

UNIT 5: MAJOR TOOLS OF EVALUATION AND THEIR USES

Paper pencil tests, Oral tests, and Performance tests – Achievement tests : standardized and teacher made tests – Diagnostic tests – Intelligence tests and aptitude tests – Rating scale – Check list – Anecdotal records – Socio-metric technique – Interview, Questionnaire and Inventory – Use of test data: placement, promotion, grouping, diagnosis and remediation. – Self reporting techniques/ Reflection as assessment technique for learning. – Interview and focus group discussion

5.1 INTRODUCTION

The comprehensive programme of evaluation involves (i) **Objectives** (well-defined purposes of education), (ii) **Content/subject matter** (Curriculum, Syllabus, Courses and Text books), (iii) **Learning activities** (Methods and Learning Activities) and (iv) **Evaluation Procedures** (testing).

5.2 EVALUATION TECHNIQUES

Techniques of evaluation are classified into three major heads namely, (A) Quantitative, (B) Qualitative and (C) Quantitative and Qualitative.

(Please check Class work for diagram on Evaluation Techniques – (A) Quantitative, (B) Qualitative and (C) Quantitative and Qualitative)

(A) QUANTITATIVE TECHNIQUES

Quantitative Techniques are classified into (a) Achievement tests and (b) Diagnostic tests.

(a) ACHIEVEMENT TESTS

Any test that measures the accomplishments of an individual after a period of training or learning.

Achievement tests are further classified into (i) Teacher-Made Achievement Tests and (ii) Standardized Tests.

(i) TEACHER-MADE ACHIEVEMENT TESTS

Achievement tests which are prepared by the teachers for a particular content to measure the accomplishment that was taught/learnt during a particular period. This Teacher-Made

Achievement Tests are classified into (a) Oral tests, (b) Written/Paper-pencil tests and (c) Performance tests.

(a) ORAL TESTS: Oral test is a situation in which two persons (an examiner and an examinee) or an examiner and many examinees are involved in a face-to-face communication. In these tests, the examinee puts the question orally in order to get either the oral response or the written response or the oral performance. These Oral tests have advantages as well disadvantages.

In the pre-school, kindergarten and lower primary grades the oral examination is probably the only way to evaluate achievement because the pupils have not yet learned to read. In higher grades, oral examinations are often desirable when the teacher wishes to evaluate a pupil's ability to discuss broad types of problems in which integration of knowledge in several areas is important. These examinations also have some value as a diagnostic tool in certain situations.

Oral tests have some limitations namely, poor sampling of content, great consumption of time and low reliability. They certainly cannot by themselves provide a satisfactory basis for determining grades for a pupil's work in a given area of instruction.

The Oral tests are classified into three types namely, (a) Oral-Oral response tests, (b) Oral-Written response tests and (c) Oral-Performance Test.

In ***Oral-Oral response tests***, the examiner puts the question orally and expects the response from the examinee in oral form. Overall Flexibility for the examiner as well for the examinee is the major advantage. It is a time consuming test and Subjectivity dilutes the verbal interaction between examinee and examiner, hence these tests are least efficient tests.

In ***Oral-Written response tests***, the questions are presented by the examiner by the use of the spoken word and are answered in writing by the examinee. The two disadvantages are namely, (a) The style of writing comes as the criterion for judging the response and (b) There is a less opportunity to probe the in-depth knowledge. These type of tests can be easily conducted to several students at the same time is the advantage.

In ***Oral-Performance Tests***, the student is orally presented with a task to carry out, and his/her skill in executing it is assessed. Many instructional objectives which cannot be measured by Oral-Oral and Oral-written tests are amenable to measurement by oral-peformance examinations. Oral performance tests are particularly well suited to the languages and related

areas. Oral questioning is a good way to discover what thought processes a pupil uses in solving problems.

(b) WRITTEN/PAPER PENCIL TESTS: Written/Paper-pencil tests play a central role in evaluation of students' class room learning. In the early stages, writing is tested objectively. Essay writing should be tested at a higher stage. These tests aim at the assessment of not only graphic skills but also to some extent of audio-lingual skills. These tests are divided into two categories namely, Objective type tests and Subjective type tests. Objective type tests contain Recall and Recognition type of questions. They are Fill ups, Multiple Choice, True/False (includes Yes/No), Matching type, Classification type and one word answer questions (Analysis type). Subjective type tests (descriptive type tests) are further divided into Short answer type and Essay type tests.

(c) PERFORMANCE TESTS:

In Performance Tests, the student is presented with a task to carry out, and his/her skill in executing it is assessed. Many instructional objectives which cannot be measured by Oral and Written/Paper-pencil tests are amenable to measurement by Performance tests.

(ii) STANDARDIZED TESTS

The characteristics of Standardized tests are as follows:

- (i) These tests consist of items of high quality.
- (ii) The items are pretested and selected on the basis of difficulty value, discrimination power and possess relationship to the clearly defined objectives in behavioural terms.
- (iii) Exact time limit and scoring are precisely stated for its administration.
- (iv) Norms (based on age, grade and sex) are representative groups of individuals, are provided as an aid for interpreting the test scores.
- (v) The reliability and validity are well established.
- (vi) A manual is supplied that explains the purposes and uses of the test, describes briefly how it was constructed, provides specific directions for administering, scoring, and interpreting results, also contains tables of norms and summarizes available research data on the test.

(vii) No two standardized tests are exactly alike. Each test measures certain specific aspects of behaviour and serves a slightly different purpose.

(viii) (Uses of Standardized tests): Standardized tests are used for evaluating the general educational development of students in the basic skills and in those learning outcomes. They are also used for evaluating student progress during the school year or over a period of years and grouping students for instructional purposes. The purpose of diagnosing relative strengths and weaknesses of students in terms of broad subject or skill areas can also be achieved through the use of standardized tests.

