

ASSESSMENT FOR LEARNING

2.9 (1st Half)

B.Ed- 2nd Sem.

Unit-I, II, III & IV

By- Chiranjib Sarkar

Department of Education

UNIT-I

Meaning of Assessment :-

In education, the term assessment refers to the wide variety of methods that educators use to evaluate, measure, and document the academic readiness, learning progress, and skill acquisition of students from preschool through college and adulthood. It is the process of systematically gathering information as part of an evaluation. Assessment is carried out to see what children and young people know, understand and are able to do. Assessment is very important for tracking progress, planning next steps, reporting and involving parents, children and young people in learning.

Techniques of Assessment:-

- Examination- Tools (Questionnaire)
- Observation- Tools (Observation shedule)
- Interview- Tools (Structured interview bshedule)
- Assignment- Tools (Criteria for developing)
- Projects- Tools (Criteria for developing)
- Debates- Tools (Time/ value portion)
- Group discussion- Tools (Rating scale)

Assessment- Quantitative+Qualitative

Measurement- Quantitative

Evaluation- Quantitative+Qualitative+Value judgement

Types Of Assessment:-

- **Pre-assessment or diagnostic assessment:**

Before creating the instruction, it's necessary to know for what kind of students you're creating the instruction. Your goal is to get to know your student's strengths, weaknesses and the skills and knowledge the posses before taking the instruction. Based on the data you've collected, you can create your instruction.

- **Formative assessment:**

Formative assessment is used in the first attempt of developing instruction. The goal is to monitor student learning to provide feedback. It helps identifying the first gaps in your instruction. Based on this feedback you'll know what to focus on for further expansion for your instruction.

- **Summative assessment:**

Summative assessment is aimed at assessing the extent to which the most important outcomes at the end of the instruction have been reached. But it measures more: the effectiveness of learning, reactions on the instruction and the benefits on a long-term base. The long-term benefits can be determined by following students who attend your course, or test. You are able to see whether and how they use the learned knowledge, skills and attitudes.

- **Confirmative assessment:**

When your instruction has been implemented in your classroom, it's still necessary to take assessment. Your goal with confirmative assessments is to find out if the instruction is still a success after a year, for example, and if the way you're teaching is still on point. You could say that a confirmative assessment is an extensive form of a summative assessment.

- **Norm-referenced assessment:**

This compares a student's performance against an average norm. This could be the average national norm for the subject History, for example. Other example is when the teacher compares the average grade of his or her students against the average grade of the entire school.

- **Criterion-referenced assessment:**

It measures student's performances against a fixed set of predetermined criteria or learning standards. It checks what students are expected to know and be able to do at a specific stage of their education. Criterion-referenced tests are used to evaluate a specific body of knowledge or skill set, it's a test to evaluate the curriculum taught in a course.

- **Ipsative assessment:**

It measures the performance of a student against previous performances from that student. With this method you're trying to improve yourself by comparing previous results. You're not comparing yourself against other students, which may be not so good for your self-confidence.

What is testing?

What is testing in education? Almost everybody has experienced testing during his or her life. Grammar tests, driving license test etc. A test is used to examine someone's knowledge of something to determine what that person knows or has learned. It measures the level of skill or knowledge that has been reached. An evaluative device or procedure in which a sample of an examinee's behavior in a specified domain is obtained and subsequently evaluated and scored using a standardized process (The Standards for Educational and Psychological Testing, 1999)

Difference between Test and Assessment:-

Test and assessment are used interchangeably, but they do mean something different. A test is a "product" that measures a particular behavior or set of objectives. Meanwhile assessment is seen as a procedure instead of a product. Assessment is used during and after the instruction has taken place. After you've received the results of your assessment, you can interpret the results and in case needed alter the instruction. Tests are done after the instruction has taken place, it's a way to complete the instruction and get the results. The results of the tests don't have to be interpreted, unlike assessment.

What is evaluation?

What's the definition of evaluation in education? Evaluation focuses on grades and might reflect classroom components other than course content and mastery level. An evaluation can be used as a final review to gauge the quality of instruction. It's product-oriented. This means that the main question is: "What's been learned?" In short, evaluation is judgmental.

Example:

You're gifted a flower.

Evaluation: "The flower is purple and is too short with not enough leaves."

Evaluation is judgmental

Assessment: "I'll give the flower some water to improve its growth."

Assessment increases the quality

Relationship between Assessment and Evaluation:-

Besides the differences, there are also some similarities between assessment and evaluation. The both require criteria, use measures and are evidence-driven.

So, what's the difference?

Assessment

Is ongoing

Improves learning quality

Individualized

Ungraded

Provides feedback

Process-oriented

Evaluation

Provides closure

Judges learning level

Applied against standards

Graded

Shows shortfalls

Product-oriented

What is an examination?

As with a test you test the knowledge of your students with an exam. The exam consists of a series of questions. They can both multiple choice or free text questions, or a different format. In the end you want to grade the answers and assign a score to every student. The end result makes a student pass or fail this exam.

What's the difference between a test and examination?

The big difference between a test and an exam is that an exam is more formal than a test. But, that said, they are used as synonyms throughout schools and courses. But, we think they serve a different purpose:

1. The test is a tool to measure the knowledge level of your students and adjust the learning material accordingly. With the purpose to have your students learn.

2. An exam or the examination is more formal and it tells you if a student passed or failed a class or course. In most cases you have to study again and re-take the exam. Or start the course or class all over again.

Measurement:-

Measurement is an essential element of education research. It is of extreme importance for determining the efficacy and legitimacy of educational practices that are now operationalized without benefit of professional consensus. Measures can be evaluated on the basis of multiple criteria; a discussion of the more important of these criteria is offered. Measures that prove viable are of special importance in three areas: the evaluation of students, communication among educators, and the investigation of hypotheses. Within the definition that is offered, measurement procedures may involve a variety of formats; these are briefly reviewed and evaluated.

What is constructivism?

Constructivism is basically a theory -- based on observation and scientific study -- about how people learn. It says that people construct their own understanding and knowledge of the world, through experiencing things and reflecting on those experiences. When we encounter something new, we have to reconcile it with our previous ideas and experience, maybe changing what we believe, or maybe discarding the new information as irrelevant. In any case, we are active creators of our own knowledge. To do this, we must ask questions, explore, and assess what we know.

In the classroom, the constructivist view of learning can point towards a number of different teaching practices. In the most general sense, it usually means encouraging students to use active techniques (experiments, real-world problem solving) to create more knowledge and then to reflect on and talk about what they are doing and how their understanding is changing. The teacher makes sure she understands the students' preexisting conceptions, and guides the activity to address them and then build on them.

Constructivist teachers encourage students to constantly assess how the activity is helping them gain understanding. By questioning themselves and their strategies, students in the constructivist classroom ideally become "expert learners." This gives them ever-broadening tools to keep learning. With a well-planned classroom environment, the students learn HOW TO LEARN.



You might look at it as a spiral. When they continuously reflect on their experiences, students find their ideas gaining in complexity and power, and they develop increasingly strong abilities to integrate new information. One of the teacher's main roles becomes to encourage this learning and reflection process.

For example: Groups of students in a science class are discussing a problem in physics. Though the teacher knows the "answer" to the problem, she focuses on helping students restate their questions in useful ways. She prompts each student to reflect on and examine his or her current knowledge. When one of the students comes up with the relevant concept, the teacher seizes upon it, and indicates to the group that this might be a fruitful avenue for them to explore. They design and perform relevant experiments. Afterward, the students and teacher talk about what they have learned, and how their observations and experiments helped (or did not help) them to better understand the concept.

Contrary to criticisms by some (conservative/traditional) educators, constructivism does not dismiss the active role of the teacher or the value of expert knowledge. Constructivism modifies that role, so that teachers help students to construct knowledge rather than to reproduce a series of facts. The constructivist teacher provides tools such as problem-solving and inquiry-based learning activities with which students formulate and test their ideas, draw conclusions and inferences, and pool and convey their knowledge in a collaborative learning environment. Constructivism transforms the student from a passive recipient of information to an active participant in the learning process. Always guided by the teacher, students construct their knowledge actively rather than just mechanically ingesting knowledge from the teacher or the textbook.

