UNIT 7 TYPES OF TOOLS

Structure

7.1 INTRODUCTION

In Unit 6, we have talked about some essential characteristics of a good tool.

In this unit we are trying to discuss a few tools of evaluation such as rating scales, intelligence and aptitude tests, inventories, anecdotal record, teacher made tools and standardized tests of achievement. Each tool and test has its own special role, its own uses and limitations.

Rating scale for instance is useful to classify opinions and judgements regarding situations, objects etc. Intelligence and aptitude tests on the other hand are used as tools to measure in numerical terms the potential performance and special ability of a person. Inventories are used to assess the expression of the inner feelings of individuals through questionnaires and personality inventories. Anecdotal records are used to assess the behaviour in the past of a person in a particular situation. The important events are recorded by the teacher. It helps to know the specific quality of a person. Teacher made and standardised tools are used to measure in numerical terms the attainment of students in various school subjects.

In the following sections and sub-sections we will discuss the definition, use and limitations of the above mentioned tools and tests.

7.2 OBJECTIVES

After going through this unit you should be able to :

- define rating scales and describe their types, uses and limitations,
- define anecdotal records, their uses and the process of its development,
- define an intelligence test and describe the types, uses and limitations of intelligence tests,
- define an aptitude test and describe its type, uses and limitations,
- differentiate between intelligence and aptitude tests,
- define inventories/questionnaires and describe their types and uses,

- define and differentiate between teacher made and standardized tests of achievement, and
- describe steps for standardising achievement tests.

7.3 RATING SCALES

It is a subjective method. Through this, one can find out what others say about a particular person. Rating is a term applied to an expression of opinion or judgement regarding some situation, object, character or an attribute. Rating scales refer to a scale with a set of points which describe varying degrees of the dimension of an attribute being observed.

Rating scales are broadly classified into five categories :

1) numerical scales;
2) graphic scales;
3) standard scales;
4) rating by cumulative points; and
5) forced choice ratings.

We discuss them below in the same order.

7.3.1 Numerical Scale

In a typical numerical scale a sequence of definite numbers is supplied to the rater or to the observer. The observer assigns, to each stimulus to be rated, an appropriate number in line with those definitions or descriptions. For example, the following scales may be used in obtaining ratings of the affective values of colours on a 9 point scale which may consist of different points.

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 9 point scale one can have 3 point, 5 point or 7 point scales too. In a 3 point scale we can have:

3. Most pleasant
2. Indifferent
1. Most unpleasant

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.

The use of negative number is not favoured as those observers who are not well versed in algebra, find it difficult to manage negative numbers; while in thinking one does think in terms of +ve, –ve, above average, below average, etc.

Numerical rating scales are the easiest to construct and to apply. They are also the simplest in terms of handling the results. However, numerical scales are often rejected in favour of other types of scales because it is believed that they suffer from various biases and errors.

7.3.2 Graphic Scale

The graphic 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 clues to help the rater. 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. An example of such scales is given below:

How effective was the teacher in the class

1	2	3	4	5
Very effective	Slightly effective	Average	Slightly ineffective	Very ineffective

There are many advantages in graphic scales. They are simple and easy to administer. Such scales are interesting to the rater and require little added motivation. However, scoring in the case of some formats of graphic scales is rather laborious.

7.3.3 Rating by Cumulated Points

The unique and distinctive feature of rating by cumulated points is its immense and easy utility of scoring. The rating score for an attribute, object or individual is the sum or average of the weighted or unweighted points. The 'check list method' and the 'guess - who technique' belong to this category of rating. Check list methods are applicable in the evaluation of the performance of personnel in a job. The weights of +1 and -1 are assigned to every favourable and unfavourable trait, characteristic or attribute respectively and the individual's score is the algebraic sum of the weights. In the 'guess-who technique', some statements in terms of some description like "here is one who is always doing bad things to make others sad", are constructed and each individual is asked to list all the members of his group who fit such descriptions mentioning an individual as many times as necessary. Each individual scores a point for each favourable or unfavourable description applied to him, and the total score is the sum total of all his points.

7.3.4 Standard Scale

In standard scale a set of standards is presented to the rater. The standards are usually objects of the same kind to be rated with pre-established scale values. This type is like that of the scales for judging the quality of hand-writing. The scales of hand-writing provide several standard specimens that have previously been spread over on a common scale by some standardised procedure like equal-appearing intervals. With the help of these standard specimens, a new sample of hand-writing can be equated to one of the standards. The 'man to man' scale and the 'portrait matching' scale are the other two forms that conform more or less to the principles of standard scales.