Similarity between Standardized and Teacher-made tests: Carefully constructed Teacher-made Achievement Tests and Standardized Achievement Tests are similar in many ways. Both are constructed on the basis of carefully planned table of specifications, both have the same type of test items, and both provide clear directions to the students.

Differences between Standardized and Teacher-made tests: Standardized tests are useful when the purpose of testing is to compare a student's performance in different content areas; to find the status of an individual student, class, or school system in a wider population; to compare classes or schools among themselves; and to measure growth over a period of years. On the other hand, teacher-made tests are used to determine whether specific curriculum goals have been achieved, giving grades and comparing them with other students. They differ in the quality of test items, the reliability of test measures, the procedures for administering and scoring and the interpretation of scores.

Comparison of Teacher Made Achievement Tests Vs. Standardize Achievement Tests

A comparison of Teacher-made Achievement tests and Standardized Achievement tests is as follows:

POINT OF COMPARISON	TEACHER-MADE ACHIEVEMENT TEST	STANDARDIZED ACHIEVEMENT TEST
Learning outcomes and content measured	They are used to evaluate the outcomes and content of what has been taught in the classroom.	They are used to evaluate outcomes and content that have been determined irrespective of what has been taught
Quality of test	Quality of test items unknown and	General quality of items is high.

items	is generally lower than items of standardized tests.	They are pre-tested and selected on the basis of difficulty and discrimination power.
Reliability	Reliability is usually unknown but it can be high if test items are carefully constructed.	Reliability is usually high.
Administration and scoring	Uniform procedure of administration and scoring may be possible. It is usually flexible.	The procedure of its administration is standardized and specific instructions for its administration and scoring are provided.
Interpretation of scores	Scores can be compared and interpreted only in the context of the local school situation.	Scores can be compared to norm groups, Test manuals and other guides for interpretation and use.

(b) DIAGNOSTIC TESTS

Diagnostic tests measure the strengths and weaknesses in the learning of the students. In order to make the teaching-learning process effective, it is essential **to** find the learning difficulties of students during instruction. The diagnostic tests consist of items based on a detailed analysis of the specific skills involved in successful performance and a study of the most common errors made by students. The test difficulty is low. A good diagnostic test will permit a student to demonstrate all aspects of the skill being measured and will pinpoint the types of errors that he/she has made.

Advantages: (i) Diagnostic tests are available for different subjects. (ii) While selecting a test, the diagnostic procedures should be evaluated in the light of the specific type of information desired. (iii) Diagnostic tests are designed for students of below average performance in the particular subject(s). If a student scores high on a sub-test, it simply means that this student does not have difficulties related to the area measured by the sub-test.

Disadvantages: (i) Diagnostic tests are useful in identifying weaknesses in learning but not for indicating the level of proficiency. (ii) Diagnostic tests indicate the typical errors a student makes, but they do not indicate the causes of the errors. Some causes can be easily inferred from the type of error made, or from a student's explanation of how he/she arrived at the answer. (iii) Diagnostic tests provide only partial information for diagnosing a student's difficulty. Other methods have to be made use of to supplement and complement this information. (iv) Results from diagnostic tests, concerning specific learning difficulties, tend to have a low reliability because there are relatively few items measuring each type of error. Therefore, the findings regarding specific strengths and weaknesses for any particular pupil should be regarded as clues to be verified by other objective evidence and by regular classroom observation.

(B) QUALITATIVE TECHNIQUES

Qualitative Techniques are classified into (a) Observational Techniques, (b) Socio-metric Techniques, (c) Self-Report Techniques and (d) Projective Techniques.

(a) OBSERVATIONAL TECHNIQUES

Classroom teachers continually observe their students. The observation is typically informal and unsystematic, carried on without any benefit of specific planned procedures. It should be as accurate and reliable as possible.

Observational Technique is the process of observing and recording an individual's behaviour. It implies the use of (any of or all) particular observational tool namely, (i) Checklist, (ii) Rating Scale and (iii) Anecdotal Records.

(i) CHECK LIST

A check list consists of a listing of steps, activities or behaviour which the observer records when an incident occurs. It enables the observer to note only whether or not a trait or characteristic is present. It does not permit the observer to rate the quality of a particular behavior or its frequency of occurrence or the extent to which a particular characteristic is present.

Advantages of check list: The following are the advantages of Check lists: (i) They are adaptable to most subject-matter areas, (ii) They are useful in evaluating those learning activities that involve a product, process and some aspects of personal-social adjustment, (iii) They are most useful for evaluating those processes that can be sub-divided into a series of clear, distinct,

separate actions, (iv) When properly prepared, they constrain the observer to direct his attention to clearly specified traits or characteristics, (v) They allow inter-individual comparisons to be made on a common set of traits or characteristics, (vi) They provide a simple method to record observations, (vii) They objectively evaluate traits or characteristics, (viii) They are objectively used for evaluating interest, attitudes and values of the learner and (ix) They may be used for evaluating teaching interest of student-teachers in the colleges of education.

Conditions for Constructing a Check List for True/False items: (i) Express each item in clear, simple language, (ii) Avoid lifting statements verbatim from the text, (iii) Avoid negative statements wherever possible, (iv) Make sure that each item is clearly true or False and (v) Review the items independently.

(ii) RATING SCALE

Rating Scales resemble check lists but are used when finer discriminations are required. Instead of merely indicating the presence or absence of a trait or characteristic, it enables us to indicate the degree to which a trait is present and provide systematic procedures for obtaining, recording and reporting the observer's judgements.