Constructivism is also often misconstrued as a learning theory that compels students to "reinvent the wheel." In fact, constructivism taps into and triggers the student's innate curiosity about the world and how things work. Students do not reinvent the wheel but, rather, attempt to understand how it turns, how it functions. They become engaged by applying their existing knowledge and real-world experience, learning to hypothesize, testing their theories, and ultimately drawing conclusions from their findings.

The best way for you to really understand what constructivism is and what it means in your classroom is by seeing examples of it at work, speaking with others about it, and trying it yourself. As you progress through each segment of this workshop, keep in mind questions or ideas to share with your colleagues.

Principles of Assessment:-

Principle 1 - Assessment should be valid Validity ensures that assessment tasks and associated criteria effectively measure student attainment of the intended learning outcomes at the appropriate level.

Principle 2 - Assessment should be reliable and consistent There is a need for assessment to be reliable and this requires clear and consistent processes for the setting, marking, grading and moderation of assignments.

Principle 3 - Information about assessment should be explicit, accessible and transparent Clear, accurate, consistent and timely information on assessment tasks and procedures should be made available to students, staff and other external assessors or examiners.

Principle 4 - Assessment should be inclusive and equitable As far as is possible without compromising academic standards, inclusive and equitable assessment should ensure that tasks and procedures do not disadvantage any group or individual.

Principle 5 - Assessment should be an integral part of programme design and should relate directly to the programme aims and learning outcomes Assessment tasks should primarily reflect the nature of the discipline or subject but should also ensure that students have the opportunity to develop a range of generic skills and capabilities.

Principle 6 - The amount of assessed work should be manageable The scheduling of assignments and the amount of assessed work required should provide a reliable and valid profile of achievement without overloading staff or students.

Principle 7 - Formative and summative assessment should be included in each programme Formative and summative assessment should be incorporated into programmes to ensure that the purposes of assessment are adequately addressed. Many programmes may also wish to include diagnostic assessment.

Principle 8 - Timely feedback that promotes learning and facilitates improvement should be an integral part of the assessment process Students are entitled to feedback on submitted formative assessment tasks, and on summative tasks, where appropriate. The nature, extent and timing of feedback for each assessment task should be made clear to students in advance.

Principle 9 - Staff development policy and strategy should include assessment All those involved in the assessment of students must be competent to undertake their roles and responsibilities.

Why Is Assessment Important?

Assessment is an integral part of instruction, as it determines whether or not the goals of education are being met. Assessment affects decisions about grades, placement, advancement, instructional needs, curriculum, and, in some cases, funding. Assessment inspire us to ask these hard questions: "Are we teaching what we think we are teaching?" "Are students learning what they are supposed to be learning?" "Is there a way to teach the subject better, thereby promoting better learning?"

Today's students need to know not only the basic reading and arithmetic skills, but also skills that will allow them to face a world that is continually changing. They must be able to think critically, to analyze, and to make inferences. Changes in the skills base and knowledge our students need require new learning goals; these new learning goals change the relationship between assessment and instruction. Teachers need to take an active role in making decisions about the purpose of assessment and the content that is being assessed.

The Difference Between Assessment Of Learning And Assessment For Learning:-

Assessment for learning is commonly referred to as formative assessment—that is, assessment designed to inform instruction. If we can agree that the purpose of assessment is to provide data to revise planned instruction, then the only **type of assessment** that's not 'assessment for learning' is 'assessment of learning,' commonly referred to as summative assessment.

Assessment is generally broken down into three categories: assessment before instruction (pre-assessment), assessment during instruction (formative assessment), and assessment after instruction (summative assessment). To further complicate matters, it could be argued that pre-assessment is both assessment of and for learning—that is, it assesses 'prior knowledge' (as a pre-assessment) and that data is then used to revise planned instruction (making it formative assessment).

In truth, most of this is semantics and a bit confusing. There are many ways to measure understanding and the primary distinction in most K-12 classrooms for most assessments is function: What is the assessment supposed to do? If you're using the 'test' so that you can see what students do and don't know so that you can more accurately plan future learning lessons

and activities, then it's assessment *for* learning (even if you're obviously doing so by performing an assessment *of* learning).

If instead, the assessment is merely a kind of benchmark to see 'how well they can do' and you're moving on, then it's primarily an assessment of learning. There is significant overlap between the two; in fact, the same test given in one circumstance would be considered an assessment of learning while in another circumstance be considered an assessment for learning.

In short then, the difference between assessment of learning and assessment for learning is a matter of function and purpose—a matter of 'who': assessment of learning is a way to see what the students can do while assessment for learning is a way to see what the teachers should do in response.

Difference between Assessment and Evaluation:-

ASSESSMENT	EVALUATION
Formative: Ongoing to improve learning	Summative: Final to gauge quality
Process-oriented: How learning is going	Product-oriented: What's been learned
Reflective: Internally defined criteria/goals	Prescriptive: Externally imposed standards
Diagnostic: Identify areas for improvement	Judgmental: Arrive at an overall grade/score
Flexible: Adjust as problems are clarified	Fixed: To reward success, punish failure
Absolute: Strives for ideal outcomes	Comparative: Divide better from worse
Cooperative: Learn from each other	Competitive: Beat each other out

Measurement Scales:

In statistics, there are four data measurement scales: nominal, ordinal, interval and ratio. These are simply ways to sub-categorize different types of data ([here's an overview of statistical data types](#)). This topic is usually discussed in the context of academic teaching and less often in the "real world." If you are brushing up on this concept for a statistics test, thank a psychologist researcher named [Stanley Stevens](#) for coming up with these terms.

These four data measurement scales (nominal, ordinal, interval, and ratio) are best understood with example, as you'll see below.

Nominal Scale:

Let's start with the easiest one to understand. Nominal scales are used for labeling variables, without any quantitative value. "Nominal" scales could simply be called "labels." Here are some examples, below. Notice that all of these scales are mutually exclusive (no overlap) and none of them have any numerical significance. A good way to remember all of this is that "nominal" sounds a lot like "name" and nominal scales are kind of like "names" or labels. A subtype of Nominal scale with only two categories is called "Dichotomous".

What is your gender?

- M - Male
- F - Female

What is your hair color?

- 1 - Brown
- 2 - Black
- 3 - Blonde
- 4 - Gray
- 5 - Other

Where do you live?

- A - North of the equator
- B - South of the equator
- C - Neither: In the international space station

Examples of Nominal Scales

Note: a sub-type of nominal scale with only two categories (e.g. male/female) is called "**dichotomous**." If you are a student, you can use that to impress your teacher.

Bonus Note #2: Other sub-types of nominal data are "nominal with order" (like "cold, warm, hot, very hot") and nominal without order (like "male/female").

Ordinal Scale:

With ordinal scales, the *order* of the values is what's important and significant, but the differences between each one is not really known. Take a look at the example below. In each case, we know that a #4 is better than a #3 or #2, but we don't know—and cannot quantify—how *much* better it is. For example, is the difference between "OK" and "Unhappy" the same as the difference between "Very Happy" and "Happy?" We can't say.

Ordinal scales are typically measures of non-numeric concepts like satisfaction, happiness, discomfort, etc.

“Ordinal” is easy to remember because it sounds like “order” and that’s the key to remember with “ordinal scales”—it is the *order* that matters, but that’s all you really get from these.

Advanced note: The best way to determine central tendency on a set of ordinal data is to use the mode or median; a purist will tell you that the mean cannot be defined from an ordinal set.

<p>How do you feel today?</p> <p><input checked="" type="radio"/> 1 - Very Unhappy</p> <p><input type="radio"/> 2 - Unhappy</p> <p><input type="radio"/> 3 - OK</p> <p><input type="radio"/> 4 - Happy</p> <p><input type="radio"/> 5 - Very Happy</p>	<p>How satisfied are you with our service?</p> <p><input checked="" type="radio"/> 1 - Very Unsatisfied</p> <p><input type="radio"/> 2 - Somewhat Unsatisfied</p> <p><input type="radio"/> 3 - Neutral</p> <p><input type="radio"/> 4 - Somewhat Satisfied</p> <p><input type="radio"/> 5 - Very Satisfied</p>
---	---

Example of Ordinal Scales

Interval Scale:

Interval scales are numeric scales in which we know both the order and the exact differences between the values. The classic example of an interval scale is Celsius temperature because the difference between each value is the same. For example, the difference between 60 and 50 degrees is a measurable 10 degrees, as is the difference between 80 and 70 degrees.

