NEHRU COLLEGE OF EDUCATION

Pillaiyarkuppam, Puducherry – 605502

Phone: 0413-2661185 / 86, Web: ncepdy.org, Email:ncepdy@gmail.com

B.Ed., - FIRST YEAR – I SEMESTER

ASSESSMENT FOR LEARNING - I

Dr. A. Murugan, Ph.D

Principal, Nehru College of Education

ASSESSMENT FOR LEARNING - I

UNIT – 2

BASICS OF EDUCATIONAL TESTING, MEASUREMENT, ASSESSMENT AND EVALUATION

UNIT – 2

BASICS OF EDUCATIONAL TESTING, MEASUREMENT, ASSESSMENT AND EVALUATION

2.1 CRYSTALLIZING THE CONCEPT OF TESTING, MEASUREMENT, ASSESSMENT AND EVALUATION

What is Testing, Measurement, Assessment, and Evaluation in Education has always been asked question by a considerable portion of my students, despite having completed B.Ed. course, always perplexing students in misconceptions state regarding the fundamental distinctions among terms like test, measurement, assessment, and evaluation within the context of education

Test, measurement, assessment, and evaluation—they're not just words; they represent distinct concepts, each with its unique significance. What exactly do these terms entail, and why is it crucial to discern?

Firstly, let's unpack the notion of **Testing**

In the contexts of education, Testing refers to the use of tools, techniques or methods to measure students' ability or their knowledge to complete a specific task successfully. It is a form of assessment that evaluates academic readiness, learning progress, skill acquisition, or educational needs.

The obvious point of testing is to assess what students have learned or to determine student strengths and weaknesses or to judge a student based on merit.

Moving on to **"Measurement,"** we're entering the quantification and precision. In educational contexts, measurement refers to the systematic process of assigning numerical values or descriptors to observed traits or behaviors.

Now, let's pivot to **"Assessment.** Unlike a mere test or measurement, assessment encompasses a broader spectrum of activities aimed at gathering information about student learning and progress. It involves not only the administration of tests but also the interpretation of results, the provision of feedback, and the adjustment of instructional strategies to meet students' needs effectively.

Assessment, in essence, is the ongoing dialogue between educators and learners, facilitating continuous improvement and growth.

Finally, we arrive at **"Evaluation."** While assessment focuses on the process of gathering information, evaluation zooms out to scrutinize the outcomes and effectiveness of educational endeavours. It involves making judgments about the quality, value, and impact of instructional practices, programs, or policies.

Evaluation serves as the critical lens through which educators assess the overall efficacy of their teaching methods and make informed decisions to enhance learning outcomes.

2.2 EVALUATION - DEFINITION

To measure means to observe or determine the magnitude of variate; evaluation means assessment or appraisal - **Encyclopedia of Education Research:**

Evaluation is a systematic process of collecting, analyzing and interpreting information to determine the extent to which pupils are achieving instructional objectives - **Gronlund and Linn**

2.3 STEPS INVOLVED IN EVALUATION PROCESS

Following are the steps involved in the process of evaluation:

(i) Identifying and Defining General Objectives
(ii) Identifying and Defining Specific Objectives
(iii) Selecting Teaching Points
(iv) Planning Suitable Learning Activities
(v) Implementing / Evaluating
(vi) Using the Results as Feedback/ Reporting

(i) Identifying and Defining General Objectives: In the evaluation process first step is to determine what to evaluation, i.e., to set down educational objectives. What kind of abilities and skills should be developed when a pupil studies, what type of understanding should be developed in the pupil who learns? Unless the teacher identifies and states the objectives, these questions will remain unanswered.

The process of identifying and defining educational objectives is a complex one; there is no simple or single procedure which suits all teachers. Some prefer to begin with the course content, some with general aims, and some with lists of objectives suggested by curriculum experts in the area.

While stating the objectives, therefore, we can successfully focus our attention on the product i.e., the pupil's behavior, at the end of a course of study and state it in terms of his knowledge, understanding, skill, application, attitudes, interests, appreciation, etc.

(ii) Identifying and Defining Specific Objectives:

It has been said that learning is the modification of behavior in a desirable direction. The teacher is more concerned with a student's learning than with anything else. Changes in behavior are an indication of learning. These changes, arising out of classroom instruction, are known as the learning outcome.

What type of learning outcome is expected from a student after he has undergone the teaching-learning process is the first and foremost concern of the teacher. This is possible only when the teacher identifies and defines the objectives in terms of behavioral changes, i.e., learning outcomes.

These specific objectives will provide direction to teaching-learning process. Not only that, it will also be useful in planning and organizing the learning activities, and in planning and organizing evaluation procedures too.

Thus, specific objectives determine two things; one, the various types of learning situations to be provided by the class teacher 10 his pupils and second, the method to be employed to evaluate both—the objectives and the learning experiences.

(iii) Selecting Teaching Points:

The next step in the process of evaluation is to select teaching points through which the objectives can be realized. Once the objectives are set up, the next step is to decide the content (curriculum, syllabus, and course) to help in the realization of objectives.

For the teachers, the objectives and courses of school subjects are ready at hand. His job is to analyze the content of the subject matter into teaching points and to find out what specific objectives can be adequately realized through the introduction of those teaching points.

(iv) Planning Suitable Learning Activities:

In the fourth step, the teacher will have to plan the learning activities to be provided to the pupils and, at the same time, bear two things in mind—the objectives as well as teaching points. The process then becomes three

dimensional, the three co-ordinates being objectives, teaching points and learning activities. The teacher gets the objectives and content readymade.

He is completely free to select the type of learning activities. He may employ the analytico-synthetic method; he may utilize the inducto-deductive reasoning; he may employ the experimental method or a demonstration method; or he may put a pupil in the position of a discoverer; he may employ the lecture method; or he may ask the pupils to divide into groups and to do a sort of group work followed by a general discussion; and so on. One thing he has to remember is that he should select only such activities as will make it possible for him to realize his objectives.

(v) Implementing/Evaluating:

In the fifth step, the teacher observes and measures the changes in the behavior of his pupils through testing. This step adds one more dimension to the evaluation process. While testing, he will keep in mind three things-objectives, teaching points and learning activities; but his focus will be on the attainment of objectives. This he cannot do without enlisting the teaching points and planning learning activities of his pupils.

Here the teacher will construct a test by making the maximum use of the teaching points already introduced in the class and the learning experiences already acquired by his pupils. He may plan for an oral lest or a written test; he may administer an essay type test or an objective type of lest; or he may arrange a practical test.

(vi) Using the Results as Feedback/ Reporting

The last, but not the least, important step in the evaluation process is the use of results as feedback. If the teacher, after testing his pupils, finds that the objectives have not been realized to a great extent, he will use the results in reconsidering the objectives and in organizing the learning activities.