Advantages of Rating Scales: The advantages of Rating Scales are as follows: (i) They can be used with a large number of students, (ii) They tend to be very adaptable and flexible, (iii) They can be efficient and economical in the use of a teacher's time, (iv) They can be comprehensive in the amount of information recorded, (v) They can help to reduce the subjectivity and unreliability that are usually associated with observation methods, (vi) They have much wider range of application and can be used for teacher ratings, personality ratings, school appraisal, sociological survey and such other areas and (vii) Rating scales can be used with any observer, since they have very little training for the purpose.

Limitation of Rating Scales: Rating scales may be affected by several limitations, particularly different types of errors. They are discussed as under:

(1) Error of Leniency: There is a constant tendency among the observers to rate those whom they know well or in whom they are closely involved, higher than they should. Such observers are called easy observers. Some observers become aware of the feeling of easy rating and consequently rate individuals lower than they should. Such observers are called hard observers. The leniency error refers to a general and consistent tendency for an observer to rate too high or

too low for whatever reasons.

(2) Error of Central Tendency: Most of the observers hesitate to rate the individuals on the extremes of the scale, instead they tend to rate the individual on the middle of the scale.

Obviously, the results get distorted.

(3) Halo-Effect: It is an error which obscures the clusters of traits within an individual. The observers form a general opinion about the person's merit and his ratings on specific traits and are greatly influenced by this general impression. It results in a spurious positive correlation among the traits which are rated.

(4) The Logical Error: It is due to the fact that judges are likely to give similar ratings for traits which they feel logically related to each other.

(5) The Contrast Error: It is due to a tendency of a rater to rate others in the opposite Direction (contrasting) from himself in a trait.

(6) The Proximity Error: It has been seen that adjacent traits on a rating tend to inter-correlate higher than remote ones, their degree of actual similarity being approximately equal. This error may be counteracted to some extent by placing similar traits further apart and the different ones close together.

Rating Scales are in general three types namely, (a) Numerical Rating Scale, (b) Graphic Rating Scale and (c) Descriptive Graphic Rating Scale

(a) Numerical Rating Scale: Numerical scale is a sequence of definite numbers supplied to the observer. The observer assigns, to each stimulus to be rated, an appropriate number in line with those descriptions.

A nine-point Rating Scale consists of 9- Most pleasant, 8- Extremely pleasant, 7- Moderately pleasant, 6- Mildly pleasant, 5- Indifferent, 4- Mildly unpleasant, 3- Moderately unpleasant, 2- Extremely unpleasant, 1- Most unpleasant

Instead of a nine-point scale, one can have three-point, five-point or seven-point scales too. In a three-point scale we can have, 3- Most pleasant, 2- Indifferent, 1- Most unpleasant.

Conditions for Construction of a Rating Scale: (i) It is always useful to have an odd number of points in a scale like 3, 5 and 7, so that there could be a middle one, a favourable one and an unfavourable one. (ii) The use of negative number is not favoured as those observers who are not

well versed in arithmetic addition and they find it difficult to manage negative numbers.

Advantages: Numerical rating scales are the easiest to construct and to apply. They are also the simplest in terms of handling the results.

Disadvantages: Numerical scales are often rejected in favour of other types of scales because it is believed that they suffer from various biases and errors.

(b) Graphic Rating Scale: The Graphic Rating Scale is the most popular and the most widely used type of Rating Scales. In this scale, a straight line is shown vertically or horizontally with various cues to help the observer. The line is either segmented into units or is continuous. If the line is segmented, the number of segments can vary from case to case. For example: (Question (i): Did the speech show good organisation? Response: 1-Very poor/2-Average/3-Very Good)

Advantages: Graphic Rating Scales are simple and easy to administer. Such scales are interesting to the observer and require little added motivation.

Disadvantages: Scoring in the case of some formats of graphic scales is rather laborious.

(c) Descriptive Graphic Rating Scale: Descriptive Graphic Rating Scale are described aspects in each point, for example, (Question (i): While preparing a blackboard summary, how was the penmanship? Response: Legible/Normally readable/Illegible)

Legible = beautiful, uniform size and slant	Normally readable = good-looking, fluent motion	Illegible = bad-looking, tends to draw outlines
---	---	---

Such specific descriptions contribute to a greater objectivity of the rating process. The description also helps to clarify and further define a particular dimension.

(ii) ANECDOTAL RECORD

Anecdotal records is a record of some significant item of conduct, a record of an episode in the life of students, a word picture of the student in action, a word snapshot at the moment of the incident, any narration of events in which may be significant about his/her personality.

An Anecdotal Record is an observation that is written like a short story. They are descriptions of incidents or events that are important to the person observing. They are short, objective and as accurate as possible. It records the behaviour of the students time to time through an informal

device used by the teacher. It provides a lasting record of behavior which may be useful later in contributing to a judgment about a student.

Characteristics of Anecdotal Record: Anecdotal Records possess the following characteristics:

(i) Anecdotal Records contain factual descriptions of what happened, when it happened and under what circumstances that the behavior was occurred. (ii) The interpretations and recommended actions should be noted separately from the description. (iii) Each Anecdotal Record should contain a record of a single incident. (iv) The incidents recorded should be considered to be significant to the student's growth and development. (v) Anecdotal Records should contain Accurate, Specific and Simple reports of behavior, Results of direct observation, Gives context of child's behavior and Typical Records of unusual behaviours.

Objectives/ Purpose of Anecdotal Record: (i) To furnish the multiplicity of evidences needed for good cumulative record, (ii) To substitute for vague generalizations about students' specific exact description of behaviour, (iii) To stimulate teachers to look for pertinent information in helping each student to realize good self- adjustment, (iv) To understand individual's basic personality pattern and his reactions in different situations, (v) To understand the pupil in a realistic manner, (vi) To provide an opportunity for healthy pupil- teacher relationship, (vii) To maintain the areas of behavior that cannot be evaluated by other systematic method, (viii) To help the students to improve their behavior and analyze his/her behavior better, (ix) To use by students for self-appraisal and peer assessment.

