Department of Open & Distance Learning

Punjabi University, Patiala

Class : M.A. I (Education) **Semester : 2**

Paper : III (Methodology of Educational Research-II) **Unit : I**

Medium : English

Lesson No.

1.1 : Sampling : Meaning, Purpose, Importance and Various Sampling Techniques

1.2 : Tools of Educational Research - Tests, Questionnaires and Rating Scales

1.3 : Tools of Educational Research : Interview, Observation and Inventories

1.4 : Reliability and Validity

1.5 : Methods of Educational Research: Historical, Descriptive

1.6 : Methods of Educational Research : Experimental, Research Proposal and Research Report

M.A. (EDUCATION) PART-I **PAPER-III**
(Semester-II) **Methodology of Educational Research-II**

LESSON NO : 1.1 **AUTHOR : Y.P. AGGARWAL**

Structure of the Lesson :

1.1.1 Objectives :

After carefully studying this chapter students will be able to :

(i) define the terms population and sample.

(ii) describe various methods of sampling : non-probability sampling and probability sampling.

(iii) explain the steps involved in the process of sampling.

(iv) describe various sampling errors.

1.1.2 Introduction

Very often our attitudes, our knowledge and our actions are based on a sample. It applies equally to everyday life and to scientific research.

The trader examining a handful of grains from the bag, the tea-taster trying different brands of tea, technician making a blood-test and a house-wife checking a few grains of rice from the panto determine whether the entire lot has been properly boiled and becomes soft are all employing the method of sampling. Their confidence in their judgments rests on the fact that the material they are sampling is so 'well mixed or homogeneous, that the few grains of wheat, a drop of blood, a few leaves of tea or a few grains of rice do adequately represent the whole. No research is possible if we think of including the entire population in our study unless the particular population to be studied is limited and confined to just one or two places. Hence the need of a sample.

1.1.3 Meaning of Sampling

A sample is a miniature picture of the entire group of aggregate, from which it has been taken. A sample, in other words, is a small representation of a whole. The entire group from which the sample has been taken is known as the "Population". The terms "Universe" and "supply" are also used but are less popular. The term population in research is used in a broader sense than its common place meaning as a population. A population may consist of persons, objects, attributes, qualities, behaviour of people, answers to various items of test, the behaviour of inanimate objects such as throws of dice or cities, families, opinions of the electorate of a nation and the like. A population is well-defined group of any of these. The definition should be explicit enough to permit anyone to say with confidence that a particular object or person of the defined population and other is not. Defining a population means fixing the limit in terms of one or more of its various aspects as geographical age or grade, sex, physical attributes, etc.

The selection of one or more of these criteria is defining a population depends on the nature and scope of research in hand. Population may be classified as follows :

(i) Finite population : When all the elements of the population can be counted and listed e.g., the population of high school in a state.

(ii) Infinite population : When all the elements of the population cannot be counted and listed e.g. population of stars in the sky, trees in jungle, fish in

the sea.

A distinction can be made between the following :

(i) Target population : The total group of subjects about whom the researcher is empirically attempting to learn something.

(ii) Experimentally Accessible Population : The population of subjects that is actually available to researcher for drawing a sample studying the whole of it.

Kempthorne has distinguished between experimentally accessible, population and 'target population'. The former is the population of subject that is available to the researcher for his study. The target population is that total group of subjects about whom the researcher is empirically attempting to learn something.

In case where the design of research allows for the inclusion of the whole population if it was accessible and also within manageable limits, sampling may not be resorted to and the study conducted on the entire group. However, in studies involving large population sampling provides a less expensive, more efficient and quicker method of data collection.

1.1.4 Purpose and Importance of Sampling :

Cochran has pointed out the following advantages that accrue from using a sample rather than the entire population.

(i) Reduced Cost : When data are collected only from a small fraction of the entire population, expenditure is smaller then when the entire group was studied. Surveys include something as small a number as one person in one thousand.

(ii) Greater Speed : The volume of the data to be collected will be smaller. Hence it can be collected tabulated and summarized more quickly with a sample than with the total population. In applied research where urgent answers to certain problems are needed, this aspect gets an added importance.