He will retrace his steps to find out the drawbacks in the objectives or in the learning activities he has provided for his students. This is known as feedback. Whatever results the teacher gets after testing his pupils should be utilized for the betterment of the students.

2.4 CHARACTERISTICS OF EVALUATION:

1. Evaluation implies a systematic process which omits the casual uncontrolled observation of pupils.
2. Evaluation is a continuous process. In an ideal situation, the teaching-learning process on the one hand and the evaluation procedure on the other hand, go together. It is certainly a wrong belief that the evaluation procedure follows the teaching-learning process.
3. Evaluation emphasises the broad personality changes and major objectives of an educational programme. Therefore, it includes not only subject-matter achievements but also attitudes, interests and ideals, ways of thinking, work habits and personal and social adaptability.
4. Evaluation always assumes that educational objectives have previously been identified and defined. This is the reason why teachers are expected not to lose sight of educational objectives while planning and carrying out the teaching-learning process either in the classroom or outside it.
5. A comprehensive programme of evaluation involves the use of many procedures (for example, analytico-synthetic, heuristic, experimental, lecture, etc.); a great variety of tests (for example, essay type, objective type, etc.); and other necessary techniques (for example, socio-metric, controlled-observation techniques, etc.).
6. Learning is more important than teaching. Teaching has no value if it does not result in learning on the part of the pupils.
7. Objectives and accordingly learning experiences should be so relevant that ultimately they should direct the pupils towards the accomplishment of educational goals.
8. To assess the students and their complete development brought about through education is evaluation.
9. Evaluation is the determination of the congruence between the performance and objectives.

2.5 NEED AND IMPORTANCE OF EVALUATION IN EDUCATION:

Now a days, education has multifold programmes and activities to inculcate in students a sense of common values, integrated approach, group feelings, and community interrelationship leading to national integration and knowledge to adjust in different situations.

- **Informing Decision-Making:** Evaluation provides data-driven insights that inform decision-making at various levels, including classroom instruction, curriculum development, resource allocation, and policy formulation.
- **Improving Educational Quality:** By identifying areas for improvement and best practices, evaluation helps educators and policymakers enhance the quality and effectiveness of educational programs and practices.
- **Ensuring Accountability:** Evaluation holds educators, schools, and educational systems accountable for achieving desired outcomes and meeting established standards of performance.
- **Facilitating Continuous Improvement:** Through ongoing assessment and feedback, evaluation supports continuous improvement efforts, allowing educators and educational systems to adapt and refine their practices over time
- **Enhancing Equity and Access:** Evaluation helps identify disparities in educational opportunities and outcomes, enabling policymakers to address inequities and ensure equal access to high-quality education for all students.
- **Knowing pupils in details:** It helps a teacher to know his pupils in details. Today, education is child-centered. So, child's abilities, interest, aptitude, attitude etc., are to be properly studied so as to arrange instruction accordingly.
- It helps the teacher to determine, evaluate and refine his instructional techniques, in setting, refining and clarifying the objectives.
- It helps him to know the entry behaviour of the students.
- It helps an administrator in educational planning and in educational decisions on selections, classification and placement.
- It helps to design better educational programmes.
- The parents are eager to know about the educational progress of their children and evaluation alone can assess the pupils' progress from time to time.

- A sound choice of objectives depends on an accurate information regarding pupil's abilities, interest, attitude and personality traits and such information is obtained through evaluation.
- Evaluation helps us to know whether the instructional objectives have been achieved or not. As such evaluation helps planning of better strategies for education.
- A sound programme of evaluation clarifies the aims of education and it helps us to know whether aims and objectives are attainable or not. As such, it helps in reformulation of aims and objectives.
- Evaluation studies the 'total child' and thus helps us to undertake special instructional programmes like enrichment programme, for the bright and remedial programmes for the backward.
- It helps a student in encouraging good study habits, in increasing motivation and in developing abilities and skills, in knowing the results of progress and in getting appropriate feedback.
- It helps us to undertake appropriate guidance services.

2.6 CHALLENGES IN EVALUATION:

- **Validity and Reliability:** Ensuring that evaluation measures accurately assess what they are intended to measure and produce consistent results over time.
- **Bias and Fairness:** Addressing potential biases in assessment tools, processes, and interpretations to ensure fairness and equity for all students.
- **Resource Constraints:** Limited resources, time, and expertise can pose challenges to conducting comprehensive and rigorous evaluations, especially in resource-constrained settings.
- **Complexity of Educational Systems:** Educational systems are multifaceted, making it challenging to capture the full range of factors that influence student learning outcomes and program effectiveness.
- **Resistance to Change:** Resistance from stakeholders, including educators, administrators, and policymakers, can hinder the implementation of evaluation findings and recommendations

2.7 PURPOSES AND FUNCTIONS OF EVALUATION:

Evaluation plays a vital role in teaching learning experiences. It is an integral part of the instructional programmes. It provides information's on the basis of which many educational decisions are taken. We are to stick to the basic function of evaluation which is required to be practiced for pupil and his learning processes.

Evaluation has the following functions:

1. Placement Functions:

- Evaluation helps to study the entry behaviour of the children in all respects.
- That helps to undertake special instructional programmes.
- To provide for individualisation of instruction.
- It also helps to select pupils for higher studies, for different vocations and specialised courses.

2. Instructional Functions:

- A planned evaluation helps a teacher in deciding and developing the ways, methods, techniques of teaching.
- Helps to formulate and reformulate suitable and realistic objectives of instruction.
- Which helps to improve instruction and to plan appropriate and adequate techniques of instruction.
- And also helps in the improvement of curriculum.
- To assess different educational practices.
- Ascertains how far could learning objectives be achieved.
- To improve instructional procedures and quality of teachers.
- To plan appropriate and adequate learning strategies.

3. Diagnostic Functions:

- Evaluation has to diagnose the weak points in the school programme as well as weakness of the students.
- To suggest relevant remedial programmes.
- The aptitude, interest and intelligence are also to be recognised in each individual child so that he may be energised towards a right direction.
- To adopt instruction to the different needs of the pupils.

- To evaluate the progress of these weak students in terms of their capacity, ability and goal.

4. Predictive functions:

- To discover potential abilities and aptitudes among the learners.
- Thus to predict the future success of the children.
- And also helps the child in selecting the right electives.

5. Administrative Functions:

- To adopt better educational policy and decision making.
- Helps to classify pupils in different convenient groups.
- To promote students to next higher class,
- To appraise the supervisory practices.
- To have appropriate placement.
- To draw comparative statement on the performance of different children.
- To have sound planning.
- Helps to test the efficiency of teachers in providing suitable learning experiences.
- To mobilise public opinion and to improve public relations.
- Helps in developing a comprehensive criterion tests.