7.3.5 Uses of Rating Scales

i) Rating methods are quite interesting to the raters, especially if graphic methods are used;

ii) Best rating can be obtained by presenting one stimulus to a rater at a time;

iii) Rating scales can be used with ratees who have very little training for the purpose;

iv) Rating scales can be used with a large numbers of stimuli;

v) Rating scales have much wider range of application and can be used for teacher ratings, personality ratings, school appraisal, sociological survey, etc.

7.3.6 Limitation of Rating Scales

Rating scales may be affected by several limitations. Some of them are discussed as under :

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

2) **Error of Central Tendency:** Most of the raters 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 raters 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.

Besides the above limitations of Rating Scale, they may suffer from the following errors as well:

- Placing of ratee
- Proximity of ratee

For example, a ratee who by chance happens to be placed immediately after ratees who have been marked on the high side, would be placed lower than where he should be, as the rater has formed a high standard in his judgement. Therefore this ratee suffers by comparison and the reverse would happen if the same ratee happended to be placed immediately after a few ratees who have been marked low or most unfavourably.

Check Your Progress 1

What are the four types of Rating Scales?

.....

.....

.....

.....

7.4 INTELLIGENCE TESTS

Tests of Intelligence or General Mental 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 subjects' 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 and individual and group tests.

7.4.1 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 replies. The posed questions are presented in the form of sentences or words and require the subject to record his 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.

7.4.2 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.

7.4.3 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.

Check Your Progress 2

i) Differentiate between Verbal and Non-Verbal tests of intelligence.

.....

.....

ii) How do Speed Tests differ form Power Tests?

.....

.....

7.5 APTITUDE TESTS

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 stduents call for tests specially directed at specialised 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 standardised aptitude test.

The important aptitude tests are:

1) Differential Aptitude Test Battery (DATB)
2) General Aptitude Test Battery (GATB)
3) Mechanical Aptitude Test Battery (MATB)
4) Mechanical Assembly Test

Check Your Progress 3

Differentiate between Intelligence and Aptitude Tests.

.....

.....

.....

.....

7.6 INVENTORIES

Personality measures are obtained by means of inventories and questionnaries. Both these terms are usually used interchangeably, but in fact it is not so.

In general, the word questionnaire referes 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.

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.

In the inventory, the statements are put in the first person. For example, "I think I am more tense than other". In the questionnaire, there is a question in a second person. For example 'Do you think you are more tense than other persons around you?'

Inventories are more exhaustive than questionnaires.

Inventories have been mostly used for measuring personality traits, interests, values and adjustment i.e., for assessing self-reporting affective behaviour.

The term inventory is also used in an other sense, i.e., inventory of various items available in an office room or inventory of clothes to be given to the washerman for ironing, etc. But in the educational evaluation in the classroom, inventories have been used for the assessment of interests, values, adjustment and personality traits.

Guidelines for administration of the inventory are :

1) The teacher should explain printed instructions very clearly to the learners.
2) The teacher should make it clear to the learners that the data will be kept confidential.
3) The teacher should remove doubts if any, regarding the manner of filling the questionnaire/inventory.
4) The teacher should take all other timely precautions for preparing a state of mind conducive to response.

Check Your Progress 4

Define an Inventory.

7.7 THE ANECDOTAL RECORD

Every teacher, (also parent, friend, relative etc.) observes his pupils from day to day and those observations can be formally recorded as a guide to his own understanding or to that of others who will later deal with the pupils. Such reports of a child by informal observations recorded in the form of anecdotes (not only of present, but past too, i.e. childhood nursery and pre-school days etc.) by parents, siblings, friends and peer group have been called Anecdotal Records (AR).

These records provide a medium for the observations of significant examples of one's behaviour. A record of events is one dependable datum that will remain unchanged from the time it is made until the time we want to refer to it. A set of such records provides stable evidence on which later appraisals can be based. Making a record of an observation of a child's behaviour, a prompt record, while the behaviour is still fresh in the mind, is dependable as it eliminates the limitations and distortions of memory. Such a record can be used, with practice, to provide a relatively direct and objective report of actions, with reactions of some observers as well. These records serve two purposes. The first purpose may be to give teachers practice in studying children with a view to deepening their understanding and increasing their sympathetic insight. Second purpose is that it provides an informal and largely qualitative picture of certain aspects of an individual's behaviour. It is a primary

aspect of social functioning or adjustment to personal problems that one hopes to illuminate by records of incidents of school behaviour. The interactions of a child with the other children in the room, aggression or withdrawal, events that throw light on the child's role in the group and his reactions to it are good materials for these records. Indications of personel tensions and adaptation to them, habitual mood and temper or special crises and adjustments are worth recording.