(iii) Greater Scope : In studies, where a complete enumeration and census of all units of population are impracticable and the research requires the use of highly trained personnel or specialized equipment, the choice may lie between collecting the information by sampling or abandoning the research itself. The use of the Rorschach, the TAT, the CAT and the Stanford Binet are some of the instances in which sampling becomes all the more important. Surveys using sampling provide greater flexibility and scope.

(iv) Greater Accuracy : With the reduction in the volume work, personnel of higher expertise and training can be employed and a more careful supervision of the field work and processing of the data are possible. Hence sampling may produce result which are more accurate than which could have been obtained through a complete census. Moreover, sampling is particularly more important in obtaining accurate results about phenomena which are

undergoing rapid changes such as opinions about political and social use.

1.1.5 Requirements of a Good Sample

There are two basic requirements of a good sample i.e. its representativeness and adequacy. If information from sample data is to be generalized to a population, it is essential that the sample would be a miniature or replica, a copy, in all respects of the population from which it has been drawn. It should apply at least to the characteristics directly under investigation or those likely to affect those characteristics indirectly. In practical use the researcher may seek a random sample rather than necessarily a perfectly representative one. Such a sample will fall within the random sampling errors, and will facilitate the estimation of the population parameter on the basis of probability theory.

A good sample not only needs to be representative, it needs to be adequate or of sufficient size to allow confidence in the stability of its characteristics.

An adequate sample is one that contains enough cases to ensure reliable results. Hence planning in advance for the size of sample is very important. The procedure of determining the sample size required varies with the characteristics under study and its distribution.

The size of the sample is determined on the basis of the variability of the population, the degree of precision required and the level of confidence at which the results also are to be accurate. If the population is homogeneous, a small sample is sufficient on the other hand, if population is heterogeneous a much larger sample is needed. The sampling measurement techniques ought to be selected and employed very carefully.

1.1.6 Bias in sampling

A sample that is not representative of the universe is generally known as a “biased sample”. The very adjective “biased” indicates that the sample has been drawn with some prejudices and prefectness in the researcher's mind or some of the units in the population have been unduly favoured in selection or that others were at the disadvantages.

1.1.7 Various sampling techniques

Sampling designs can be broadly classified into two categories.

(i) Probability Sampling Designs and

(ii) Non-Probability Sampling Designs

(i) Probability sampling technique has the following kinds :

i) Simple Random Sampling

ii) Systematic Sampling

iii) Stratified Sampling

iv) Cluster Sampling

v) Multiphased Sampling

(ii) Non-Probability Sampling has :

i) Incidental Sampling
ii) Purposive Sampling
iii) Quota Sampling
iv) Sequential Sampling

The probability designs are based on random selection as the fundamental element of control and permit the specification of the precision that can be obtained and the size of the sample required for that purpose. The non-probability designs are based on the judgment of the investigator as to the most important element of control. An investigator may be instructed to interview 100 persons passing a certain market crossing, interview on phone a specific number of car-owners, or house-wives. The guiding principles in non-probability designs are the availability of the subjects, the personal judgment of the interviewer, and the convenience in carrying out a survey.

However, in non-probability designs, there is a risk of over weightage of the cooperative and the available elements. Moreover, dependence exclusively on the investigator's insight does not lend to design any statistical procedure for the purpose of determination of the margin of sampling errors. The new researcher should not use non-probability techniques as he lacks proper judgement and experience.

It may be possible in some studies to combine the probability and the non-probability procedures. But such mixed designs are complicated and beyond the comprehension of an average student.

1.1.8 Probability Sampling Methods

1.1.8.1 Simple Random Sampling : Sample studies deal with samples drawn from finite population or population which contain a finite number of units.

Simple random sampling, theoretically is a method of selecting in units out the N units in such a way that everyone of the Non samples has an equal chance of being selected In practice, however, simple random sample is drawn unit by unit through the following steps of procedures.

1. Define the population
2. List up all the unit in the population and number them from n1, n2 to N
3. Decide upon the size of the sample, or the number of units to be included in the sample.
4. Use lottery-method, random number tables or some mechanical/ electronic device such as computer to pick up the required elements.