Advantages of Anecdotal Records: The advantages of Anecdotal Records are as follows: (i) Anecdotal Records record unusual events, children's behavior, skills and interests for the purpose of planning, (ii) It records how an individual is progressing in a specific area of development, (iii) It provides a means of communication between the members of the health care team and facilitates coordinated planning and continuity of care, (iv) It acts as a medium for data exchange between the health care team, (v) It provides a reliable, permanent record of patient care, since Anecdotal Records are clear, complete, accurate and factual documentations, (vi) It Supplements and validates of other structured instruments, (vii) It provides insight into total behavioral incidents, (viii) It needs no special training and uses formative feedback, (ix) It is economical

and easy to develop, (x) It is Open ended and can catch any unexpected events in the life of the student, (xi) It can select behaviors' or events of interest and ignore others, or can sample a wide range of behaviors' (different times, environments and people).

Disadvantages of Anecdotal Records: (i) If Anecdotal Records are carelessly recorded, the purpose will not be fulfilled, (ii) Quality of the record depends on the memory of the observer, (iii) Incidents can be taken out of context may give hasty opinions, (iv) Subjectivity of the observer distracts the behaviour of the student, (v) Lacks the process of standardization, (vi) Its scoring is difficult, (vii) It is time consuming, (viii) It may miss recording specific types of behaviours, (ix) It results with limited application.

Guidelines for Making Anecdotal Record: (i) Keep a notebook handy to make brief notes to remind you of incidents you wish to include in the record. Also include the name, time and setting in your notes. (ii) Write the record as soon as possible after the event. The longer you leave it to write your anecdotal record, the more subjectivity and vague the observation will become. (iii) In your anecdotal record identify the time, child, date and setting. (iv) Describe the actions and what was said. (v) Include the responses of other people if they relate to the action. (vi) Describe the event in the sequence that it occurred. (vii) Record should be complete in all nature. (viii) They should be compiled and filed. (ix) They should be emphasized as an educational resource. (x) The teacher should have practice and training in making observations and writing records.

Items in Anecdotal Records: To relate the incident correctly for drawing inferences the following items to be incorporated: (a) The first part of an anecdotal record should be factual, simple and clear: Name of the student, Date of Event and time....., Brief report of what happened..... (b) The second part of an anecdotal record may include additional comments, analysis and conclusions based on interpretations and judgments.

Example: Anecdotal Record of Mr. Rajesh

Anecdotal Record of Mr. Rajesh

<p style="text-align: center;"><i>December 10, 10.30 a.m.</i></p> <p style="text-align: center;"><i>Recess</i></p> <p>1. Rajesh is seen in the Library, sitting in a corner, solving riders in geometry.</p>
<p style="text-align: center;"><i>December 15, 1.15 p.m.</i></p> <p style="text-align: center;"><i>Geometry Period</i></p> <p>2. Rajesh is found to be only boy responding to the teacher's questions solving a rider</p>
<p style="text-align: center;"><i>December 21, 4.45 p.m.</i></p> <p style="text-align: center;"><i>Language Period</i></p> <p>3. Rajesh was caught reading a book “Fun in Geometry “during the language period and was asked to leave the room.</p>
<p><i>Interpretation:</i> Rajesh is very much interested in Geometry</p> <p><i>Recommendation:</i> Rajesh should be encouraged to do more in mathematics, and also not to neglect languages.</p>

Example for Anecdotal Recording Form

<p>George Washington Elementary School</p> <p>Anecdotal Recording Form</p>	
Observer: _____	Observation Date: _____
	Observation Time: _____
Student Name: _____	
Description of the incident:	

Description of the location/setting:	

Notes/Recommendations/Actions:	

Signature _____	

(b) SOCIO-METRIC TECHNIQUES

The word Sociometry comes from Latin word *Socius* and *Metrum* meaning Partner/Friend/Companion and Measure. Jacob Levy Moreno coined the term Sociometry and conducted the first long-range Sociometric study from 1932-38 at the New York State Training School for Girls in Hudson, New York.

Moreno defines Sociometry as the mathematical study of psychological properties of populations, the experimental technique of and the results obtained by application of quantitative methods.

Sociometry is a quantitative method for measuring social relationships, *the measurement of Social Interaction*. It is a way of measuring the degree of relatedness among people. It is the study of interpersonal relationships between people in a group. It is the inquiry into the evolution and organization of groups and the position of individuals within them. It is a method used for the discovery and manipulation of social configurations by measuring the attractions and repulsions between individuals in a group. It is based on the fact that people make choices in interpersonal relationships. Whenever people gather, they make choices-- where to sit or stand; choices about who is perceived as friendly and who not, who is central to the group, who is rejected, who is isolated.

Objectives/ Purpose of Sociometry:

- (i) To facilitate group task effectiveness and satisfaction of participants by bringing about greater degrees of mutuality amongst people and greater authenticity in relationships.
- (ii) To measure the relatedness that can be used not only in the assessment of behavior within groups, but also for interventions to bring about positive change and for determining the extent of change.
- (iii) To reduce the conflict and improve communication because it allows the group to see itself objectively and to analyze its own dynamics.
- (iv) To assess dynamics and development in groups devoted to therapy or training.
- (v) To study the patterns of how individuals associate with each other when acting as a group toward a specified end or goal.

Types of Sociometry

Sociometry has two main branches: Research Sociometry, and Applied Sociometry.

Research Sociometry is action research with groups exploring the socio-emotional networks of relationships using specified criteria.

