiliar information.

During the recruitment process, it is important for employers to assess the skills and abilities of candidates because these often reflect their capacity to do the job. Assessing important skills through an aptitude test can help highlight exemplary candidates who can really make a valuable contribution to the future of a business.

The best innovations are derived from highly motivated and skilled employees. Finding a high caliber candidate during the recruitment process without the use of an aptitude test could prove challenging to say the least. An aptitude test can help a business differentiate the good from the great in an industry where employees need to generate ideas that are the best of the best.

Q.5 Explain the advantages and disadvantages of objective type test items. Also highlight the importance and significance of the objective type tests.

Advantages and disadvantages of the most commonly used test questions and the test banks that now frequently provide them.

Multiple-choice questions

Advantages

o Quick and easy to score, by hand or electronically

o Can be written so that they test a wide range of higher-order thinking skills

o Can cover lots of content areas on a single exam and still be answered in a class period

Disadvantages

o Often test literacy skills: "if the student reads the question carefully, the answer is easy to recognize even if the student knows little about the subject" (p. 194)

o Provide unprepared students the opportunity to guess, and with guesses that are right, they get credit for things they don't know

o Expose students to misinformation that can influence subsequent thinking about the content

o Take time and skill to construct (especially good questions)
True-false questions

Advantages

o Quick and easy to score

Disadvantages

o Considered to be "one of the most unreliable forms of assessment" (p. 195)

o Often written so that most of the statement is true save one small, often trivial bit of information that then makes the whole statement untrue

o Encourage guessing, and reward for correct guesses
Short-answer questions

Advantages

o Quick and easy to grade

o Quick and easy to write

Disadvantages

o Encourage students to memorize terms and details, so that their understanding of the content remains superficial
Essay questions

Advantages

o Offer students an opportunity to demonstrate knowledge, skills, and abilities in a variety of ways

o Can be used to develop student writing skills, particularly the ability to formulate arguments supported with reasoning and evidence

Disadvantages

o Require extensive time to grade

o Encourage use of subjective criteria when assessing answers

o If used in class, necessitate quick composition without time for planning or revision, which can result in poor-quality writing
Questions provided by test banks

Advantages

o Save instructors the time and energy involved in writing test questions

o Use the terms and methods that are used in the book

Disadvantages

o Rarely involve analysis, synthesis, application, or evaluation (cross-discipline research documents that approximately 85 percent of the questions in test banks test recall)

o Limit the scope of the exam to text content; if used extensively, may lead students to conclude that the material covered in class is unimportant and irrelevant.

We tend to think that these are the only test question options, but there are some interesting variations. The article that promoted this review proposes one: Start with a question, and revise it until it can be answered with one word or a short phrase. Do not list any answer options for that single question, but attach to the exam an alphabetized list of answers. Students select answers from that list. Some of the answers provided may be used more than once, some may not be used, and there are more answers listed than questions. It's a ratcheted-up version of matching. The approach makes the test more challenging and decreases the chance of getting an answer correct by guessing.

Objective-type tests have two characteristics viz.:

1. They are pin-pointed, definite and so clear that a single, definite answer is expected.

2. They ensure perfect objectivity in scoring. The scoring will not vary from examiner to examiner.

Merits of Objective Type Test:

1. Objective type test gives scope for wider sampling of the content.

2. It can be scored objectively and easily. The scoring will not vary from time to time or from examiner to examiner.

3. This test reduces (a) the role of luck and (b) cramming of expected questions. As a result, there is greater reliability and better content validity.

4. This type of question has greater motivational value.

5. It possesses economy of time, for it takes less time to answer than an essay test. Comparatively, many test items can be presented to students. It also saves a lot of time of the scorer.

6. It eliminates extraneous (irrelevant) factors such as speed of writing, fluency of expression, literary style, good handwriting, neatness, etc.

7. It measures the higher mental processes of understanding, application, analysis, prediction and interpretation.

8. It permits stencil, machine or clerical scoring. Thus scoring is very easy.

9. Linguistic ability is not required.