The foregoing discussions would remove the layman's misconceptions of random selection as something haphazard, careless, unplanned hit and miss and something involving exercise of non-conscious selection on the part of the investigator. Such ideas are far from correct. However, simple random sampling is the simplest probability design calling for no special expertise and training or even insight. It can be used mechanically by anybody who has all the population elements listed and a set of tables of random numbers.

At the same time, it allows for the control of sampling error, but a randomly chosen sample may look most "unrandom". If a sample, by chance, contains no women although there may be fair proportion of women in the population, the sample is unrepresentative of the population and hence cannot reflects its true characteristics. But such "extraordinarily unlikely", events may occur, such a sample being close to the tails of the sampling distribution. If the investigator goes on the taking large number of such samples in precisely the same way, on the average, the estimates of the characteristics derived from them all would correspond to the true characteristics in the population.

The unrepresentativeness of a single sample does not throw doubt on the randomness of the sampling procedure. In cases where reoccurrence of such "unrepresentative looking" sample is detected, the investigator will have grounds for suspicion and will be advised to review the procedure of sampling used by him. It may, however be understood that the randomness of process can be gauged by studying the results of repeated samples, not by the appearance of a single one.

Before concluding the discussion on simple random method, it would be appropriate to mention the precautions necessary to avoid inadvertent departures from randomness :

(i) The definition of the population and of the observation should precise and coinciding with each other.

(ii) The definition or list of all the population elements should be complete.

(iii) The mechanical procedure of drawing the sample should be easy to carry on and should not allow any biases to enter.

1.1.8.2 Systematic Sampling

A variation of the random process of sampling is the systematic selection of the required number of the elements in the population to be included in the sample. The process involves the following steps :

1. List up the population elements in some order-alphabetical, seniority, street and house number and the like.

2. Determine the desired sampling fraction, Say 100 out of 1000; and also

3. Starting with a randomly chosen number between I and K, both inclusive select every kth element form the list. If in the above example the

randomly chosen first number of the sample is 5, the sample will be composed of the following : 5th, 15th, 25th, 35th, 45th, 55th, 65th, 75th, 85th, 95th, to 955th elements of the list. Since the elements are chosen from regular intervals, the technique is also known as "Sampling by Regular Intervals", "Sampling by fixed intervals, and "Sampling by every kth unit".

(a) Systematic Selection compared with the Simple Random Sampling :

The procedure of systematic selection is easier and more even spread of the sample over the population list and hence leads to greater precision. Strictly speaking Systematic sampling is not equivalent to simple random sampling except in cases where the population list is in a random order. But generally ordinary list are not so. Order the list on any basis, or get hold of a list already prepared, there is invariably some systematic arrangement. Furthermore, the choosing of the first number randomly does not make it so. It does not provide all the possible samples the equal chance of being selected. Moreover, the random selection of the starting points determines the other units of the sample.

This dependence or linkage of one member of the sample on the previous one makes the process different from simple random method, in which selection of every member is independent of other.

Moser and Kalton (1971) have used the term, Quasi - random sampling for systematic samples selected from lists arranged more or less at random or when the feature, on the basis of which it is arranged, is not related to the subject of the survey. The selection at regular intervals, from such a list, can be considered approximately equivalent to simple random sampling.

(b) Advantages

Cochran (1972) mentions the advantages of systematic sampling over simple random sampling in the following words :

1) It is easier to draw a sample and often easier to execute without mistake. This procedure is speedy whereas simple random sampling would be slow.

2) Initially, systematic sampling seems likely to be more precise than simple random sampling. Infact it stratifies the population into a strata, which consists of the first units, the second k units and so on. We might, therefore, expect the systematic sample to be about as precise as the corresponding stratified random sample with one unit per stratum. The difference is that, with the systematic sample the unit occur at the position in the stratum is determined separately by randomization within each stratum. The systematic sample is spread more evenly over the population, and this fact sometimes makes systematic sampling considerably more precise than stratified random sampling.