Interval scales are nice because the realm of statistical analysis on these data sets opens up. For example, *central tendency* can be measured by mode, median, or mean; standard deviation can also be calculated.

Like the others, you can remember the key points of an “interval scale” pretty easily. “Interval” itself means “space in between,” which is the important thing to remember—interval scales not only tell us about order, but also about the value between each item.

Here’s the problem with interval scales: they don’t have a “true zero.” For example, there is no such thing as “no temperature,” at least not with Celsius. In the case of interval scales, zero doesn’t mean the absence of value, but is actually another number used on the scale, like 0 degrees Celsius. Negative numbers also have meaning. Without a true zero, it is impossible to compute ratios. With interval data, we can add and subtract, but cannot multiply or divide.

Confused? Ok, consider this: 10 degrees C + 10 degrees C = 20 degrees C. No problem there. 20 degrees C is not twice as hot as 10 degrees C, however, because there is no such thing as “no temperature” when it comes to the Celsius scale. When converted to Fahrenheit, it’s clear: 10C=50F and 20C=68F, which is clearly not twice as hot. I hope that makes sense. Bottom line, interval scales are great, but we cannot calculate ratios, which brings us to our last measurement scale...



Example of Interval Scal

e

Ratio Scale:

Ratio scales are the ultimate nirvana when it comes to data measurement scales because they tell us about the order, they tell us the exact value between units, AND they also have an absolute zero—which allows for a wide range of both descriptive and inferential statistics to be applied. At the risk of repeating myself, everything above about interval data applies to ratio scales, plus ratio scales have a clear definition of zero. Good examples of ratio variables include height, weight, and duration.

Ratio scales provide a wealth of possibilities when it comes to statistical analysis. These variables can be meaningfully added, subtracted, multiplied, divided (ratios). Central tendency can be measured by mode, median, or mean; measures of dispersion, such as standard deviation and coefficient of variation can also be calculated from ratio scales.



This
weight)

Device Provides Two Examples of Ratio Scales (height and

Measurement error:-

Measurement error in education generally refers to either (1) the difference between what a test score indicates and a student's actual knowledge and abilities or (2) errors that are introduced when collecting and calculating data-based reports, figures, and statistics related to schools and students.

Because some degree of measurement error is inevitable in testing and data reporting, education researchers, statisticians, data professionals, and test developers often publicly acknowledge that performance data, such as high school graduation rates or college-enrollment rates, are not perfectly reliable (they may even report the "margin of error" for a given statistic or finding) or that test scores don't always accurately reflect what students know or can do—i.e., that there is no such thing as a perfectly reliable test of student knowledge and skill acquisition.

Measurement errors in testing may result from a wide variety of factors, such as a student's mental and emotional state during the test period or the conditions under which the test was administered. For example, students may have been unusually tired, hungry, or emotionally distressed, or distractions such as loud noises, disruptive peers, or technical problems could have adversely affected test performance. Test scores for young children are often considered to be especially susceptible to measurement error, given that young children tend to have shorter attention spans and they may not be able to fully comprehend the importance of the test and take it seriously. In addition, young children of the same chronological age or grade level may be at very different stages of social, cognitive, and emotional development, and if a young child experiences a rapid developmental growth spurt, test results could quickly become outdated and therefore misrepresentative.

The following is a representative list of a few additional factors and problems that may give rise to measurement error in testing:

- Ambiguously phrased questions or inaccurate answers.
- Test items, questions, and problems may not address the material students were actually taught.
- Performance levels and cutoff scores, such as those considered to be "passing" or "proficient" on a particular test, may be flawed, poorly calibrated, or misrepresentative.
- The scoring process may be poorly designed, and both human scorers and computer-scoring systems may make mistakes.
- Test administrators could give students incorrect directions, help students cheat, or fail to create calm and conducive test-taking conditions.
- Test-result data may be inaccurately recorded and reported.

UNIT-II

Instructional Goals and Objectives: Foundation for Assessment:-

Clearly stating learning goals and objectives plays a key role both in the instructional and assessment processes. These learning goals serve as guides as to what will be taught, how it will be taught, and whether the material has been mastered. Instructional goals and objectives work best when they describe, in terms of observable behaviors, what the student is expected to learn. This model of behaviorally stated goals results in the child producing or demonstrating products that indicate learning. However, other models of stating learning goals exist which are more concerned with the processes of learning and with the complexity of understanding and application by the student. Most individuals think of educational objectives as confined to cognitive or "school" learning goals. While it is true that schools and states are most concerned with these cognitive skills, other types of learning goals such as affective or psychomotor learning goals exist and are taught in schools. Cognitive, affective, and psychomotor learning have been grouped separately into their own taxonomies of objectives. Learning goals for students can originate from a variety of sources and professional organizations. A variety of criteria exist for choosing learning goals for students. Learning goals may be stated both in terms of general instructional objectives and specific learning outcomes. Both types of goals should be stated in clear, behavioral terms so that mastery can be clearly measured.

What is a Goal?

- Goals are broad, generalized statements about what is to be learned. Think of them as a target to be reached, or "hit."
- **Goal Example:** Students will gain an understanding of world cultures.

What is an Objective?

- Objectives are the foundation upon which you can build lessons and assessments that you can prove meet your overall course or lesson goals. Think of objectives as tools you use to make sure you reach your goals. They are the arrows you shoot towards your target (goal).
- **Objective Example:** Given a list of comparison points, students will compare the three major points of two given cultures in writing with 90% accuracy.

What are Instructional Objectives?

- Instructional objectives are specific, measurable, short-term, observable student behaviors.
- Objectives are the foundation upon which you can build lessons and assessments that you can prove meet your overall course or lesson goals.
- Think of objectives as tools you use to make sure you reach your goals. They are the arrows you shoot towards your target (goal).
- The purpose of objectives is not to restrict spontaneity or constrain the vision of education in the discipline; but to ensure that learning is focused clearly enough that both students and teacher know what is going on, and so learning can be objectively measured. Different archers have different styles, so do different teachers. Thus, you can shoot your arrows (objectives) many ways. The important thing is that they reach your target (goals) and score that bull's eye!

Characteristics of a Learning Objective:-

1. **Specific performance-** An objective always states what a learner is expected to be able to do and/or produce to be considered competent.

E.g., to write, to name, to compare and contrast, to analyze, to evaluate.

2. **Conditions-** An objective describes the important conditions under which the behavior is to occur.

E.g., during a cooperative activity, after reading chapter 1.

3. **Criterion, or standard-** An objective describes the criteria of acceptable performance; that is, it states how well someone would have to perform to be considered competent.

E.g., correct to the nearest ml, 80% correct, with no grammatical errors.

Learning Outcomes:-

Learning outcomes are an essential part of any unit outline. A learning outcome is a clear statement of what a learner is expected to be able to do, know about and/or value at the completion of a unit of study, and how well they should be expected to achieve those outcomes.

It states both the substance of learning and how its attainment is to be demonstrated. Learning outcomes not only serve the purpose of directing the content and design of a unit of study, they form the basis of assessment and are also linked to the larger outcomes of learning set by the University in the form of generic and/or course/discipline-specific graduate attributes. Because of their clear linkage to assessment, students will achieve the learning outcomes to differing degrees.