6. Guidance Functions:

- Assists a person in making decisions about courses and careers.
- Enables a learner to know his pace of learning and lapses in his learning.
- Helps a teacher to know the children in details and to provide necessary educational, vocational and personal guidance.

7. Motivation Functions:

- To motivate, to direct, to inspire and to involve the students in learning.
- To reward their learning and thus to motivate them towards study.

8. Development Functions:

- Gives reinforcement and feedback to teacher, students and the teaching learning processes.
- Assists in the modification and improvement of the teaching strategies and learning experiences.
- Helps in the achievement of educational objectives and goals.

9. Research Functions:

- Helps to provide data for research generalisation.
- Evaluation clears the doubts for further studies and researches.
- Helps to promote action research in education.

10. Communication Functions:

- To communicate the results of progress to the students.
- To intimate the results of progress to parents.
- To circulate the results of progress to other schools.

2.8 EVALUATION – TYPES

Evaluation can be classified into different categories in many ways.

According to function

1. Placement Evaluation
2. Formative evaluation
3. Diagnostic evaluation
4. Summative evaluation

According to Approaches

1. Formative evaluation
2. Summative evaluation

According to nature of

1. Norm referenced
2. Criterion referenced

2.9 PLACEMENT EVALUATION:

Placement evaluation is designed to place the right person in the right place. It ensures the entry performance of the pupil. The future success of the instructional process depends on the success of placement evaluation. Placement evaluation aims at evaluating the pupil's entry behaviour in a sequence of instruction. In other words the main goal of such evaluation is to determine the level or position of the child in the instructional sequence.

We have a planned scheme of instruction for classroom which is supposed to bring a change in pupil's behaviour in an orderly manner. Then we prepare or place the students for planned instruction for their better prospects.

When a pupil is to undertake a new instruction, it is essential to know the answer of the following questions:

- Does the pupil possess required knowledge and skills for the instruction?
- Whether the pupil has already mastered some of the instructional objectives or not?
- Whether the mode of instruction is suitable to pupil's interests, work habits and personal characteristics?

We get the answer to all the probable questions by using a variety of tests, self-report inventories, observational techniques, case study, attitude test and achievement tests. Sometimes past experiences, which inspire for present learning also lead to the further placement in a better position or admission. This type of evaluation is helpful for admission of pupils into a new course of instruction.

Examples:

Aptitude test, Self-reporting inventories, Observational techniques, Medical entrance exam and Engineering or Agriculture entrance exam.

2.10 FORMATIVE EVALUATION:

Formative evaluation is used to monitor the learning progress of students during the period of instruction. Its main objective is to provide continuous feedback to both teacher and student concerning learning successes and failures while instruction is in process. Feedback to students provides reinforcement of successful learning and identifies the specific learning errors that need correction. Feedback to teacher provides information for modifying instruction and for prescribing group and individual remedial work.

Formative evaluation helps a teacher to ascertain the pupil-progress from time to time. At the end of a topic or unit or segment or a chapter the teacher can evaluate the learning outcomes basing on which he can modify his methods, techniques and devices of teaching to provide better learning experiences.

The teacher can even modify the instructional objectives, if necessary. In other words, formative evaluation provides feedback to the teacher. The teacher can know which aspects of the learning task were mastered and which aspects were poorly or not at all mastered by pupils. Formative evaluation helps the teacher to assess the relevance and appropriateness of the learning experiences provided and to assess instantly how far the goals are being fulfilled.

Thus, it aims at improvement of instruction. Formative evaluation also provides feedback to pupils. The pupil knows his learning progress from time to time. Thus, formative evaluation motivates the pupils for better learning. As such, it helps the teacher to take appropriate remedial measures. **"The idea of generating information to be used for revising or improving educational practices is the core concept of formative evaluation."**

It is concerned with the process of development of learning. In the sense, evaluation is concerned not only with the appraisal of the achievement but also with its improvement. Education is a continuous process. Therefore, evaluation and development must go hand in hand. The evaluation has to take place in every possible situation or activity and throughout the period of formal education of a pupil.

Cronback is the first educationist, who gave the best argument for formative evaluation. According to him, the greatest service evaluation can perform is to identify aspects of the course where education is desirable. Thus, this type of evaluation is an essential tool to provide feedback to the learners for improvement of their self-learning and to the teachers for improvement of their methodologies of teaching, nature of instructional materials, etc.

It is a positive evaluation because of its attempt to create desirable learning goals and tools for achieving such goals. Formative evaluation is generally concerned with the internal agent of evaluation, like participation of the learner in the learning process.

Functions of formation evaluation are:

1. **Diagnosing** -Diagnosing is concerned with determining the most appropriate method or instructional materials conducive to learning.
2. **Placement** - Placement is concerned with the finding out the position of an individual in the curriculum from which he has to start learning.
3. **Monitoring** - Monitoring is concerned with keeping track of the day-to-day progress of the learners and to point out changes necessary in the methods of teaching, instructional strategies, etc.

Characteristics of Formative Evaluation:

The characteristics of formative evaluation are as follows:

- It is an integral part of the learning process.

- It occurs, frequently, during the course of instruction.
- Its results are made immediately known to the learners.
- It may sometime take form of teacher observation only.
- It reinforces learning of the students.
- It pinpoints difficulties being faced by a weak learner.
- Its results cannot be used for grading or placement purposes.
- It helps in modification of instructional strategies including method of teaching, immediately.
- It motivates learners, as it provides them with knowledge of progress made by them.
- It sees role of evaluation as a process.
- It is generally a teacher-made test.
- It does not take much time to be constructed.

Examples:

Monthly tests, Class tests, Periodical assessment, Teacher's observation, etc.

2.11 DIAGNOSTIC EVALUATION:

It is concerned with identifying the learning difficulties or weakness of pupils during instruction. It tries to locate or discover the specific area of weakness of a pupil in a given course of instruction and also tries to provide remedial measure.

"Formative evaluation provides first-aid treatment for simple learning problems whereas diagnostic evaluation searches for the underlying causes of those problems that do not respond to first-aid treatment."- **N.E. Gronlund**

When the teacher finds that inspite of the use of various alternative methods, techniques and corrective prescriptions the child still faces learning difficulties, he takes recourse to a detailed diagnosis through specifically designed tests called 'diagnostic tests'. Diagnosis can be made by employing observational techniques, too. In case of necessity the services of psychological and medical specialists can be utilised for diagnosing serious learning handicaps.

2.12 SUMMATIVE EVALUATION:

Summative evaluation is done at the end of a course of instruction to know to what extent the objectives previously fixed have been accomplished. In other words, it is the evaluation of pupils' achievement at the end of a course. The main

objective of the summative evaluation is to assign grades to the pupils. It indicates the degree to which the students have mastered the course content. It helps to judge the appropriateness of instructional objectives. Summative evaluation is generally the work of standardised tests.