Anecdotal Record should be an accurate factual report of an event in a child's life reported with enough detail so that it is a meaningful evidence of behaviour.

A good Anecdotal Record has the following features:

1) It provides an accurate description of specific events.
2) It describes the setting sufficiently to give the events meaning.
3) If it includes interpretation or evaluation by the recorder, this interpretation is separated from the description and its different status is clearly identified.
4) The event it describes is one that relates to the child's personal development or social interactions.
5) The event it describes is either representative or the typical behaviour of the child or significant because it is strikingly different from his usual form of behaviour. If it is unusual behaviour for the child that fact is to be noted.

Maintenance of Anecdotal Records

Each original Anecdotal Record is an item of information about an individual. A series of records provides whole sets of such items. But for data to be useful they must be organised, summarised and interpreted. At intervals the anecdotal record should be brought out. An attempt should be made to relate the anecdotal materials to other facts that are known about the child i.e., his health, intellectual ability, academic achievement, home surroundings and family pattern. A tentative interpretation of the pattern may be attempted, if it is recognised that any interpretation is to be thought of as a set of very tentative hypotheses. Observations should be a factual report of significant items of behaviour. They should avoid evaluation, interpretation and vague generalities. Recording of observations should be kept as simple as possible and the record should be viewed periodically to give an organised picture of the person who has been observed.

It is one of the important tools to record a single specific, but significant event or the behaviour of a learner observed by the teacher. It is a systematic and significant record of learner's behaviour that generally forms a part of the permanent record passed on from grade to grade. In studying a behaviour problem, it is important to look for signs of its development because change can be evaluated through knowledge of previous events.

In Anecdotal Record, an attempt is made to produce a factual report of one spontaneous behaviour of a learner. This written description of a specific incident may be very useful later on in helping to understand the learner. These records also open up a possibility of describing events that are perhaps exceptional but potentially significant.

Significant Guidelines to Use Anecdotal Record

1) The teacher should try to describe the incident as objectively as possible. He should not write his personal comments in the description column.
2) If a teacher is to write his comments or interpretation of incidents, he should be very objective.
3) Description of the incident and the teacher's interpretation of comments should be reported separately.
4) The teacher should record these descriptions as soon as the event has occured. Simplicity and accuracy in recording should be maintained.
5) Conclusions should not be based on a single incident/event. A series of related anecdotal records are necessary for objective, reliable and valid conclusions. In other words adequate sampling is neccessary for a valid interpretation of behaviour. We can draw conclusions regarding temperament, anxiety and other behaviour of learners through it.

A sample Anecdotal Record is shown below:

Cumulative Anecdotal Record

Date : Time :

Name of the Pupil _____

Sl. No.	Description of Anecdote	Name of the Observer	Comments if any
1.			
2.			
3.			
4.			
5.			

Check Your Progress 5

i) Define Anecdotal Record.

.....

.....

.....

.....

ii) Write five features of Anecdotal Record.

.....

.....

.....

.....

7.8 TEACHER MADE AND STANDARDIZED TOOLS

Teachers construct various tools for the assessment of various traits of their students. The most commonly used tools constructed by a teacher are the achievement tests. The achievement tests are constructed as per the requirements of a particular class and subject area they teach. Besides achievement tests, for the assessment of the traits, a teacher observes his students in a classroom, playground and during other co-curricular activities in the school. The social and emotional behaviour is also observed by the teacher. All these traits are assessed. For this purpose too tools like rating scales are constructed.

Tools used by a teacher may be both non-standardized and standardized. A standardized tool is one which has got systematically developed norms for a population. It is one in which the procedure, apparatus and scoring have been fixed so that precisely the same test can be given at different time and place as long as it pertains to a similar type of population. The standardized tests are used in order to :

1) compare achievements of different skills in different areas;

2) make comparisons between different classes and schools;

- They have norms for the particular population. They are norm-referenced.