Characteristics of good learning outcomes:-

- Should be **flexible**: while individual outcomes should be specific, instructors should feel comfortable adding, removing, or adjusting learning outcomes over the length of a course if initial outcomes prove to be inadequate.
- Are **focused on the learner**: rather than explaining what the instructor will do in the course, good learning outcomes describe knowledge or skills that the student will employ, and help the learner understand why that knowledge and those skills are useful and valuable to their personal, professional, and academic future.
- Are **realistic**, not aspirational: all passing students should be able to demonstrate the knowledge or skill described by the learning outcome at the conclusion of the course. In this way, learning outcomes establish standards for the course.
- Focus on the **application** and **integration** of acquired knowledge and skills: good learning outcomes reflect and indicate the ways in which the described knowledge and skills may be used by the learner now and in the future.
- Indicate useful **modes of assessment** and the specific elements that will be assessed: good learning outcomes prepare students for assessment and help them feel engaged in and empowered by the assessment and evaluation process.
- Offer a **timeline** for completion of the desired learning.

Learning:-

Learning is a process that is often not under our control and is wrapped up with the environments we inhabit and the relationships we make. It involves encountering signals from the senses; attending to them; looking for connections and meanings; and framing them so that we may act.

Nature of Learning:-

- a. Learning is adaptation or adjustment: Friends, we all continuously interact with our environment. We often make adjustment and adapt to our social environment. Through a process of continuous learning, the individual prepares himself for necessary adjustment or adaptation. That is why learning is also described as a process of progressive adjustment to ever changing conditions, which one encounters.
- b. Learning is improvement: Learning is often considered as a process of improvement with practice or training. We learn many things, which help us to improve our performance.
- c. Learning is organizing experience: Learning is not mere addition of knowledge. It is the reorganization of experience.
- d. Learning brings behavioural changes: Whatever the direction of the changes may be, learning brings progressive changes in the behaviour of an individual. That is why he is able to adjust to changing situations.
- e. Learning is active: Learning does not take place without a purpose and self-activity. In any teaching learning process, the activity of the learner counts more than the activity of a teacher.
- f. Learning is goal directed: when the aim and purpose of learning is clear, an individual learns immediately. It is the purpose or goal, which determines what, the learner sees in the learning situations and how he acts. If there is no purpose or goal learning can hardly be seen.
- g. Learning is universal and continuous: All living creatures learn. Every moment the individual engages himself to learn more and more. Right from the birth of a child till the death learning continues.

The Five (or Six) Dimensions of Learning:-

Learning theorists have argued that learning and development are not like an assembly-line which can be broken down into discrete steps occurring with machine-time precision, but an organic process that unfolds in complex ways according to its own pace and rhythm. Teaching and learning occurs in complex ecosystems, dynamic environments where teachers, students, materials and supplies, texts, technologies, concepts, social structures, and architectures are interdependently related and interactive. Using the Learning Record, the teacher (and student) is actively searching for, and documenting, positive evidence of student development across five dimensions: confidence and independence, knowledge and understanding, skills and strategies, use of prior and emerging experience, and critical reflection. These five dimensions cannot be "separated out" and treated individually; rather, they are dynamically interwoven. Our goals for a

particular class should describe a trajectory of learning across multiple dimensions, and our measurements should be able to identify the paths taken by students and their progress from their individual starting points along that trajectory.

Individually, learners can expect to make progress across these five dimensions:

1) Confidence and independence:

We see growth and development when learners' confidence and independence become congruent with their actual abilities and skills, content knowledge, use of experience, and reflectiveness about their own learning. It is not a simple case of "more (confidence and independence) is better." In a science class, for example, an overconfident student who has relied on faulty or underdeveloped skills and strategies learns to seek help when facing an obstacle; or a shy student begins to trust her own abilities, and to insist on presenting her own point of view in discussion. In both cases, students are developing along the dimension of confidence and independence.

2) Skills and strategies:

Skills and strategies represent the "know-how" aspect of learning. When we speak of "performance" or "mastery," we generally mean that learners have developed skills and strategies to function successfully in certain situations. Skills and strategies are not only specific to particular disciplines, but often cross disciplinary boundaries. In a writing class, for example, students develop many specific skills and strategies involved in composing and communicating effectively, from research to concept development to organization to polishing grammar and correctness, and often including technological skills for computer communication.

3) Knowledge and Understanding:

Knowledge and understanding refers to the "content" knowledge gained in particular subject areas. Knowledge and understanding is the most familiar dimension, focusing on the "know-what" aspect of learning. In a psychology class, knowledge and understanding might answer a wide range of questions such as, What is Freud's concept of ego? Who was Carl Jung? What is "behaviorism"? These are typical content questions. Knowledge and understanding in such classes includes what students are learning about the topics; research methods; the theories, concepts, and practices of a discipline; the methods of organizing and presenting our ideas to others, and so on.

4) Use of prior and emerging experience:

The use of prior and emerging experience involves learners' abilities to draw on their own experience and connect it to their work. A crucial but often unrecognized dimension of learning is the capacity to make use of prior experience as well as emerging experience in new situations. It is necessary to observe learners over a period of time while they engage in a variety of activities in order to account for the development of this important capability, which is at the heart of creative thinking and its application. With traditional methods of evaluating learning, we cannot discover just how a learner's prior experience might be brought to bear to help scaffold new understandings, or how ongoing experience shapes the content knowledge or skills and strategies the learner is developing. In a math class, students scaffold new knowledge through applying the principles and procedures they've already learned: algebra depends on the capacity to apply basic arithmetic procedures, for example.

5) Reflection:

Reflection refers to the developing awareness of the learner's own learning process, as well as more analytical approaches to the subject being studied. When we speak of reflection as a crucial component of learning, we are not using the term in its commonsense meaning of reverie or abstract introspection. We are referring to the development of the learner's ability to step back and consider a situation critically and analytically, with growing insight into his or her own learning processes, a kind of metacognition. It provides the "big picture" for the specific details. For example, students in a history class examining fragmentary documents and researching an era or event use reflection to discover patterns in the evidence and construct a historical narrative. Learners need to develop this capability in order to use what they are learning in other contexts, to recognize the limitations or obstacles confronting them in a given situation, to take advantage of their prior knowledge and experience, and to strengthen their own performance.

6) An optional dimension: Creativity, originality, imagination:

As learners gain confidence and independence, knowledge and understanding, skills and strategies, ability to use prior and emerging experience in new situations, and reflectiveness, they generally become more playful and experimental, more creative in the expression of that learning. This is true not only in "creative" fields such as the arts, but in nearly all domains: research, argumentation, history, psychology, mathematics. In all fields the primary contributions to the field at the highest levels are the result of creative or imaginative work. Even in the early stages of learning in a discipline, exploration and experimentation, taking new or unexpected perspectives, and playfulness should be recognized and encouraged as a natural part of the learning process. This optional dimension may be adopted as part of the Learning Record by teachers or schools to make explicit the value of creativity, originality, and imagination in students' development and achievement. Among other things, it recognizes the value of creative experimentation even when the final result of the work may not succeed as the student may hope. If we hope to foster this quality in students' thinking and development, it is important to

encourage it, to document it, and to explicitly make it a value. We make this dimension optional because there are certain classes that depend on the transfer of information (as in human anatomy, for example) or the acquisition of fundamentally technical skills (calculus, for example) where creativity and imagination may not play a significant role.

UNIT-III

❖ CHARACTERISTICS OF INSTRUMENTS OF EVALUATION

It is important that instruments of measurement used for research purposes should be properly standardized. Individual score can be compared with the scores of others in a defined group. The evaluation tools serve a variety of uses. Selecting a proper tool is the base for accurate result of evaluation. An evaluation tool may be defined as a sophisticated means of intelligently and scientifically designed to evaluate what is required. Regardless of the type of tool used or how the results of evaluation are to be used, all types of evaluation should possess certain characteristics. The most important characteristics are validity, reliability, objectivity and usability.