It tries to compare one course with another. The approaches of summative evaluation imply some sort of final comparison of one item or criteria against another. It has the danger of making negative effects. This evaluation may brand a student as a failed candidate, and thus causes frustration and setback in the learning process of the candidate, which is an example of the negative effect.

The traditional examinations are generally summative evaluation tools. Tests for formative evaluation are given at regular and frequent intervals during a course; whereas tests for summative evaluation are given at the end of a course or at the end of a fairly long period (say, a semester).

Functions of summative evaluation are:

1. **Crediting -** Crediting is concerned with collecting evidence that a learner has achieved some instructional goals in contents in respect to a defined curricular programme.
2. **Certifying -** Certifying is concerned with giving evidence that the learner is able to perform a job according to the previously determined standards.
3. **Promoting -** It is concerned with promoting pupils to next higher class.
4. **Selecting -** Selecting the pupils for different courses after completion of a particular course structure.

Characteristics of Summative Evaluation:

- It is terminal in nature as it comes at the end of a course of instruction (or a programme).
- It is judgemental in character in the sense that it judges the achievement of pupils.
- It views evaluation “as a product”, because its chief concern is to point out the levels of attainment.
- It cannot be based on teachers observations only.
- It does not pin-point difficulties faced by the learner.
- Its results can be used for placement or grading purposes.
- It reinforces learning of the students who has learnt an area.

- It may or may not motivate a learner. Sometimes, it may have negative effect.

Examples:

Traditional school and university examination,Teacher-made tests, Standardised tests, Practical and oral tests, and Rating scales, etc.

2.13 NORM-REFERENCED AND CRITERION-REFERENCED EVALUATION:

Two alternative approaches to educational testing that must be thoroughly understood are norm-referenced testing and criterion-referenced testing. Although there are similarities between these two approaches to testing, there are also fundamental differences between norm and criterion referenced testing.

There have been disputations about the relative virtues of norm and criterion-referenced measurements for a long time. However, a fundamental fact is recognised by most of concerned people that norm-referenced and criterion-referenced testing are complementary approaches.

2.14 CRITERION REFERENCE TEST (CRT)

Criterion-referenced test (CRT) is a type of test that measures a student's academic performance against a predetermined standard or criteria. The standard can be a percentage of items answered correctly or a state test benchmark. The student's score shows how close they are to meeting the standard.

CRTs compare a person's knowledge or skills against a predetermined standard, learning goal, performance level, or other criterion. With CRTs, each person's performance is compared directly to the standard without considering how other students perform on the test.

CRTs can be used to assess whether an individual has a particular set of competencies or skills. For example, a CRT might ask if a child can use a spoon to feed themselves or tie their shoelaces.

Some criterion-referenced tests are standardized, while others are not. An example of a high-stakes CRT is the IELTS.

CRTs can help teachers structure instruction and intervention for students who need it. However, they may not provide a comprehensive view of a student's abilities compared to their peers.

Definition:

A criterion-referenced test is used to ascertain an individual's status with respect to a defined achievement domain.

When the evaluation is concerned with the performance of the individual in terms of what he can do or the behaviour he can demonstrate, is termed as criterion- referenced evaluation. In this evaluation there is a reference to a criterion. But there is no reference to the performance of other individuals in the group. In it we refer an individual's performance to a predetermined criterion which is well defined.

Examples:

Raman got 93 marks in a test of Mathematics.

A typist types 60 words per minute.

Amit's score in a reading test is 70.

In the above examples there is no reference to the performance of other members of the group. Thus criterion-referenced evaluation determines an individual's status with reference to well defined criterion behaviour. It is an attempt to interpret test results in terms of clearly defined learning outcomes which serve as referents (criteria). Success of criterion-reference test lies in the delineation of all defined levels of achievement which are usually specified in terms of behaviourally stated instructional objectives.

The purpose of criterion-referenced evaluation/test is to assess the objectives. It is the objective based test. The objectives are assessed, in terms of behavioural changes among the students. Such type of test assesses the ability of the learner in relation to the criterion behaviour.

Glasar (1963) first used this term, 'Criterion-reference test' to describe the learner's achievement on a performance continuum.

Hively and Millman (1974) suggested a new term, 'domain-referenced test' and to them the word 'domain' has a wider connotation. A criterion referenced test can measure one or more assessment domain.

2.15 NORMS REFERENCED TESTS (NRT)

Norms-referenced tests are types of tests that compare an individual's performance to a predefined standard or norm. These tests are commonly used in educational settings to measure a student's abilities about their peers.

Norms-referred tests can provide valuable information about an individual's strengths and weaknesses, as well as their overall performance relative to others in the same age group or grade level.

Norm-referenced evaluation is the traditional class-based assignment of numerals to the attribute being measured. It means that the measurement act relates to some norm, group or a typical performance. It is an attempt to interpret the test results in terms of the performance of a certain group. This group is a norm group because it serves as a referent of norm for making judgements.

Test scores are neither interpreted in terms of an individual (self-referenced) nor in terms of a standard of performance or a pre-determined acceptable level of achievement called the criterion behaviour (criterion-referenced). The measurement is made in terms of a class or any other norm group.

Almost all our classroom tests, public examinations and standardised tests are norm-referenced as they are interpreted in terms of a particular class and judgements are formed with reference to the class.

Examples:

Raman stood first in Mathematics test in his class.

The typist who types 60 words per minute stands above 90 percent of the typists who appeared the interview.

Amit surpasses 65% of students of his class in reading test.

Norm referenced Test - Definition:

A norm-referenced test is used to ascertain an individual's status with respect to the performance of other individuals on that test.

In the above examples, the person's performance is compared to others of their group and the relative standing position of the person in his/her group is mentioned. We compare an individual's performance with similar information about the performance of others. That is why selection decisions always depend on norm- referenced judgements. A major requirement of norm-referenced judgements is that individuals being measured and individuals forming the group or norm, are alike. In norm-referenced tests very easy and very difficult items are discarded and items of medium difficulty are preferred because our aim is to study relative achievement.

2. 16 EVALUATION – USES

1. Ensures organization in delivering the best programs and services

It provides vital services to the community, how do you know your delivering the best programs and services to the organization. Evaluation provides quality control to support your efforts and ensures your time, money, and resources are well spent.

2. Makes programs, services, and systems more efficient and effective

The deeper your organization dives into evaluation approaches and adopts program evaluation into your culture, the more opportunities to leverage evaluation to make programs, services, and systems more efficient and effective. In other words, evaluation becomes a bridge that connects different individuals, departments, and partners around shared goals.

3. **Provides direction and informs strategic decisions**

Evaluation becomes a tool and resource to provide your organization with direction and guides strategic decisions. Sometimes it helps to make difficult decisions and evaluation provides your organization with vital Intel.