On the other hand, teachers make tests as per the requirements of a particular class and the subject area they teach. Hence they are purposive and criterion-referenced. They want :

- to assess how well students have mastered a unit of instruction;
- to determine the extent to which objectives have been achieved;
- to determine basis for assigning course marks and find out how effective their teaching has been.

You have already learnt about Achievement tests in detail in Unit 4 and you will remember that a good tool has the following characteristics:

1) Validity;
2) Reliability;
3) Comprehensiveness;
4) Practicability;
5) Discriminating Power;
6) Acceptability;
7) Motivating students to put up their best performance.

Teacher made tools can be divided into :

1) Oral;
2) Written;
3) Practical;

If you want to standardize your own tool for the assessment of students in a large population, you must follow the following steps :

1) Planning;
2) Preparation;
3) Try out;
4) Preparation of Norms;
5) Manual, i.e. instructions for administering a tool, scoring of a tools and interpretation of data.

You have already read all these steps for standardization of a tool in detail in Chapter 4.

Check Your Progress 6

i) What is the need of constructing Achievement Tests, for a teacher?

.....

.....

.....

.....

.....

.....

ii) Differentiate between teacher made tools and standardized tool?

.....

.....

.....

.....

iii) Write the steps for constructing and standardizing a tool.

.....

.....

.....

.....

7.9 LET US SUM UP

In this unit, an attempt has been made to make you aware about the importance of evaluation in education and how it is to be done in school situations. Various tools like Rating Scale for classifying opinions and judgements on a Scale are discussed in detail alongwith its various types. Where to use what type of Scale is explained for your practical use.

The importance of the use of Intelligence and Aptitude tests has been highlightened. The type of tests and its use must be very clear to you. Similarly, Inventories, Anecdotal Records as reports of informal observation and Achievement Tests play a very important role in the evaluation of students. We are sure this brief picture of various tools will prove very helpful to you in your institutions.

7.10 UNIT-END EXERCISES

1) What are the various tools of evaluation?

.....

.....

.....

.....

2) Differentiate between a Rating Scale and an Anecdotal Record.

.....

.....

.....

.....

3) Mention two uses of Intelligence Tests and Aptitude Tests.

.....

.....

.....

.....

7.11 ANSWERS TO CHECK YOUR PROGRESS

1) Numerical, Graphic, Standard, Rating by Cumulative points.

2) i) In verbal test, the subject responds to items by writing the replies or ticking, underlying or encircling the replies given in words while in Non-verbal test replies/responses are presented in designs or diagrams.

   ii) In speed test, time limit is important while in power test, time limit is not important. Subject has a chance to attempt each item.

3) Intelligence is a general mental ability where as Aptitude is a special mental ability.

4) Inventory consists of a series of questions or statements to which a subject responds by answering 'Yes' or 'No'; 'agree' or 'disagree' or in some similar way to indicate preferences.

5) i) The formal record of a teacher's observation of pupils behaviour is Anecdotal Record.

   ii) Accurate description of event, sufficient setting, interpretation of the event, relation with personal development and representative typical behaviour.

6) i) There is a need of constructing Achievement Tests for a teacher in order to assess and compare various traits of their students which will help the teacher to diagnose the difficulties of their students and to guide them accordingly.

ii) Teacher made tools are used as per the requirements of a particular class and subject areas they teach. Hence they are purposive and criterion-referenced. On the other hand a standardized tool is systematically developed for a population with norms for the same.

iii) The steps for constructing and standardizing the tool are – Planning, Preparation, Try-out, Preparation of Norms and Manual.

7.12 SUGGESTED READINGS

Anastasi, Anne, (1960), *Psychological Testing*. The Macmillan company, New York.

Ebel, E.L. (1965), *Measuring Educational Achievement*, Englewood Cliffs, New Jersey.

Frederick, Davis B.(1981), *Educational Measurement and their interpretation,* Woodworth Publishing Company Inc. Belmont, california.

Gronlund E. Norman, (1970), *Reading in Measurement and Evaluation*. The Macmillan Company, Collier, Macmillan Ltd, London.

Remmers, N.H. and et. all, (1965), *"Measurement and Evaluation"*, Harper and Row, New York.

Thorndike. R.L., (1970), *"Measurement and Evaluation in Psychology and Education"*. Wiley Eastern Pvt. Ltd., New Delhi.

Jaiprakash and Tiwari, (1974), *"Measurement, Evaluation and Examination"* (in Hindi). Kala Tample Publishers, Agra.