1. Validity:

Validity is the quality of data gathering instrument which enables to measure what it is supposed to measure. Validity refers to the degree to which the test actually measures what it claims to measure. Validity is also the extent to which inferences, conclusions and decisions made on the basis of test scores are appropriate and meaningful. Validity also refers to whether or not a test measures what it intends to measure. A test with high validity has items closely linked to the test's intended focus. A test with poor validity does not measure the content and

competencies it ought to. Validity encompasses the entire experimental concept and establishes whether the results obtained meet all of the requirements of the scientific research method. A quality of a measurement indicating the degree to which the measure reflects the underlying construct, that is, whether it measures what it purports to measure.

a. Different Methods of Validity-

Sometimes validity is also thought as utility. Basic to validity of a tool is to measure the right thing or asking right questions. The items of a questionnaire, inventory must appropriately sample a significant aspect of the purpose of the investigation. Validity is not absolute characteristic; it depends on purpose and method used. The six categories of validity are **content validity, construct validity, criterion-related validity, concurrent validity, predictive validity and face validity.**

i. Content Validity-

Content validity refers to the connections between the test items and the subject related tasks. It is judged by the degree of relationship between diagnostics techniques and achievements in curriculum. The content validity of academic achievement test in subjects is examined by checking the test items against the complete courses of study. The test should evaluate only the outline of the content related to the field of study in a manner sufficiently representative, relevant, and comprehensible. **Based on the outline of the content indicating the kinds of knowledge and abilities which the students answer correctly.** The overall judgment is based on the extent of agreement between the test and the instructional plan.

ii. Construct Validity -

Construct validity is the relationship between the results of a **technique of measurement and other indicators of the characteristics that are measured.** It implies using the construct (concepts, ideas, and notions) in accordance to the state of the art in the field. Construct validity seeks agreement between updated subject-matter theories and the specific measuring components of the test. This type of validation is often used for measures of a psychological characteristic that is assumed to exist by

empirical or theoretical deduction. The general mental ability comprises independent factors such as verbal ability, number ability, perceptual ability, special ability, reasoning ability and memory ability. In order to establish the construct validity of a test, it may be necessary to **correlate the results of other tests.**

iii. Criterion-Related Validity-

It referred to as instrumental validity; it is used to demonstrate the accuracy of a measure or procedure by comparing it with another process or method which has been demonstrated to be valid. For example, imagine a **hands-on driving test** has been proved to be an accurate test of driving skills. A **written test** can be validated by using a criterion related strategy in which the **hands-on driving test is compared to it.**

iv. Concurrent Validity-

Concurrent validity refers to the usefulness of a test in closely relating to measures or scores on another test of known validity. **Tests are validated by comparing their results with a test of known validity.** Concurrent validity indicates the relationship between a measure and more or less immediate behavior or performance of identifiable groups. Concurrent validity is considered when any test is used for the purpose of distinguishing between two or more groups of individuals whose status at the time of testing is different. Concurrent validity is used for statistical methods of correlation to other measures. Once the tests have been scored, the relationship between the examinees' status and their performance (i.e., pass or fail) is estimated based on the test.

v. Predictive Validity -

Predictive validity refers to the usefulness of a test in **predicting some future performance.** Predictive validity is measured by the degree of relationship between a measured and subsequent criteria measure of judgments. This type of validity is used in tests of intelligence, test of aptitudes, vocational interest inventories and projective techniques. This type of validity is especially useful for test purposes such as selection or admissions.

vi. Face Validity-

Face validity is the characteristics which appear to **measure those which are actually sought to be measured**. It determined by a review of the items and not through the use of statistical analyses. Unlike content validity, face validity is not investigated through formal procedures. Instead, anyone who looks over the test, including examinees, may develop an informal opinion as to whether or not the test is measuring what it is supposed to measure. Face validity is not however suitable measure of validity, sometimes it might be misleading.

b. Factors Affecting the Validity-

The factors which influences the validity are

- i. Factors in the test itself
 - ii. Factors in test administration and scoring
 - iii. Factors of pupil response
 - iv. Nature of group
-
- i. **Factors in the Test Itself** Each test consists of number of items with close scrutiny. Tests the subject matter content only. Some factors lower the validity. **The unclear direction, complicated vocabulary, inappropriate level of difficulty, poorly constructed test items, misinterpretation, test is too short and improper arrangement of items.**
 - ii. **Factors in Test Administration and Scoring** The **test administration and scoring procedures may affect** the interpretation of the results. Teacher made test or standardized test are conducted during the adverse physical and psychological conditions, it may affect the validity.

- iii. Factors of Pupil Response **The economically disturbed students, lack of student's motivation and student's fear of test situation may ultimately affect the validity.**
- iv. Nature of Group Validity is always specific to a particular group to be measured. The nature of criterion used is **age, sex, ability level, educational and cultural background** influences the validity.

2. Reliability:

Reliability refers to the consistency of measurement, that is, **how consistent are evaluation results from one measurement to another.** Reliability is concerned with the extent to which an evaluation test is consistent in measuring what it is intended to measure. If the measurements are not consistent over different occasions or over different samples of the same performance domain, the evaluator can have little confidence in the results. **A test scores would be reliable when there is good reason to believe that it is stable and trust worthy.** These characteristics will depend on the extent to which the score is free from chance error. It is to be expected that the same test which is repeatedly administered on the same group of individuals, should yield the same pattern of scores.

Types of Reliability:

There are four general types of reliability.

- Inter-rater or Inter-observer Reliability : Measures the degree to which different observers give consistent estimates of the same persons.
- Test - Re-test Reliability : Measures the consistency of measurement on two separate occasions.
- Parallel-Forms Reliability : Measures the consistency of results of two parallel forms of same test constructed in the same way.

- Internal consistency Reliability : Measures the consistency of results across items within a test.

Different Methods of Reliability-

The various methods of estimating reliability are explained as follows-

i. **Test-Retest Method**

In this test, the **same tool or instrument is administered to the same sample on two different occasions**. The resulting test scores are correlated and the correlation coefficient provides a measure of stability over a given period of time. If the results are highly stable, those respondents who are high on one administration of test will also be high on the other administration and the other respondents tend to remain in their same relative positions on both administrations. An important factor to be kept in mind is the time interval between tests when interpreting measures of stability. If the time interval is short (say 1-2 days), the consistency of results will be inflated because respondents will remember some of their answers from the first test. If the time interval is quite long (say 1 year), the results will be influenced by the actual changes in the respondent over that period of time. Therefore, the best time interval between test administrations will mainly depend on the use to be made of results.

ii. **Equivalent-Forms Method**

This method uses **two versions of an instrument given to the same sample** of respondents. The two forms of the instrument are administered to the same group of respondents in close succession, and the resulting scores are correlated. The correlated coefficient provides a measure of equivalence. It indicates the degree to which both forms of the test are measuring the same aspects of behaviour. The equivalent forms method reflects short term constancy of respondents' performance and the extent to which the test represents an adequate sample of the characteristics being measured.

iii. Split-Half Method

Reliability is also estimated from a single administration of a single form of a test. The test is administered to a group of respondents in the usual manner and then is divided in halves for scoring purposes. To split the test into halves that are most equivalent, the usual procedure is to score the even numbered and the odd numbered items separately. This produces two scores for each respondent, which, when correlated, provide a measure of internal consistency. A reliability coefficient is determined by correlating the scores of two half-tests. The split half method is similar to the equivalent forms method in that it indicates the extent to which the sample of test items is a dependable sample of the content being measured. A high correlation between the scores on the two-halves of a test denotes the equivalence of the two-halves and consequently the adequacy of the sampling. The advantage of this method is that all data for calculation of the reliability coefficient can be collected in one sitting thereby avoiding variations due to two sessions.

iv. Kuder-Richardson Method

Another method of estimating the reliability of test scores from a single administration of a single form of a test is by means of **formulas developed by Kuder and Richardson. These formulas provide a measure of internal consistency** as with the split-half method but do not require splitting the test in halves for scoring purposes. Kuder-Richardson estimates of reliability provide information about the degree to which the items in the test measure similar characteristics. For a test with relatively homogeneous content, the reliability estimate generally will be similar to that provided by the split half method. In fact, Kuder-Richardson estimate can be thought of as an average of all of the possible split half coefficients for the group tested. It is an advantage when considering tests with relatively homogenous content since the estimate does not depend on the way in which the items are confined to the two half test as in the split-half method. However, for tests designed to measure more heterogeneous learning outcomes, the Kuder-

Richardson estimate will be smaller as compared to split half method and the later method is to be preferred.