Example: (i) Who in this group do you want to sit beside you at work? (ii) Who in the group do you go to for advice on a work problem? (iii) Who in the group do you see providing satisfying leadership in the pending project?

Applied Sociometry utilizes a range of methods to assist people and groups review, expand and develop their existing psycho-social networks of relationships.

Both fields of Sociometry exist to produce through their application, greater spontaneity and creativity of both individuals and groups.

Sociometric Techniques are as follows: (i) Sociogram, (ii) Sociometric Matrices, (iii) Guess-Who Technique and (iv) Social Distance Scale

Uses of Sociometr: The uses of Sociometry are as follows: (i) Sociometric measurements can be used to study the type of interaction and factors affecting the interaction. (ii) It is also used to measure group characteristics such as, cohesiveness, morale and bonding. (iii) It can also be used to measure individual characteristics such as leadership, aggressiveness, liberalism and conservatism. (iv) Any personality factor that is related to interpersonal attraction can be measured through Sociometry. (v) Sociometry provides a bird's eye view of group structure and at the same time indicates the degree of social attraction possessed by the members of the group.

(c) **SELF-REPORT TECHNIQUES**

Self-report techniques require the respondent to react to items concerning his/her own behavior or characteristics. The items generally require expression regarding likes, dislikes, fears, hopes, religious beliefs, ideas about sex and many other matters that reflect the way which the person copes with his own needs and demand of his environment. The Self-report techniques are viewed through three techniques namely, (i) Interview, (ii) Questionnaire and (iii) Inventory.

(i) **INTERVIEW**

The interview is a face-to-face (Interviewer and Interviewee) personal conference, in which the required information is obtained directly from an individual. There are *at least three main aspects to the technique* of interviewing and they are described as follows: **(1) Parts of an interview:** (a) A *beginning* (establishing rapport and clarifying the purposes of the interview), (b) A *middle* (data gathering or exploration of alternative solutions to a problem) and

(c) *An end* (summarizing and terminating the face-to-face relationship). (2) **Securing Responses:**

(a) *Structured or controlled interview:* (Here, there is an attempt to cover a definite ground.

Schedules of questions are generally drawn up beforehand, (b) *Unstructured or Uncontrolled interview:* (Here, conditions are least controlled. The interviewer raises various questions orally with the interviewee and takes advantage of the flexibility and informality of the situation). (3)

Recording Results: The teacher takes brief notes during the course of the interview and summarizes the main points after the interviewee leaves. With the consent of the interviewee, a tape recorder may be used.

Advantages of Interview: (i) Interview provides flexibility. It allows the interviewee to ask for a clarification of a question and the interviewer to raise all kinds of follow-up questions to the answers of the respondent. The interviewer can skim over certain areas and probe others intensively. (ii) It lends itself to dealing with confidential and personal material which cannot be obtained through a questionnaire. (iii) It gives the interviewer the opportunity to hear how an interviewee has said something (for example, accent, fluency, tone of voice, meaning) as well as what he/she has said. (iv) An informal interview on how he/she solves an arithmetic problem may reveal to the teacher what techniques of thinking he/she employs. (v) Quietly chatting with a student at his seat, the teacher can gather valuable information about him. (vi) The interview is a face-to-face situation. Here two people enter into a structural conversation. To this extent the interview is a familiar, real-life situation. (vii) It enables each side to have a look at the other. No employer would like to take an employee without seeing him, not an employee would like to join a firm without meeting somebody from the firm. (viii) The interview is a method acceptable to both sides. An unacceptable selection method (for example, tests of intelligence) could cause considerable initial tension which might impair performance. (ix) The skillfully-handled interview provides a highly flexible situation. Different approaches can be tried. (x) The interview is relatively quick and economical.

Limitations of Interview: (i) There is always a chance that time will be wasted, unnecessary questions raised, irrelevant materials covered, or that constructive person-to-person relationship cannot be sustained because of the personality of either the interviewer or the interviewee. (ii) It is time-consuming. (iii) It makes it possible for an interviewer to ride a personal hobby. (iv) Its utility depends on the willingness of the interviewee to give honest answers. (v) It cannot be used in group situations. (vi) There is lack of definite aim or purpose in the interview. (vii) The

candidate is not allowed to develop a theme. Instead of that there are too many questions which can be answered in one word. (viii) Asking questions which only elicit information already possessed by the interviewer through application forms and confidential reports. This only shows that the interviewer has no clear idea of procedure. (ix) Asking the questions in such a tone that the candidate is led to believe that the question must be answered in a certain way. (x) Restricting the time so much that there is no real chance of making maximum use of the interview possibilities. (xi) Sometimes the interviewee is unduly influenced by the views and techniques of the other interviewers. (xii) Failure by interviewers to discuss techniques of interviewing among themselves.

(ii) QUESTIONNAIRE

Questionnaire refers to a device for securing answers to a question by using a form which the respondent fills in by himself. It is a popular tool of collecting all kinds of data. It is widely used to obtain data about certain conditions and practices and to enquire into opinions and attitudes of an individual or a group.

Advantages of Questionnaire:

- ***Economical:*** The questionnaire requires paper, printing and postage only. There is no need in it to visit the respondents personally or continue the study over a long period. Accordingly, in comparison with other methods, the mailed questionnaire is by far the cheapest and quickest research method.
- ***Time saving:*** Besides saving money, the mailed questionnaire also saves time. Simultaneously, hundreds of persons are approached through it whereas if they are to be interviewed it may take very long time.
- ***Most reliable in special cases:*** As noted above in certain problems a mailed questionnaire is not only suitable but preferred techniques of research.
- ***Research in wide area:*** When we have to make research about people living at great distances, the mailed questionnaire can be the only means of research if a researcher can undertake to travel the whole world; but be mailed questionnaire, people living in all countries can be approached.
- ***Suitable in special type of response:*** The information about certain problems can be best obtained through questionnaire method.