1.1.8.3 Stratified Sampling

In the previous discussion, simple random sampling technique was explained as the one basic many other sampling techniques. It was also pointed out that the precision of simple random sample can be raised by increasing its size. However, it is not the only way. A very popular method to increase the precision is the stratification of the main population into a number of sub-populations each of which is homogenous with respect to one more characteristics and then to select randomly the required number of cases from each sub-population.

In symbolic terms, the population of N units is divided into sub-population of N_1, N_2, N_3 .....N_L units respectively. The sub-populations are non-overlapping and together comprise the whole of population so that $N_1 + N_2 + N_3 +$ .....$N_L = N$.

A sample is then drawn from each sub-population independently of other sub-population. This may involve random, systematic or any other mode of selection. If sampling from each stratum is done randomly, the method can be designed as "stratified random" sampling. If systematic sampling is used it may be called "stratified systematic sampling". The latter is not equivalent to the former. At the most, it can be called 'stratified' quasi-random sampling" However, stratification does not improve the efficiency of both the random selection and the systematic selection. The steps involved in the stratified random sampling are enumerated below :

1) Decide upon the relevant stratification factor as sex residence, age, courses of studies etc.
2) Divide the entire population in sub-population based on the stratification.
3) List up the units separately in each sub-population.
4) Select the required number of units from each sub-population by using an appropriate selection technique.
5) All the sub samples thus stratification make up the main sample.

There are various factors in which the stratification is often done. Selection of these factors depends upon the nature of the study. The factors commonly used for the purpose are ; sex age, income, height, weight, educational status, IQ, residence, volume of business, number of employees, strength of students, course of studies, grade, caste, cultural level, religion and the like.

(a) Problem of Proportionality

When the strata have already been constructed and the population divided into sub population, the question arises whether an equal number of cases is to be taken from each stratum. In other words, whether an equal

weightage is to be given to every sub-sample, or a proportional weightage is desired. The most popular and widely used procedure is to select from each stratum in proportion to the contribution which that stratum makes to the total population. This can be checked from the census records or other reliable source of descriptive idea.

For the purpose of illustrations, suppose a survey of general intelligence of school-going children of the age, sex and residence (Urban, Suburban and Rural). The school census data of district show the following percentage of each category in the population of 14 + to 15 + school - going children.

Population (All 14 + to 15 + school - going children.)

Boy 60%			**Girls 40%**		
Urban	Suburban	Rural	Urban	Suburban	Rural
20%	25%	55%	60%	25%	15%

Suppose further that a sample of 1000 has been considered to be adequate for the purpose. The distribution of various strata in the final sample will then be as follows:

Urban Boys	120	Urban Girls	240
Suburban Boys	150	Suburban Girls	100
Rural Boys	330	Total Girls	60
Total Boys	**600**	**Total Girls**	**400**

In the way, the stratified proportional random sampling requires selection of units at random from each stratum in proportion to the actual size the group in the total population.

Stratification, if done wisely, improves the representativeness of the sample. The proportional weighting of each criterion improves it further and investigator may use a smaller sample and achieve higher efficiency at reduced cost.

1.1.8.4 Cluster Sampling

A cluster is an intact group. Cluster sampling involves division of the population of elementary units under consideration into groups or clusters that serve as primary sampling units. A selection of the clusters is then made to make up the samples. This is cluster sampling. The sampling unit contains groups of elements (cluster) instead of individual members or items in the population.

For example, the purpose of selecting 10% sample from all primary school children in New Delhi, the investigator may list up all the primary schools instead of all the primary school children and select randomly or systematically a 10% sample of the school (clusters of the units) and include all the children in the selected schools in the sample. Although the unit of inquiry may be individual child, yet the unit of sample is school or cluster of children.

Let us consider another example. Suppose, a survey of households of a large town is to be conducted the results about individual household age are

to be obtained. The unit of investigation is the individual household. Suppose further, that there are 30,000 households of all of them listed conveniently in the records of the local corporation and a sample of 300 household is to be selected, it can be done by picking up 300 households randomly from the list of households available in the corporation office. As a consequence of the use of simple random method, the sample would be spread over the whole town which, in turn, involves high field work costs and much inconvenience. But if suppose the town was divided into 600 blocks of 50 household each, a simple course would be to select at random six blocks i.e. 1% and include all the household in them is the sample.