Relationship between Validity and Reliability-

Validity and reliability are closely related. A test cannot be considered valid unless the measurements resulting from it are reliable. Likewise, results from a test can be reliable and not necessarily valid. Test validity is requisite to test reliability. If a test is not valid, then reliability is moot. In other words, if a test is not valid there is no point in discussing reliability because test validity is required before reliability can be considered in any meaningful way. Likewise, if a test is not reliable it is also not valid. Therefore, the two studies do not examine reliability.

At the same time, the evaluation results cannot be perfectly consistent. There are many factors that influence the results. If a single test is administered to the same group twice in a close succession, some variations in the scores can be expected because of temporary fluctuations in memory, attention, effort, fatigue and guessing etc. Variation in scores can occur due to intervening learning experiences if long time gap exists between two tests. Such extraneous factors introduce certain amount of measurement error in all types of evaluation. The method of determining reliability is, in fact, the means of determining the measurement error under different conditions. Methods of estimating reliability involve comparing at least two applications of the same instruments or equivalent instruments and determining the extent to which they agree. The closer the agreement, the greater is the reliability.

3. OBJECTIVITY-

Objectivity is a **central philosophical concept, related to reality and truth**, which has been variously defined by sources. Generally, objectivity means the **state or quality of being true even outside of a subject's individual biases, interpretations, feelings and imaginings**. A proposition is generally considered objectively true (to have objective truth) when its truth conditions are met and are "bias-free"; that is, existing without biases caused

by, feelings, ideas, etc. of a sentient subject. A second, broader meaning of the term refers to the ability in any context to judge fairly, without bias or external influence; this second meaning of objectivity is sometimes used synonymously with neutrality.

a. High objectivity tests-

Standardized group tests have high objectivity because they are provided with scoring keys. For example **intelligence, achievement, attitude and aptitude** tests. These tests have certain standard answers.

b. Moderate objectivity-

Binet and Wechsler Bellevue intelligence test administered individually have moderate objectivity. The evaluation techniques Rorschach and Thematic Appreciation test also have moderate objectivity. **The results obtained from test or other evaluative techniques require interpretation.** The raw scores on a standardized test becomes much more meaningful when compared with average score obtained by reference groups arranged according to age, grade, years of study and type of person.

4. USABILITY-

While selecting evaluation tool / instrument, **practical considerations** need to be kept in mind. Generally, the tests are administered by teachers having very limited training. The time available for testing is also limited. The cost of testing is also taken into consideration. All these factors must be taken into account when selecting evaluation tools.

a. Ease of Administration

Ease of administration is an important aspect if the evaluation instrument or test is to be administered by persons with limited training. For this, it is to be kept in mind that the **questions asked are few in numbers. The time provided should be sufficient and the directions should be simple and clear,** otherwise, persons who are not properly trained in administering tests

may err in giving directions as to how to fill the test. This can have an adverse effect on the validity and reliability of test score.

b. Appropriate Time for Administration

The **time provided for administering a test should be appropriate**. If an attempt is made to cut down too much on the time allotted to testing, the reliability of the test score may reduce drastically. A safe procedure is to allot as much time as is necessary to obtain valid and reliable results.

c. Ease of Interpretation of Results

The success of an evaluation is determined by the use made of the evaluation results. If they are interpreted correctly, they will contribute effectively in decision-making process. **If the results are misinterpreted, they will be of little value**. Therefore, ease of interpretation of evaluation results is important, when the results are to be presented.

d. Cost of Administering Tests

The cost of administering test should not be a major consideration as it is comparatively inexpensive. However, **in large-scale testing programs, use of separate answer sheets, machine scoring and reusable booklets can reduce the cost appreciably**.

5. NORMS-

A norm represents a typical level of performance for a particular group. A raw score on any Psychological test alone is meaningless unless we have additional interpretive data. So the score on psychological test are most commonly interpreted by reference to norms that represent the test performance of the standardised sample. Norms are empirically established by determining what persons in a representative group actually do on a test. In order to ascertain more precisely the individual's exact position with reference to the standardised sample, the raw score is converted into some relative measure. These derived scores serve two purposes.

1. They indicate the individual's relative standing in the normative sample and facilitate evaluation of performance.
2. They provide comparable measures that permit a direct comparison of the individual's performance on different tests.

Types of Norms :

Fundamentally, derived scores are expressed in one of two major ways (1) Development Norms and (2) Within group Norms.

- (1) **Developmental Norms** :- These type of norms generally indicate the normal developmental path the individual has progressed. They are very helpful for descriptive purpose but they are not compatible to precise statistical treatment. The types of development norms are (a) Mental Age Norms, (b) Grade Equivalent Norms and (c) Ordinal Scale Norms.
- (2) **Within Group Norms** :- Such type of norms help in comparing the individual's performance with the most nearly comparable standardised group's performance. Within group norms have a uniform and clearly defined quantitative meaning and can be appropriately employed in most types of statistical analyses.
 - a. **Percentiles (P(n) and PR)**:- Percentile scores represent the percentage of persons in the standardised sample who fall below a given raw score. They indicate an individual's relative position in the standardized sample. In case of percentiles, the counting begins from the bottom so lower the percentile, poorer the standing / rank.
 - b. **Standard Score** :- Standard score express the individual's distance from the Mean in terms of the standard deviation of the distribution. They are obtained by linear or nonlinear transformation of the original raw scores. T Scores and Z scores are known as standard scores.
 - c. **Age Norms** :- To establish age norms, the Mean of raw scores obtained by all in the same age group within a standardized sample is taken. So Mean raw score of 12 year old students would represent the 12 years norm.

d. Grade Norms :- Grade norms are found by computing the Mean row score obtained by students in particular grade.

a. Characteristics of a Norm:

- Its basic purpose is to measure student's achievement in curriculum based skills.
- It is prepared for a particular grade level.
- It is administered after instruction.
- It is used for forming homogeneous or heterogeneous class groups.
- It classifies achievement as above average, average or below average for given grade.
- It is generally reported in the form of Percentile Rank, Linear Standard Score, Normalized Standard Score and grade equivalent.

b. Merits of Norms

- To make differential predictions in aptitude testing.
- To get a reliable rank ordering of the pupils with respect to the achievement
- To identify the pupils who have mastered the essentials of the course more than others.
- To select the best of the applicants for a particular programme.
- To find out how effective a programme is in comparison to other possible programmes.

c. Demerits of Norms

- Test items answered by the students are not included in these test items because of their inadequate contribution to response variance.
- There is lack of congruence between what the test measures and what is stressed in a local curriculum.

- This promotes unhealthy competition and injurious to self-concepts of low scoring students.

Norm-referenced measurement is the traditional class based assignment. The measurement act relates to some norm, group or a typical performance. It is an attempt to interpret the test results in terms of performance of a certain group of students. So, this group is a norm group test scores. Thus norm-referenced test typically attempts to measure more general category of competencies.

UNIT-IV

4.6 OBSERVATION

In school, during the school hours, the students undergo various institutional environments. During this time the students behave differently. Observation is useful in evaluating students' behavior in different situations. In English we use the phrases to see and to observe. To see means to study the external features of the body whereas to observe. To see means to study the external body feature along with the internal features. This process of observation is simple as far as it is concerned with the gathering of information about the behavior and personality of the students.

Observation helps the observer to observe the activities of students, class-behaviour, and hence ascertain the emotional development, mental development and maturity etc. During observation care should be taken that the person who is observed is unaware that he is being observed. Thus the observation should be out of the knowledge of person. Further care should be taken that the behavior which is to be observed should be pre-decided. At a time one person and one characteristic should be observed. Observation can be direct or indirect, controlled or uncontrolled, known or unknown.

4.6.1 Types of Observation

Observations can be categorized as controlled observation and uncontrolled. The observation done in the laboratory is called controlled observation. Controlled observation means the observation is done with the knowledge of the person means that the person who is being observed is aware that he or she is being observed by the observer. Controlled observation thus means conscious observation. Uncontrolled observation means taking observation without the knowledge of person. Uncontrolled observation is thus conducted in a natural situation or condition.