4. **Supports funding opportunities**

evaluation sets a base for allocation of funding to be invested in every walks of educational process.

6. **Produces results you can trust**

Evaluation produces results you can trust. This is powerful and vital information that confident about taking decisions and what needs to be improved.

7. Assessing the effectiveness of teaching:

Evaluation is concerned with assessing the effectiveness of teaching, teaching strategies, methods and techniques. It provides feedback to the teachers about their teaching and the learners about their learning.

8. Curriculum improvement:

The improvement in courses/curricula, texts and teaching materials is brought about with the help of evaluation.

9. Accountability to the Society:

Evaluation provides accountability to society in terms of the demands and requirements of the employment market.

10. Regular reporting to Parents:

Evaluation mainly manifests itself in a perceived need for regular reporting to parents.

In brief, evaluation is a very important requirement for the education system. It fulfills various purposes in systems of education like quality control in education, selection/entrance to a higher grade or tertiary level.

It also helps one to take decisions about success in specific future activities and provides guidance to further studies and occupation. Some of the educationists view evaluation virtually synonymous with that of learner appraisal, but evaluation has an expanded role. It plays an effective role in questioning or challenging the objectives.

NEHRU COLLEGE OF EDUCATION
PILLAIYARKUPPAM, PUDUCHERRY - 605 502
Love Knowledge Wisdom

ASSESSMENT FOR LEARNING

UNIT – 3

TEACHER MADE ACHIEVEMENT TESTS

3.1 TEACHER-MADE ACHIEVEMENT TESTS

3.2 TEACHER-MADE ACHIEVEMENT TESTS - EXAMPLES

3.3 PURPOSE OF TEACHER MADE ACHIEVEMENT TESTS

3.4 CONSTRUCTION OF TEACHER MADE ACHIEVEMENT TESTS

3.5 ADMINISTRATION OF TEACHER MADE ACHIEVEMENT TESTS

3.6 SCORINGAND RECORDING OF TEST RESULTS

3.7 OBJECTIVE TYPE TEST:

3.8 ESSAY TYPE TEST:

3.9 TYPES OF ESSAY TEST:

3.10 ADVANTAGES OF THE ESSAY TESTS:

3.11 LIMITATIONS OF ESSAY TESTS:

3.12 SUGGESTIONS FOR IMPROVING ESSAY TESTS:

3.13 CHARECTERISTICS/ CRITERIA OF A GOOD TEST

NEHRU COLLEGE OF EDUCATION ★ PILLAIYARKUPPAM, PUDUCHERRY - 605 502 ★
Love Knowledge Wisdom

ASSESSMENT FOR LEARNING

UNIT – 3

TEACHER MADE ACHIEVEMENT TESTS

3.1 TEACHER-MADE ACHIEVEMENT TESTS

Teacher-made tests are those that are constructed by teachers for use largely within their classrooms. Thus, teacher made achievement tests are those that are constructed by the teacher to assess learning progress of the students and also to identify if there is any learning difficulty to that particular content/ concept. Preparing and using teacher-made test in teaching and learning is a

regular task of teachers. For preparing teacher-made tests, objective type or essay type items or both can be constructed.

3.2 TEACHER-MADE ACHIEVEMENT TESTS - EXAMPLES

Classroom tests, chapter-end or unit-end test, quarterly, half-yearly, annual, and pre-board examinations, etc.

3.3 PURPOSE OF TEACHER MADE ACHIEVEMENT TESTS

Teacher made achievement tests especially provides an idea to the teachers about the progress and mastery of learning of the students. It also provides teachers feedback about how effective the teaching was, and whether the learning objectives have been achieved or not. It also provides teachers enough feedback to re-design his/her teaching and learning conditions to make it more effective.

- It can be used for a variety of instructional purposes.
- To measure whether students possess the pre-requisite skills needed to succeed in a unit or course or to what extent students have already achieved the objectives of the planned instruction (Placement Evaluation)
- To monitor students' learning progress and to provide ongoing feedback to students and teacher about the success of the teaching-learning process (Formative Evaluation)
- To identify students' learning difficulties in any areas of learning, to investigate the causes of the learning difficulties and to provide adequate remedial instruction to maintain the gap of learning.
- To assign grades and certify the students at the completion of any semester, year or entire programme (Summative evaluation).

3.4 CONSTRUCTION OF TEACHER MADE ACHIEVEMENT TESTS

There are five major steps involved in the process of construction of teacher made achievement tests they are:

1. Identifying instructional objectives
2. Making the design
3. Preparing blueprint
4. Writing the test items
5. Developing marking scheme

Let us now discuss the above steps in detail.

1. **Identifying Instructional Objectives**

The first and the most important step in planning a test is to identify the instructional objectives. Each subject has a different set of instructional objectives. Generally, in the subjects of Science, Social Science, and Mathematics the major objectives are categorized into knowledge, understanding, application and skill, while in Languages the major objectives are categorized into knowledge, understanding, and expression. Knowledge objectives are considered to be the lowest level of learning whereas understanding, application of knowledge are considered higher levels of learning.

For detailed understanding of instructional objectives you must read Bloom's Taxonomy of Educational Objectives. For writing instructional objectives, appropriate action verbs should be used. A few examples of action verbs that can be used are: name, list, write, differentiate, compare, contrast, describe, illustrate, define, explain, apply, underline, select, analyse, describe, etc.

2. Making the Design

The second step in planning a test is to make the 'Design'. The design specifies weightages to different aspects of the test, such as : (a) instructional objectives, (b) types (or forms) of questions, (c) units and sub-units of the course content covered, and (d) levels of difficulty. It also indicates as to whether there are any options in the question paper, and if so, what their nature is. The design, in fact, is termed as an instrument which reflects major policy decisions of the examining agency, whether it is a Board or a school.

3. **Preparing Blueprint**

The third step is to prepare the 'Blueprint'. The policy decisions, as reflected in the design of the question paper, are translated into action through the blueprint. It is the stage where the paper setter decides on how many questions are to be set for different objectives. Further she/he decides under which unit/topic a particular question is to be set. Furthermore, she/he picks up various forms of questions.

Thereafter, the paper setter decides how all the questions are to be distributed over different objectives and content areas so as to obtain the weightages decided in the design.

The blueprint is basically a two-dimensional chart that consists contents and objectives. Apart from these, form of questions and in terms of marks are also other aspects of blueprint. Once the blueprint is prepared, the paper setter can write/select the items and prepare the question pape

4. Writing the Test Items

The next step, after the finalization of the blueprint is writing appropriate questions in accordance with the broad parameters set out in the blueprint.