In this way, the sample instead of being scattered over all the blocks, will be confined to only six blocks. Suppose the information were to be collected by direct interviewing, it would be more economical to contact the household because of the proximity if the same in six compact areas. In addition of this, the advantage in the use if a cluster and hot household as a sampling unit receives added importance in case where inadequate listing of all the units of investigation in the population was not readily available. Listing of these units through direct contact would be highly expensive and time-consuming. Cluster sampling has been described in social survey's as ?" area sampling and "area probability sampling". The latter term is somewhat misleading because simple random sampling can be used with a real unit Moreover, the area probability sampling is based to some extent on the principle that a real probability can be used for sampling purpose.

1.1.8.5 Multi-phase Sampling

Multi-phase sampling involves collection of some information from the whole sample and additional information either at the same time or later forms sub-samples of the full sample. After making a simple inexpensive survey of a large sample, an investigator may select sub-samples from its for a more comprehensive investigation.

For example, an investigator plans to conduct a household expenditure survey, but each household is not questioned on all the matters being covered. Only the basic data size and composition of the household, occupation of the head, income, are collected. Information regarding both important aspects such as distribution of expenditure over the main items may also be sought from the entire sample, because the analysis of results are to be based on this information. However, there may be some less important matters on which the investigator does not require detailed analysis of such high precision. Seeking less important type of information from the entire sample may impose a considerable burden on the respondents. The question then is : can the burden be reduced ? The investigator then may select sub-sample from the main sample and question them on the less important issues and thus, reduce

the burden of answering on a very large selection of the main sample. Moreover, information about certain factors known to be fairly constant in the population may be obtained by sub-samples without loss in precision.

In addition to the consideration of lessening the burden of the respondents, the factors of high costs and difficulty in collecting some information may also warrant the use of only a small part of the entire sample for the purpose, thus involving the use of multiple phase sampling.

Multi-phase sampling may be distinguished from multi-stage sampling which it can indeed be combined. The main distinction between the two lies in the use of unit of sampling different level. In the multi-stage sampling, different types of sampling units (administrative districts, polling constituencies) individuals are sampled at different sampling stages : in the Multi-phase Sampling the investigator is concerned with the same type of sampling unit at each phase but some units are asked for more information than others. With only one sub-sampling the technique may be called two phase sampling or double sampling.

The likely advantages of the procedure include reduction of burden on the respondent considerable economies in terms of expenditure, time and labour and improvement of the precision of the sub-sample data though the information collected on the entire sample. Data collected during the first phase may be used for stratification purposes in the selection of the sub-sample.

Post stratification ration and regression estimation techniques can also be used to improve the precision of the sub-sample results. The effect of non-response in the sub-sample can be estimated, unrepresentativeness of the remaining sample can be gauged on the basis of the basic data and if necessary, re-writing can be done to counteract the effects of unrepresentativeness.

1.1.9 Non-Probability Sampling Method

Introduction : Samples which are selected through non-random methods are called non-probability samples. Depending upon the technique used, they are sometimes called accidental, incidental, purposive and quota samples. Since the term non-random is the most expressive of their nature, these may be designated as non-random sample. The main feature of these samples is the lack of control of the sampling error. Hence these can also be referred to as 'uncontrolled' sampling methods. Since the selection of the units in these samples is based on judgement and not on equal or known probability, the same as a class are known as non-probability sampling methods. In this section a description of the various non-probability methods of sampling is given. It is followed by a critical appraisal of the same.

1.1.9.1 Incidental Sampling

The term "incidental sampling" (also called accidental sampling) is generally applied to those groups which are used chiefly because they are

easily or readily obtainable. A researcher employed in a school of psychology may use the students enrolled in the department, a professor of education may undertake a study using and student enrolled in education classes and a research worker may use the children of a local school or animals available in the local laboratory. These subjects are available in the number and under conditions none of which may be of the experimenter's choosing. Such casual groups rarely constitute random samples of any definable population.