4.6.2 Steps of Observation

Following are the steps for observation:

- Planning
- Execution
- Recording and interpretation

Planning

The characteristic, topic or the thing to be observed is decided in this stage. Whether it is group observation or personal observation, when and how many times the observation would be done, the tools useful for recording the observation etc. is also decided here. The specific type of training, if necessary for observer, is provided. These training are useful during interpretation. Who will be doing observation or the observer is also pre decided.

Execution

The arrangement for observation is done. The necessary arrangement for the observation as such the natural or artificial arrangement is done. Then after that environment of that opportunity is given so that the person is motivated to behave in some manner and that behavior is observed. Moreover the type of observation, its time and place for observation is also decided.

Recording and Interpretation

If the tools or instruments are ready then observation turns a fast process. The observation or the recording is hereby evaluated and interpreted.

4.7 CHECKLIST

It is one of the specific instruments for evaluation. Checklist is in the forms questionnaire. In this the answers of the questions are given checklist can be used for selfevaluation or for other's evaluation. It exhibits if the student has any particular characteristics or not and thus helps in the evaluation of the students.

4.7.1 Characteristics of Checklist

Checklist is used for evaluation of self and others. It is used as an instrument of observation. It involves questions and its answers. It involves signs by the respondent. It involves the characteristics about a particular subject to be evaluated.

4.7.2 Construction and Application of Checklist

The first horizontal line of the check list is used to write the name or number of the subject under observation. The characteristics of the subject or thing to be evaluated are arranged in vertical column of the evaluation sheet with the corresponding blank options to place the tick mark in the adjacent columns. Then the characteristics present in the subjects under observation are decided and if that characteristic is present in the subject then the tick mark is placed in that column. Then after the frequency of all tick mark is counted and marks are given to students on the bass of predefined norms or standards. Then the percentage, mean, median or correlation is used.

4.7.3 Uses of Checklist

1. It is useful for survey and research.
2. The amount of characteristics or traits of subjects can be known.
3. It is helpful to give the appropriate guideline to the subjects.
4. To know the developmental direction of the specific behavior pattern check list is used.
5. It is useful for self-evaluation and other's evaluation.

4.7.4 Limitations of Checklist

1. As only sign is used in checklist therefore no other options are found.
2. It is subjective and biased.
3. It is difficult to evaluate the personality of student or adjustment capacity through checklist.

4.8 RATING SCALE

By observing the various school and college activities we find change in behavior of students. Over and above that various personal characteristics are also observed. These characteristics separate the human behavior. The teacher observes such type of behavior of students by his insight and intelligence and hence evaluates the personality of the student. If this behavior of the students is evaluated through rating scale then it becomes more reliable. The technique of observation or the tool with the help of which the researcher or observer observes externally the amount of the various characteristics developed in a person and takes a note of it methodologically is called rating scale. Here the evaluation is done in relation to their opinion. Such a tool or instrument which converts the opinion into numbers is called rating scale. It can be used to evaluate the personality traits, creative skills, individual or social adjustment etc.,

4.8.1 Types of Rating Scales

The following are the main scales-

- Numerical Scales,
- Graphic Scale,
- Standard Scales,
- Check Lists,
- Forced Choice Scale,
- Ranking method and Q-Sort-method.

i. Numerical Scales

One of the simplest scales to construct and easiest to use, is the numerical rating scale. This type of tool usually consists of several items each of which names or describes the behavior to be rated, and then offers as alternative responses a series of numbers representing points along the scale. This simple numerical scale does have face validity and therefore seems to be widely accepted. It is more subjective or bias tool.

ii. Graphic Scale

If the format of the rating scale is such that the characteristics to be rated is represented as a straight line along which are placed some verbal guides, the tool is referred to as a graphic rating scale. It is easy to construct and easy to administer therefore it is widely used of all the specific types of rating scales, but it is less reliable measure.

iii. Standard Scale

In the standard scale approach an attempt is made to provide the rater with more than verbal cues in describe various scale points. Ideally, several samples of the objects to be rated are included each with a given scale value which have been determined in experimental studies prior to the use of the scale.

iv. Check Lists

An approach which is widely popular because it is simple to administer and still permits wide coverage in short time is the behavior check list. It contains a long list of specific behaviors which supposedly represented individual differences, and rater simply checks whether the item applies. The behavior index of individual is obtained by summing up the items, which have been checked. The modified check list or for reliable result, it is essential for each item as applicable or not applicable or not known.

v. Forced Choice Scale

One of the most recent innovations in the rating scale area has been developed a forced choice technique which has been designed to overcome the major difficulties faced on with earlier techniques. In a forced choice rating the rater is required to consider not just one attribute, but several characteristics all at one time. Assuming that relevant item is difficult for a better to distinguish from which is not predictive if both are equally favourable to the person, the format requires that only few of several behaviours listed in each item be selected as applicable.

vi. Ranking Method

It is not possible that rater can accurately judge equivalent distances at various points along the scale. Under these conditions a ranking method which requires only that subjects who are being rated to be placed in order of each trait can be used. This approach is essential for large number of persons are to be rated. The raking approach has the advantage of forcing the judge to make a definite discrimination among these rates by eliminating the subjective differences faced by the judges, second advantage that group ranking is uniform.

vii. Q Method

Another relative ranking method is so called Q-Sort developed by Stephenson 1953. It is one of the best approaches to obtain a comprehensive description of an individual while ranking method gives the comprehensive friction of a group of the individuals. Therefore Q-Sort is widely used for rating person's school or one the hob for individual guidance.

4.8.2 Importance of Rating Scale

- Any characteristic can be measured through rating scale.
- It is helpful to evaluate the behaviour which other tools can hardly deal with.
- Abstract characteristics can be evaluated by rating scales.
- It is helpful to personality or the social development of person.
- The level of each characteristic of each student of the class can be known.
- It is helpful to deliver all the necessary information related to the progress of students.
- The rating scale is also useful for the measurement of other methods or techniques.
- Within less time more opinions can be obtained.

4.8.3 Limitations of Rating Scale

- The evaluation being totally based on observation, the bias, liking, disliking, beliefs and assumptions etc., of the evaluator are the hindering factors for unbiased evaluation.
- The unawareness about the characteristics leads to the wrong observation.
- If large number of behavioral evaluation is to be done then the evaluator being bored of the tick mark generalizes the results.

4.9 INTERVIEW

—Interview is a purposeful conversation. - **John Darle**

Interview means a serious conversation which is done by some purpose (Goode and Hatt)ll. Interview means communication or conversation between two persons initiated by interviewer for collecting the information about research keeping in mind the objectives of the interview. Here the information is collected directly by verbal communication between two or more persons and the responses of the respondents are noted. It is a purposeful and serious conversation. The important aspect of interview is establishment of intimacy and to get response from respondent. Thus, the interview is a process of communication or interaction in which the

respondent delivers the required information to the interviewer face-to-face. It is used effectively to collect the useful information in many research situations.

When the researcher is extremely conscious about asking the questions in his presence to exhibit his personal interactive objectives, the researcher uses this process of questioning which is called interview. Here the information is collected from the people verbally with their physical presence. The responses of the respondent are then collected by the interviewer in a separate sheet. It can be conducted by the interviewer in person or in group. When the interviewer is conducted in group, the size of the group should not be so large that it inhibits participation of most of the members and at the same time it should not be so small that it lacks substantially greater convergence than in the individual interview. The optimum size is approximately 10-12 persons. Social, intellectual and educational homogeneity is important for effective participation of all group members. A circular seating arrangement, with the interviewer as one of the group, is conducive to full and spontaneous reporting and participation. The interview can be conducted one or more times as per requirement. As a tool for research interview is used as formal and informal, directional and non-directional interview.

4.9.1 Characteristics of Interview

- It is social interaction
- It is a sincere method
- It is direct purposeful conversation.
- It involves various direct involvements of interviewer and respondent.
- It involves various forms of questions to be asked to the respondent.
- It is a purposeful and serious conversation.
- It involves establishment of intimacy between the interviewer and respondent.
- It is a process of communication or interaction.
- It involves the note of responses delivered by the respondent.
- It involves the face-to-face involvement of the respondent and the interviewer.
- It can be conducted one or more times.
- It is a tool to collect the useful information in many research situations,
- It can be in person or in group.