Limitations of Questionnaire:

- ❖ ***Limited response:*** As noted above this method cannot be used either on illiterate or semi-literate persons. Furthermore, very preoccupied, travelling and aristocrat individuals do not reply. Therefore, the number of persons who cooperate and respond to the questionnaire is very small. The percentage of such persons is extremely low in the underdeveloped countries.
- ❖ ***Lack of personal contact:*** There is no provision in this method for coming face to face with the respondents. However much a researcher may try to simplify and rationalize his questionnaire, it is not possible to avoid each and every technical term.
- ❖ ***Useless in Depth-problems:*** If a problem requires deep and long study, it is obvious, it cannot be studied by the questionnaire method.
- ❖ ***Possibility of wrong answers:*** A respondent may not really understand a question or may give the answer in a casual manner. In both cases there is a strong likelihood of misleading information being given.
- ❖ ***Illegibility:*** Some persons write so badly that even they themselves find it difficult to read their own hand.
- ❖ ***Incomplete Response:*** Some persons give answers which are so brief that the full meaning is incomprehensible.

(iii) INVENTORY

An Inventory is constructed in the form of a questionnaire. It consists of a series of questions or statements to which the subjects respond by answering ***Yes*** or ***No***, ***Agree*** or ***Disagree*** or in some similar way to indicate preferences, or to make those items that describe his typical behaviour. Personality measures are obtained by means of inventories and questionnaires. (Similarity between Inventory and Questionnaire)

In the inventory, the statements are put in the first person. For example, "I think I am more tensed than others". In the questionnaire, there is a question in a second person. For example 'Do you think you are more tensed than other persons around you?' (Dissimilarity between Inventory and Questionnaire)

Uses of inventories

- Inventories are more exhaustive than questionnaires.
- Inventories have been mostly used for measuring personality traits, interests, values and adjustment that is for assessing self-reporting affective behaviour.

Guidelines for administration of the Inventory are: (i) The teacher should explain printed instructions very clearly to the learners. (ii) The teacher should make it clear to the learners that the data will be kept confidential. (iii) The teacher should remove doubts if any, regarding the manner of filling the questionnaire/inventory. (iv) The teacher should take all other timely precautions for preparing a state of mind conducive to response.

Some important inventories are given below:

1. The Ascendance-Submission Reaction Inventory
2. Bernreuter Personality Inventory by Bernreuter
3. Bell's Adjustment Inventory by Bell
4. Minnesota Multiphasic Personality Inventory by Starke R. Hathaway and J.Charnley Mckinely
5. Comrey Personality Scales by Comrey
6. California Psychological Inventory
7. Self-analysis Form by R.B.Cattell

(iv) PROJECTIVE TECHNIQUES

Projective Techniques refers to the techniques used to reveal (to project) the inner world of repressed feelings, wishes, hopes, fears and ambitions of an individual by going deep in to his/her unconscious. It is called Projective techniques because it projects the inner world of unconscious.

In Projective Techniques indefinite and unstructured stimuli are presented to the subjects and asked them to structure those stimuli or explain them. While structuring or explaining the stimuli the subjects project the repressed feelings, wishes, hopes, fears and ambitions from the unconscious.

- Some common Projective Techniques are (i) The Rorschach Inkblot Test, (ii) Thematic Apperception Test, (iii) Children Apperception Test, (iv) Word Association Test and (v) Sentence Completion Test.

The Rorschach Inkblot Test was developed by Hermann Rorschach, a Swiss psychiatrist. It consists of ten cards containing inkblots. Five of them are in black and white and five of them are multi-coloured. These are unstructured and do not have any specific meaning. It can be administered on subjects and responses can be scored and interpreted.

Thematic Apperception Test was first introduced by Henry Murray (1943) to measure the need for achievement. Later it was developed for the assessment of personality with the help of C.D.Morgan. It consists of thirty vague and indefinite pictures portraying human beings in a variety of actual life situations. Ten are meant for males, ten for females and ten are common to both sexes. The test is administered in two sessions using ten pictures in each session. The pictures are presented one at a time. The subject has to make up a story for each of the pictures within a fixed period of time. The following aspects are to be considered while the subject is making up the story. (i) What is going on in the picture? (ii) What has lead to this scene? (iii) What would likely happen in such a situation?

□

Children Apperception Test: Thematic Apperception Test is not suitable for children. So Dr. Leopold Bellak developed this test for children between three and ten years old. It consists of ten cards having pictures of animals. These are meant for both genders. Children are asked to make up stories as a game.

Word Association Test: Word Association Test includes a number of selected words. The examiner speaks a series of words, one word at a time. The subject should immediately say the first word that comes to his mind. There is no right or wrong answers. Uttering of unusual words, if any and behaviour manifestations are considered for evaluation.

□

Sentence Completion Test consists of a list of incomplete sentences, generally open ended. These require completion by the subject in one or more words. The subject has to go through the list of words and answer as quickly as possible

Item Examples: • I am worried over.....

• My hope is

- I feel proud when
- My hero is

Indian Adaptations of Personality Tests are (i) Indian adaptations of senior apperception test developed by Uma Chaudhary, (ii) Indian adaptations of thematic apperception test by Uma Chaudhary, (iii) Indian adaptations of children's apperception test by Uma Chaudhary, (iv) A pragmatic view of Rorschach inkblot technique by B.L Dubey, (v) Sentence Completion test by L.N Dubey and Archana Dubey and (vi) Mosaic test of personality by B.B Chatterjee.