It proceeds in several ways, like

Either writing all questions (one by one) belonging to one objective at a time i.e. knowledge or understanding or application followed by other objectives, or By taking up questions according to their form or type i.e. Long Answer or Essay Type followed by Short Answer and Very Short Answer Type or in any other order, or By writing questions for one unit of the syllabus or portion to be covered by the test at a time. Each approach has its advantages and disadvantages, too. Irrespective of the method followed, the questions then have to be arranged in a logical sequence.

5. Marking Scheme

The fifth step is to prepare the 'Marking Scheme'. The marking scheme helps in preventing inconsistency in judgement. In the marking scheme, possible responses to items in the test are structured. The various value points for response are graded and the marks allowed to each value point indicated. The marking scheme ensures objectivity in judgement and eliminates differences in score which may be due to subjectivity of the evaluator.

The marking scheme, of course, includes the scoring key, which is prepared in respect of objective type questions.

3.5 ADMINISTRATION OF TEACHER MADE ACHIEVEMENT TESTS

Having prepared a good test, you should plan to administer it in such a way that each of your students will do their best. Motivating students is very important, and this is an area in which each teacher will have her/his own special technique. Let them understand the advantages of the class test. Make them understand that such tests help them to get a feedback on their weaknesses and the concept that they have not understood; which can be corrected before they face external examinations

Some tips to be kept in mind while planning for the administration of a test

- Time Schedule be sure that time schedule is planned carefully, ensuring teacher and pupil readiness. Much preparation may be done a day before
- It will be wise to schedule enough time for briefing the invigilators. If there is a deadline for finishing the test and leaving the room
- The Room It is important for any examination to provide a quiet, comfortable atmosphere, in which the students are encouraged to do their best. As much as possible, try to test in a quiet place with a minimum of distracting noises.
- Avoid rooms near cafeterias, common rooms, playing fields or other noisy places. Request nearby loudspeaker owners to shut them off for the duration of the examination hours.
- Hang signs on the door, saying ËXAMINATION IN PROGRESS: DO NOT DISTURB". Objective examinations generally require more intense concentration than essay type exams.
- Remember that the students will be writing on a single – thick answer sheet, not a thick answer book. Be sure the writing surfaces are at least 30 X 80 cm. and as smooth as possible.
- Be sure the room is clear of any charts, posters, etc. that might help some candidates. d) Equipment It is wise to make up a check-list, ahead of time, of what you will have to take with you to the examination hall.
- Be sure to include chalk to write necessary notices on the black board. If there is no black board; then make placards or poster ahead of time.
- For exact timing of the test (much more important for objective tests), it is better to have two watches or clocks.
- Invigilators For anything more that an informal, half-period quiz, you will probably need the help of one or more invigilators. Chose persons who are willing to give their full attention to the task. Neither you nor your invigilators should talk, read, correct papers or do any other work during the examination time. They should observe closely, circulating constantly, checking that the students are answering in the right place, not copying,

etc. However, they should not hover too long over any student, as this makes the examinee nervous.

3.6 SCORINGAND RECORDING OF TEST RESULTS

Despite the objectivity of scoring short answer tests, certain procedures are indispensable if scoring is to be done with maximum accuracy and efficiency. The necessity for extreme care in scoring has been indicated by several studies showing that scoring errors occur with appalling frequency. “Constant” errors can be due to failure to understand scoring directions, with resultant scores which are consistently too low or too high. “Variable” errors can be due to carelessness in marking, adding, computing, or transcribing scores.

Order of Scoring - With essay tests it may be desirable to have one person score all answers to the first question, then to the second, and so on. If, for objective tests separate answer sheets are provided, the scorer may score a given page in all booklets first, then the next page, and so on, rather than scoring all of one booklet before going on to the next. If so many booklets are to be scored that several scorers are needed, each person may specialize on a given page or group of pages of the booklet but should score only one page in all booklets at a time.

Rescoring - With a large number of booklets to be scored and sufficient help available, it is always worthwhile to re-score them so as to eliminate errors that otherwise are almost inevitable in a clerical task like this. If complete rescoring is not feasible, then every fifth or tenth booklet should be rescored to get a rough idea of the frequency and magnitude of scoring errors. Rescoring a sample sometimes uncovers such an inaccuracy as to make it desirable to re-score the remainder.

KEEPING RECORDS

As soon as possible after the tests have been administered, the answer sheet should be checked and scored, and the scores should be recorded on the permanent records of the school. Each teacher should be given copies of the score reports for the pupils in his/her classes. Usually schools have some type of permanent record for each pupil which provides space for recording test results.

The form in which test results are recorded is often meaningless to anyone except the persons recording them.

REPORTING AND INTERPRETATION OF TEST SCORES

After administering and scoring of a test, next step is reporting and interpretation of the scores. Learning and development must be presented on a very brief report form that is understandable to a variety of users (e.g., students, parents, teachers, counselors, and administrators). In most of the cases, the school policies guide a teacher in reporting the test scores. School grading and reporting systems are designed to serve a variety of functions, like instructional use, report to parents, and administrative and guidance uses. The focus of the grading and reporting system should be towards the improvement of student learning and development.

3.7 OBJECTIVE TYPE TEST:

An objective type test is one which is free from any subjective bias either from the tester or the marker. It refers to any written test that requires the examinee to select the correct answer from among one or more of several alternatives or supply a word or two and that demands an objective judgement when it is scored.

Objective-type tests - characteristics

- They are pin-pointed, definite and so clear that a single, definite answer is expected.
- They ensure perfect objectivity in scoring. The scoring will not vary from examiner to examiner.

Types Objective Type Test

Matching Test, Multiple Choice Test, True False Tests, correct/Incorrect Test, Simple Recall Test, Best Answer Test, Completion Test, and Classification Test.

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.

Limitations of Objective Type Test:

1. Objectives like ability to organise matter, ability to present matter logically and in a coherent fashion, etc., cannot be evaluated.
2. Guessing is possible. No doubt the chances of success may be reduced by the inclusion of a large number of items.
3. If a respondent marks all responses as correct, the result may be misleading
4. Construction of the objective test items is difficult while answering them is quite easy.
5. They demand more of analysis than synthesis.
6. Linguistic ability of the testee is not at all tested.
7. Printing cost considerably greater than that of an essay test.

Guidelines for Constructing Better Objective Type Test Items:

To be a good item writer, one should have:

- A thorough understanding of the subject matter;
- A thorough understanding of the pupils teste
- Perseverance; and A little creativity to prepare fertile kind of items.
- It is of paramount importance for him to be cognizant of the pitfalls involved in writing objective type test items.