- It exhibits the response excitement by the respondent.
- It is a behavioral method.
- It indicates social, intellectual and educational homogeneity.
- It may be formal, informal, directional or non-directional interview.

4.9.2 Nature of Interview

1. It is in the form of individual or group interview.
2. It is directional or non-directional interview.
3. It is formal or informal interview.

4.9.3 Types of Interview

Interviews are helpful as well resourceful in the research study especially in the personal study, case study, trend studies, historical research, experimental problems etc., to collect the relevant important information.

Following are the types of interview:

1. **Diagnostic Interview:** The problems related with education like adjustment, selfconcept, anxiety etc. can be known through diagnostic interview. It may be personal or in group. It consists of variety of questions necessary for the diagnosis. The respondent is questioned and thus diagnosis is done.
2. **Remedial Interview:** Such type of interview is conducted to resolve the diagnosed related problems for their remedy. It may be personal or in group. It consists of variety of questions necessary for the remedy. The respondent is questioned and thus remedial work is planned.
3. **Structured Interview or Controlled Interview:** In this type of interview the subject matter, questions and methods are pre-decided and fixed. The order of the questions and words are pre-decided and asked accordingly. Thus the content, method, number of question, words and the order of questions are fixed in this type of interview. The questions and answers are manipulated and controlled.
4. **Unstructured Interview:** In this type of interview the attitude, aspirations, beliefs and characteristics of respondent are used to collect the information. It is uncontrolled and flexible whereby pre-decided or pre-organized orders of questions are not emphasized. Here the interview is un manipulated and flexible. To get information about the attitudes, motivations,

characteristics and beliefs of the respondent the respondent is questioned in the way he feels comfortable. Unlike structured interview the respondent can give free responses.

5. Individual and Group Interview: Here as per the need the individual or group interviews are conducted. Individual interviews are conducted to evaluate the behavior, attitude or development of the individuals whereas group interviews are taken for specific or general problems etc.

6. Directional Interview: For fulfilling the needs related to interaction process, uncontrolled interviews are taken. As such subject matter and the area of checking are accurate and pre-organized and as it is used for getting related information they are considered as directional. Here interviewer is independent of form and order of question to be asked in interview means that the interviewer is free to ask any question of any form and in any order.

7. Non-Directional Interview: Non-directional interviews are included in psychoanalysis and in the field of medicine. There is freedom to show sympathy to the respondent by the interviewer. The respondents are motivated towards the subject matter. There is no pre-decided structure. In case of any doubt the answer of the respondent can be modified. It is also used to know the attitude, beliefs, ideas and feelings etc. In this type of interview free from of questions can be asked to find the solution of the problem. This type of interview being irrespective of order or sequence of questions, the interviewer should have talent of asking the appropriate questions. The respondent can be motivated about the subject of research.

8. Focused Interview: Focused interview focuses on events or occasions or the known situation of the respondent. Prior to the interview, the interviewer does the analysis and accordingly the questions are decided by during the interview, if it is required the pre-decided hypothesis can be rejected or can be changed. Interviewer interprets and evaluates the excitement of excited samples. Focused interview may be structured or un-structured.

9. In depth Interview: Through the in depth interview the amount of the experience and characteristics of respondent can be decided. Here exciting situations are motivated and concentrated over previous experience. To get the relevant and in depth information the intimacy between the interviewer and the respondent is established. Here the statements are repeated, misunderstanding is created, description of situation is done or situations are compared in different ways to get the in depth information from the researcher.

4.9.4 Steps of Interview

For interview the talent, patience and potential of interviewer, organization of interview and the intimacy between the respondent and the interviewer are of much importance. To conduct an interview, it should be carefully planned and designed whereas the interviewer, it should be skilled and able to develop intimacy with the respondent. The steps of interview include preparation of interview, execution of interview, note taking and analysis of the information. Following are the steps of interview:

- **Preparation for Interview:** Preparation for interview includes the objectives of interview and preparing the interview register. It is actually the mental preparation of the interviewer for the interview. It includes thinking for the objectives, type of interview, number of interviewer, position, place and time of interview etc. by the investigator.
- **Objectives of interview:** In this step, the general aims of research are converted into specific objectives. The area, information to be collected, the respondents and the type of interview is decided according to objective.
- **Prepare an Interview Register:** While preparing for the interview register the objectives of research are used to frame the questions. The research problem, related variables and the samples are considered. The information of good questions is based on subject matter, inspiration, realities, attitude, expectation of information and the intellect of interviewer and his rapport to develop relations. These questions can be objective or subjective, specific or general, fixed answer or free of giving any answer etc. Proper training, guidance and experience assure a good interview. It is a chain of appropriate questions and answers. The answers to the effective questions depend on the content, motivation, attitudes, expectation of information, time of interview and ability of the interviewer to establish the intimacy. The responses can be objective or subjective, special or general, free response or restricted.
- After the careful evaluation and critical thinking of the above aspects the appropriate types of questions are planned and a register is prepared whereby the investigator can use the appropriate type of questions. It can be in the form of questions, fill in the blanks, rating scale, checklist etc. The responses can be worked of accordingly.
- **Execution of Interview:** The execution of interview means conducting the interview. As per the preplan whether be it the personal or group interview, before starting the interview it is necessary to disclose the personal identity and the objectives and type of interview. The investigator should bear the tape recorder, camera, if necessary and the interview register. The

instructions, if any and necessary is delivered to the respondents. The execution of interview included establishing the rapport and eliciting information.

□ **Establishing Rapport:** To get the necessary, relevant, important and all the information related to the subject it is necessary to gain the confidence of respondent and thus leading towards a good and successful interview. It is necessary that the interviewer should be polite, well dressed, cool, calm, patient, decent and capable of questioning and must bear good understanding. The investigator should himself be clear with the questions and their responses and the objectives of interview. The investigator should be skillful, positive, joyous, unbiased, capable, and free of any rational and bear the attitude of sympathy thus establishing a good rapport with the respondents.

□ **Seeking the Information:** In pre-planned series asking appropriate questions without hurting the feelings of respondent and getting necessary and relevant information is important hence care should be taken that if in any case the respondent gets distracted of the point then those points should be flexible and the respondent is not bored of the interview and thus the information could be obtained.

□ **Note taking:** The final step of the interview with the respondent used a paper sheet, predesigned answer sheet, tape recorder or video recorder as per the requirement. Information is then minimized through analysis. To note the complete information from the respondent various activities, skill and talent could be used.

□ **Analysis of the collected Information:** In this step the investigator does the assessment of the respondent's view as per the pre-decided structure. Here the information provided by the respondent is analysed and transformed into specific group or class or category. Then with reference to the objectives of research the analysis and interpretation of the data is done.

ANECEDOTAL RECORDS :-

Anecdotal records are systematically kept notes of specific observations of student behaviors, skills, and attitudes in the classroom. Anecdotal records provide cumulative information regarding progress, skills acquired, and directions for further instruction. Anecdotal notes are often written as the result of ongoing observations during the lessons but may also be written in response to a product or performance the student has completed. Systematic collection of anecdotal records on a particular student provides excellent

information for evaluation of learning patterns and consistency of student progress. Well-kept anecdotal records provide a valuable, practical, and specific reference about a student. Anecdotal notes are used to record specific observations of individual student behaviours, skills and attitudes as they relate to the outcomes in the program of studies. Such notes provide cumulative information on student learning and direction for further instruction. Anecdotal notes are often written as the result of ongoing observations during the lessons but may also be written in response to a product or performance the student has completed. They are brief, objective and focused on specific outcomes. Notes taken during or immediately following an activity are generally the most accurate. Anecdotal notes for a particular student can be periodically shared with that student or be shared at the student's request. They can also be shared with students and parents at parent-teacher-student conferences.

The purpose of anecdotal notes is to: 1. provide information regarding a student's development over a period of time 2. provide ongoing records about individual instructional needs 3. capture observations of significant behaviours that might otherwise be lost 4. provide ongoing documentation of learning that may be shared with students, parents and teachers.