(C) QUANTITATIVE AND QUALITATIVE TECHNIQUES

Quantitative and Qualitative Techniques of Evaluation are mainly Psychological tests for (i) Intelligence and (ii) Aptitude. Other psychological tests are considered in the earlier sections of this same chapter.

(i) INTELLIGENCE TESTS

Intelligence is the ability to understand world, think rationally or logically, and use of resources effectively when faced with challenges of problems. Intelligence includes the ability demanded in the solution of problems which require the comprehension and use of symbols.

Tests of Intelligence or General Menial Ability measure general ability which enters into performance of all activities and which differs in magnitude from individual to individual. The items in such tests assess the subject's ability to perceive relationships, solve problems and apply knowledge in a variety of ways.

Intelligence tests are classified as

- verbal and non-verbal tests, paper-pencil and performance tests,
- speed and power tests
- Individual and group tests.

Verbal and Non-Verbal; Paper-Pencil and Performance Test: Verbal or paper-Pencil tests require the subject to respond to the item by writing his/her replies. The posed questions are presented in the form of sentences or words and requires the subject to record his/her answers either by underlying, ticking or encircling one of the alternative answers which are provided, or by writing a word, phrase or sentence or sentences in a blank space provided for the purpose.

In the Non-verbal test, problems are posed in the form of designs and require the subject to record his answer either by underlying, ticking or encircling one of the alternative answers which are provided.

On the other hand in the performance test, problems are presented in a concrete form and the subject **is** required to respond not by writing but by manipulating blocks or picture cards etc., depending on the level of the test.

Speed vs. Power Tests: A speed test is defined as one in which no subject has time to attempt all items. Such tests limit the time in which a subject has to complete certain tests. On the other hand 'power test' is a test in which every subject has a chance to attempt each item of the test. It has no time limit and **the** subject goes on attempting test items until he is unable to continue successfully any more.

Individual vs. Group Tests: The tests which are administered on one individual at a time are known as individual tests. These tests are useful in situations where a precise and detailed assessment of some characteristics of an individual is desired.

The test which is administered to many subjects at the same time is termed as 'group test'. These tests are particularly useful where a large number of subjects have to be tested at the same time.

Assessment of Intelligence can be done through Stanford-Binet Intelligence Scale, Wechsler Test, Uzgiris Hurt Scale, Bhatia's Battery Test, Cattell's Culture-Fair Tests, Koh's Block Design Test, Leiter International Performance Scale and Miller Analogies Test.

Advantages of Intelligence Tests:

- Intelligence tests measure a wide variety of human behaviors.
- Allow professionals to have a uniform way of comparing a person's performance with that of other people who are similar in age.
- Also provide information on cultural and biological differences among people.
- Excellent predictors of academic achievement and provide an outline of a person's mental strengths and weaknesses.

Disadvantages of Intelligence Test:

- Many intelligence tests produce a single intelligence score.
- This single score is often inadequate in explaining the multidimensional aspects of intelligence.
- Individuals with similar intelligence test scores can vary greatly in their expression of these talents. For example, strong verbal skills vs. strong skills in perceiving and organizing various tasks may lead two people to have identical scores on intelligence tests.

(ii) APTITUDE TEST

Aptitude is a natural ability or propensity. It is the capacity of person to achieve along specific ways. It makes the special abilities and achievements in special area.

Aptitude test attempts to determine and measure a person's ability to acquire, through future training, some specific set of skills (intellectual, motor, and so on). The tests assume that people differ in their special abilities and that these differences can be useful in predicting future achievements. Components of Aptitude Test are Verbal Reasoning, Abstract Reasoning, Numeric Reasoning, Spatial Reasoning, Mechanical Reasoning, Data Checking and Work Sampling.

Though intelligence tests seek to measure general mental abilities which are valuable in almost any type of thinking, yet effective educational and vocational guidance and appropriate placement of students call for tests specially directed at specialized abilities. Such types of tests are called **tests** of special abilities, traits or aptitudes. These traits are indicative of the future success of an individual in a particular field. Therefore, these aptitude tests are used for guidance, as well as prediction of success in some occupation. Training or academic courses are possible on the basis of scores on a standardized aptitude test.

Assessment of Aptitude tests can be done through Differential Aptitude Test Battery (DATB), General Aptitude Test Battery (GATB), Mechanical Aptitude Test Battery (MATB) and Mechanical Assembly Test.

Advantages of Aptitude tests: (i) They are excellent predictors of future scholastic achievement, (ii) They provide ways of comparing a child's performance with that of other children in the same situation, (iii) They provide a profile of strengths and weaknesses, (iv) They assess

differences among individuals, (v) They have uncovered hidden talents in some children, thus improving their educational opportunities, (vi) They are valuable tools for working with handicapped children, (vii) Group aptitude tests---quick and inexpensive and (viii) Aptitude tests are valuable in making program and curricula decisions.

Difference between Aptitude Test and Intelligence Test

Intelligence Test	Aptitude Test
IQ denotes intelligence as a single, measurable trait.	Aptitude breaks that intelligence down into several different characteristics.
It refers more to a broad range of mental abilities.	Aptitude score reflects specialized abilities.
It is not a precise tool for career guidance.	Aptitude tests is a precise tool for career guidance.
A high IQ score, by itself, would not indicate whether a person is strong or weak.	Provide strengths and weakness of person.

5.3 USES OF TEST DATA

Placement: The test scores are used for determining grading, promotion, placement in the same school and in other institutions.

Promotion: Many factors enter into the important decision of moving a student into the next grade. Intuition is an important part of any decision but that intuition is enhanced when coupled with data. Standardized tests and records of classroom performance tests are essential for supplying much of the data upon which these decisions are based.