1. Each item must be clearly expressed i.e. there must be precision in writing the test items.
2. Test for important facts and knowledge and not for trivial details; e.g.,

3. Give the name of the ship that Columbus was on when he discovered America.
4. Give the date (and/or time) when Edison invented the light bulb.
5. These items test the ability to recall or supple trivial details and therefore are unsound.
6. Avoid ambiguous statements. Each item should be subjected to one and only one interpretation.
7. Quantitative rather than qualitative words should be used. Words such as few, many, low, high, large, etc. are vague, indefinite, and, therefore, should be avoided.
8. Use good grammar and sentence structure to improve clarity.
9. Avoid lifting statements verbatim from the text-book. The use of text book language in a test encourages a pupil to memorise rather than to understand the subject matter.
10. Avoid negative questions whenever possible. Directions to questions should be specific. Ambiguous wording and double negatives should be avoided in questions

3.8 ESSAY TYPE TEST:

The essay tests are still commonly used tools of evaluation, despite the increasingly wider applicability of the short answer and objective type questions. There are certain outcomes of learning (e.g., organising, summarising, integrating ideas and expressing in one's own way) which cannot be satisfactorily measured through objective type tests. The importance of essay tests lies in the measurement of such instructional outcomes.

An essay test may give full freedom to the students to write any number of pages. The required response may vary in length. An essay type question requires the pupil to plan his own answer and to explain it in his own words. The pupil exercises considerable freedom to select, organise and present his ideas. Essay type tests provide a better indication of pupil's real achievement in learning. The answers provide a clue to nature and quality of the pupil's thought process.

That is, we can assess how the pupil presents his ideas (whether his manner of presentation is coherent, logical and systematic) and how he

concludes. In other words, the answer of the pupil reveals the structure, dynamics and functioning of pupil's mental life.

The essay questions are generally thought to be the traditional type of questions which demand lengthy answers. They are not amenable to objective scoring as they give scope for halo-effect, inter-examiner variability and intra-examiner variability in scoring.

3.9 Types of Essay Test:

There can be many types of essay tests:

Some of these are given below with examples from different subjects:

1. Selective Recall.

e.g. What was the religious policy of Akbar?

2. Evaluative Recall.

e.g. Why did the First War of Independence in 1857 fail?

3. Comparison of two things—on a single designated basis.

e.g. Compare the contributions made by Dalton and Bohr to Atomic theory.

4. Comparison of two things—in general.

e.g. Compare Early Vedic Age with the Later Vedic Age.

5. Decision—for or against.

e.g. Which type of examination do you think is more reliable? Oral or Written. Why?

6. Causes or effects.

e.g. Discuss the effects of environmental pollution on our lives.

7. Explanation of the use or exact meaning of some phrase in a passage or a sentence.

e.g., Joint Stock Company is an artificial person. Explain 'artificial person' bringing out the concepts of Joint Stock Company.

8. Summary of some unit of the text or of some article.

9. Analysis

e.g. What was the role played by Mahatma Gandhi in India's freedom struggle?

10. Statement of relationship.

e.g. Why is knowledge of Botany helpful in studying agriculture?

11. Illustration or **examples** (your own) of principles in science, language, etc.

e.g. Illustrate the correct use of subject-verb position in an interrogative sentence.

12. Classification.

e.g. Classify the following into Physical change and Chemical change with explanation. Water changes to vapour; Sulphuric Acid and Sodium Hydroxide react to produce Sodium Sulphate and Water; Rusting of Iron; Melting of Ice.

13. Application of rules or principles in given situations.

e.g. If you sat halfway between the middle and one end of a sea-saw, would a person sitting on the other end have to be heavier or lighter than you in order to make the sea-saw balance in the middle. Why?

14. Discussion.

e.g. Partnership is a relationship between persons who have agreed to share the profits of a business carried on by all or any of them acting for all. Discuss the essentials of partnership on the basis of this partnership.

15. Criticism—as to the adequacy, correctness, or relevance—of a printed statement or a classmate's answer to a question on the lesson.

e.g. What is the wrong with the following statement?

The Prime Minister is the sovereign Head of State in India.

16. Outline.

e.g. Outline the steps required in computing the compound interest if the principal amount, rate of interest and time period are given as P, R and T respectively.

17. Reorganization of facts.

e.g. The student is asked to interview some persons and find out their opinion on the role of UN in world peace. In the light of data thus collected he/she can reorganise what is given in the text book.

18. Formulation of questions-problems and questions raised.

e.g. After reading a lesson the pupils are asked to raise related problems-questions.

19. New methods of procedure

e.g. Can you solve this mathematical problem by using another method?

3.10 Advantages of the Essay Tests:

1. It is relatively easier to prepare and administer a six-question extended-response essay test than to prepare and administer a comparable 60-item multiple-choice test items.
2. It is the only means that can assess an examinee's ability to organise and present his ideas in a logical and coherent fashion.
3. It can be successfully employed for practically all the school subjects.
4. Some of the objectives such as ability to organise idea effectively, ability to criticise or justify a statement, ability to interpret, etc., can be best measured by this type of test.
5. Logical thinking and critical reasoning, systematic presentation, etc. can be best developed by this type of test.
6. It helps to induce good study habits such as making outlines and summaries, organising the arguments for and against, etc.
7. The students can show their initiative, the originality of their thought and the fertility of their imagination as they are permitted freedom of response.
8. The responses of the students need not be completely right or wrong. All degrees of comprehensiveness and accuracy are possible.
9. It largely eliminates guessing.
10. They are valuable in testing the functional knowledge and power of expression of the pupil.

3.11 Limitations of Essay Tests:

1. One of the serious limitations of the essay tests is that these tests do not give scope for larger sampling of the content. You cannot sample the course content so well with six lengthy essay questions as you can with 60 multiple-choice test items.
2. Such tests encourage selective reading and emphasise cramming.
3. Moreover, scoring may be affected by spelling, good handwriting, coloured ink, neatness, grammar, length of the answer, etc.
4. The long-answer type questions are less valid and less reliable, and as such they have little predictive value.

5. It requires an excessive time on the part of students to write; while assessing, reading essays is very time-consuming and laborious.
6. It can be assessed only by a teacher or competent professionals.
7. Improper and ambiguous wording handicaps both the students and valuers.
8. Mood of the examiner affects the scoring of answer scripts.
9. There is halo effect-biased judgement by previous impressions.
10. The scores may be affected by his personal bias or partiality for a particular point of view, his way of understanding the question, his weightage to different aspect of the answer, favouritism and nepotism, etc.

- Poor predictive validity,
- Limited content sampling,
- Scores unreliability, and
- Scoring constraints.

3.12 SUGGESTIONS FOR IMPROVING ESSAY TESTS:

The teacher can sometimes, through essay tests, gain improved insight into a student's abilities, difficulties and ways of thinking and thus have a basis for guiding his/her learning.