When a pupil transfers from one school to another for admission in the same class or in a higher class, a decision must be reached as to the grade level and group into which he is to be placed.

Grouping: If the teacher has specified instructional objectives, testing can yield information that

will aid the teacher in assigning specific students to instructional groups. The teacher can change the groups later after more teaching and testing has taken place.

Diagnosis: No one source of data can be sufficient to assess what a pupil knows about school-related content. What is called for is a triangulation of several kinds of data drawn from various types of tests: standardized tests of achievement and aptitude, teacher-made quizzes, observations of behavior and the like. Diagnosis does not necessarily mean prescription unless the data collected have demonstrated high reliability and validity.

Remediation: The test scores will help the students can know their strengths and weakness in respective subjects. It provides feedback to the students. It also provides a basis for checking the adequacy of their own progress in a particular subject, as well as their study habits, interest, home influence etc. Based on the test scores, teachers may infer about the success of instruction process adopted by them. Also they may provide more appropriate instructional guidance for individual students or the class as a whole.

5.4 SELF-REPORTING TECHNIQUES/ REFLECTION AS ASSESSMENT

TECHNIQUE FOR LEARNING–INTERVIEW

(Self-Reporting Techniques were dealt earlier in this Section)

The assessment interview is very similar to a regular job interview but focuses more directly on the person's personality. Assessment interview techniques allow an organization to quantitatively determine if an applicant has the necessary skills and knowledge to complete the tasks required for a particular role. This technique is also used in educational settings as well, for certification and promotion purposes by using the Interview Assessment Form.

The Interview Assessment Form incorporates a scoring grid to enable the researcher to assess each candidate as fairly and objectively as possible. It involves five hierarchical steps as follows:

Step 1: List the competencies and criteria from the Person Specification on the interview assessment form and apply a weighting according to how important you consider them to be in fulfilling the duties of the post. The weighting scale is from 1 to 3, with 3 being most important and 1 being less important.

Step 2: Carry out the interview, using questions and, if appropriate, exercises to gain evidence to enable you to make a judgement as to how far the candidate meets each competency and/or criteria. Make notes during the interview.

Step 3: At the end of each interview, agree as a panel how far you consider the candidate met each competency or criteria and apply a rating as follows: *Zero rating for evidence demonstrated that the candidate did not meet the competency or criteria; One rating for evidence demonstrated that the candidate met the competency or criteria; Two for evidence demonstrated that the candidate exceeded the competency / criteria.*

Record the rating on the Interview Assessment form.

Step 4: Multiply the weighting for each competency or criteria by the rating you have applied and record on the interview assessment form. Add together to arrive at a total score and record on the form. The candidate you appoint should therefore be the person who achieves the highest score.

Step 5: Record any training or support needs you have identified.

5.6 FOCUS GROUP DISCUSSION

Focus groups are called as a data collection method. The data is obtained through a semi-structured interview process.

Focus groups are moderated by a group leader. Focus group methods emerged in the 1940s with the work of Merton and Fiske who used focus groups to conduct audience studies.

Characteristics of Focus Groups

The Characteristics of Focus Group researches are as follows:

- (i) Standardization of questions: Focus groups can vary in the extent to which they follow a structured protocol or permit discussion to emerge.
- (ii) Stratification in Focus Groups: Sampling in focus Group research depend on the different stratifications (e.g. age, sex, socioeconomic status, health status) that the researcher identifies as important to the research topic
- (iii) Number of participants per group: The rule of thumb has been 6-10 homogeneous strangers.
- (iv) Level of moderator involvement: Level can vary from high to low degree of control exercised during focus groups (e.g. extent to which structured questions are asked and group

dynamics are actively managed)

When Focus Groups might be used:

Focus groups might be used (i) To explore new research areas, (ii) To explore a topic that is difficult to observe, (iii) To explore a topic that does not lend itself to observational techniques (e.g. decision-making), (iv) To explore sensitive topics, (v) When you want to collect a concentrated set of observations in a short time span, (vi) To ascertain perspectives and experiences from certain people on a topic, (vii) In combination with other methods, (viii) To gather preliminary data, (ix) In the development of surveys and interview guides and (x) To clarify research findings from another method.

Recording Focus Group data

- (i) One of the challenges in recording focus group data knows who is speaking at any particular time, since often multiple people speak in overlap. Consider audio- or video-recording Focus Group sessions (or even both). Video will be helpful for identifying who is speaking. Recordings also provide access to tones of the discussion and the ability to replay sessions during analysis.
- (ii) Transcribe focus group discussions.
- (iii) Have a least 2-3 researchers (in addition to the moderator) attend the focus group and take notes. The focus of each researcher's note-taking efforts might be different (e.g. nonverbal behaviour, group dynamics, emergent themes).
- (iv) Note taking is important to capture nonverbal data. Even if one is video-recording a group, some nonverbal behaviour will be lost that might be recorded by a note-taker.

Advantages of Focus Groups

The advantages of Focus Groups are as follows: (i) Ability to produce a large amount of data on a topic in a short time, (ii) Access to topics that might be otherwise unobservable, (iii) Can ensure that data directly targets researcher's topic, (iv) Provide access to comparisons that focus group participants make between their experiences, (v) This can be very valuable and provide access to consensus/diversity of experiences on a topic, (vi) Captures real life data in a social environment, (vii) Flexible, (viii) High face validity, (ix) Speedy results, (x) Economical and (xi) The Researcher can increase sample size for qualitative research.

Disadvantages of Focus Groups

The Disadvantages of Focus Groups are as follows: (i) Facilitator has less control, (ii) Data more difficult to analyse, (iii) Special skills required, (iv) Nature of group varies, (vii) Groups may be difficult to assemble, (viii) Venue must be conducive.