While Framing Questions:

- Give adequate time and thought to the preparation of essay questions, so that they can be re-examined, revised and edited before they are used. This would increase the validity of the test.
- The item should be so written that it will elicit the type of behaviour the teacher wants to measure. If one is interested in measuring understanding, he should not ask a question that will elicit an opinion;
- e.g., "What do you think of Buddhism in comparison to Jainism?"
- Use words which themselves give directions e.g. define, illustrate, utline, select, classify, summarise, etc., instead of discuss, comment, explain, etc.Give specific directions to students to elicit the desired response.
- Indicate clearly the value of the question and the time suggested for answering it.

- Do not provide optional questions in an essay test because—
- Prepare and use a relatively large number of questions requiring short answers rather than just a few questions involving long answers.
- Do not start essay questions with such words as list, who, what, whether. If we begin the questions with such words, they are likely to be short-answer question and not essay questions, as we have defined the term.
- Adapt the length of the response and complexity of the question and answer to the maturity level of the students.
- The wording of the questions should be clear and unambiguous.
- It should be a power test rather than a speed test. Allow a liberal time limit so that the essay test does not become a test of speed in writing.
- Supply the necessary training to the students in writing essay tests.
- Questions should be graded from simple to complex so that all the testees can answer atleast a few questions.
- Essay questions should provide value points and marking schemes.

While Scoring Questions:

Prepare a marking scheme, suggesting the best possible answer and the weightage given to the various points of this model answer. Decide in advance which factors will be considered in evaluating an essay response.

While assessing the essay response, one must:

- Use appropriate methods to minimise bias;
- Pay attention only to the significant and relevant aspects of the answer;
- Be careful not to let personal idiosyncrasies affect assessment;
- Apply a uniform standard to all the papers.
- The examinee's identity should be concealed from the scorer. By this we can avoid the "halo effect" or "biasness" which may affect the scoring.
- Check your marking scheme against actual responses.
- Once the assessment has begun, the standard should not be changed, nor should it vary from paper to paper or reader to reader. Be consistent in your assessment.
- Grade only one question at a time for all papers. This will help you in minimizing the halo effect in becoming thoroughly familiar with just one set of scoring criteria and in concentrating completely on them.

- The mechanics of expression (legibility, spelling, punctuation, grammar) should be judged separately from what the student writes, i.e. the subject matter content.
- If possible, have two independent readings of the test and use the average as the final score.

3.13 CHARECTERISTICS OF A GOOD TEST

In order for a test to be a good tool for measuring students' knowledge and skills, it should have the following characteristics of examination that are essential for the success of any test.

1. Reliability or Consistency

Reliability or consistency of a test means that learners should perform the same or get the same score if they are exposed to different questions in different times and places. A test is considered reliable when the same result is achieved over different tests.

"The reliability of test scores is the extent to which they are consistent across different occasions of testing, different editions of the test, or different raters scoring the test taker's responses."

How to Make Sure Your Test Is Reliable?

I. **Score Distribution:** The percentage of test takers at each score level.

II. **Mean Score**: The average score, computed by summing the scores of all test takers and dividing by the number of test takers.

III. **Standard Deviation**: A measure of the amount of variation in a set of scores. It can be interpreted as the average distance of scores from the mean. (Actually, it is a special kind of average called a "root mean square," computed by squaring the distance of each score from the mean score, averaging the squared distances, and then taking the square root.)

IV. **Correlation**: A measure of the strength and direction of the relationship between the scores of the same people on two tests.

There are several methods of estimating reliability coefficients. They include:

1. Test-retest method that yields measures of stability

2. Equivalent forms method that yields measures of equivalence
3. Split half method that yields measures of internal consistency
4. Kuder-Richardson method that yields measures of internal consistency
5. Cronbach Alpha method that yields measures of internal consistency.

2. Validity

A validity of a test can be achieved when the test measures what it is really intended to measure. Therefore, a certain criteria must be selected.

Validity is very important to gauge the quality of a given test as questions must be in line with the selected criteria and measures.

Here are some of the top different types of validity:

Content Validity: A test should fairly represent the content of the course or the field of study.

Criterion Validity: It is used to predict the performance of a job applicant or a student.

Convergent validity: This is mostly used in the field of sociology or psychology.

Discriminant Validity: Discriminant validity means that a test of a concept is not highly correlated with other tests that are set to measure theoretically different concepts

3. Objectivity

"Objectivity of a test refers to the degree to which equally competent scores obtain the same results," the test should be away from any personal or subjective judgment. It should be based only on the evaluation of human development.

For example, in an essay-type test, students answer differently as each one has his/her own style of writing.

Hence, when more than one instructor check the test, they may give different scores according to whether they like the style or not. So, here, the test is less objective.

To avoid such bias, sharp rules should be set in evaluating such types of tests. There should be a unified guide for teachers to use while correcting such tests.

Personal judgment does not occur in true or false or multiple choice tests. Besides, teachers should receive training on how to score a test as untrained

teachers may give wrong scores and not be able to maintain the required fairness and accuracy.

4. Comprehensiveness

A test should fully cover the entire field of study that students are exposed to during the course. Vague questions should not be included especially during online tests when students are confused and short in time.

Absence of Ambiguity

There has to be no place for ambiguity especially in online tests where examiners are absent. Students should not be left in confusion and all questions have to be crystal clear.

Ambiguous questions often result when instructors put off writing test questions until the last minute. Careful editing and an independent review of the test items can help to minimize this problem."

6. Preparation

To ensure the success of any test, instructors should take into consideration the following factors:

- Students have to be well-prepared for the test through extensive revisions and discussions.
- There should not be any gaps between the revision period and the exam.
- Examiners should make it clear to students which topics are expected to be tackled in the exam.
- Students should be well-trained for the test type.

7. Appropriateness of Time

One of the top characteristics of a good test is when students have appropriate time to answer all questions. For example, essay questions require more time than multiple choice or true/false questions. A good test is supposed to be practical and comprehensive.

8. Discrimination

Discriminating power of a test shows how well an item discriminates between the bright and the dull students. It indicates whether the item is measuring the same ability as the test measures. It is a measure of correlation between the item and the total test score.

9. Usability

Test should solve desired problems or solve intended constructs, then such a test is as good test

10. Acceptability

A good test should be acceptable to the learners for whom it is intended. It should also be acceptable to teachers, parents, and other members of the society

11. Adequacy

We cannot assume that a comprehensive test can measure all the elements of knowledge and skills that a learner must acquire in completing a course.

12. Purpose

A test may possess all the important characteristics of a good test and yet it may be of no value for use in a situation. Unless tests are selected or constructed for definite purposes and used in an intelligent manner to achieve the desired results, they are of little value and may even be harmful

13. Economy

This may be economy of time and/or economy in cost. On the economy of time, tests requiring long time are not acceptable to students, parents, as well as markers.

14. Meaningfulness of test score

Generally, a single score is obtained from a test which is likely to be more meaningful than the several different scores. It has to be specified what the overall score conveys, what scores on separate sub-tests convey, or what the various combinations of scores convey to the tester.