The main considerations in terms of merits of such procedure are : the administrative convince of obtaining subjects for study, the case testing, saving in time and completeness of the data collected, the demerits of such a procedure lie in the fact that since there is no-well defined population and no random method of selecting the sample has been used, standard error formulas apply with a high degree of approximation if at all to incidental samples and no valid generalizations can be drawn. Any attempt at generalizations based on such data will be-misleading.

1.1.9.2 Purposive sampling or Judgement Sampling

Samples are sometimes expressively chosen because in the light of the available information they mirror some larger group with reference to one or more given characteristics. The controls in such samples are usually identified as representative areas (city, country, state district) representative characteristics of individuals (age, sex, marital status, socio-economic status, race or type of group, school administrators, school counsellors, elementary teachers, secondary school teachers college teachers editors of dailies, house-wives, visitors to a religious shrine). The control may be further sub-divided by special categories within classes such as amount of training, years of experience, or attitudes towards a specific categories phenomenon. Upto this stage, these controls are somewhat similar to those used in stratification. The researcher may select the number of cases in the sample proportionate to the total number of such cases with control characteristics in the population. It is also important to know whether they were based on adequate information or scant and partial information about the population.

Some examples of purposiveness sampling in social science research are newspaper editors reflecting accurately the public opinion upon various social and economic questions on their own : a sample of housewives representing accurately the buyers of canned goods: and a sample of brokers reflecting the opinion of financiers on a new stock issue. If the saying “As Maine goes, so goes the nation is accepted as correct, the Maine becomes an important barometer (a purposive sample) of political thinking.

Purposive sampling differs stratified random sampling in that actual selection of the units to be included in the sample in each group is done purposively rather than by random method. The former is not a satisfactory

procedure if high precision is required. There is absolutely no reason to believe that important characteristics are representative of the population when two or three characteristics are representative. A considerable research evidence has been put forward against such a belief. From the statistical point of view also, this technique is particularly weak. There is no way of calculating the limits of permissible error, on the required number for the sample, if strict probability sampling is not used.

The purposive sampling approach may be useful-where it is necessary to include a very small number of units in the sample. Thus, if one were faced with the problem of finding one district or even five of perhaps a dozen districts to represent the whole of a big country purposive selection might be best approach. The sample of one district or five districts cannot ordinarily be found to represent big country in a number of characteristics unless the investigator was extremely fortune or unless the answer to the problem was known for all purposes before the sample was selected, or unless there is no variability between areas in terms of the desired characteristics.

1.1.9.3 Quota Sampling

Quota Sampling involves the selection of the sample units within each stratum or quota, on the basis of the judgment of the interviewers rather than on calculateable chance of being include in it.

A wide variety of procedures go under name of quota sampling, but what distinguishes them all fundamentally from probability sampling is that, once the general breakdown of the sample is decided (e.g. how many men and women, how many people in each age group and in each social class it is to include) and the quota assignments are allocated to interviewers, the choice of the actual sample units to fit into this framework is left to the interviewers. Quota Sampling is, therefore, a method of stratified sampling in which the selection of the cases within the population is non-random. It is non-random elements that constitute its greatest weakness.

On the issue of quota sampling versus probability sampling. Moser and Kalton hold the quota method to be so unreliable and prone to bias as to be almost worthless: others think that, although it is obviously less sound theoretically than probability sampling it can be used safely on some subjects; still others believe that with adequate safeguards quota a sampling can be made highly reliable and the extra cost of probability sampling is not worth while. In general, statisticians have criticized the method for theoretical weakness, while market and opinion researchers have depend on it for cheapness and administrative convenience."

1.1.9.4 Sequential Sampling

Sequential Sampling is another sampling design of relatively recent origin. In this design, sampling is continued until a significant result on which

decision is based is obtained. Thus, instead of carrying out a study of four hundred cases, it might be able to carry out, for instance, a four stage sequential research programme of one hundred cases each. As soon as a decisive answer is provided, the study is dropped, may it be at the end of the first of second stage. In case, conclusive results have not been obtained at the expiry time of the first or second stages, the study is continued until an answer is obtained, or until the four hundred cases are exhausted.

Summing up, in the matter of choosing a sample from the total population, on final choice of the sample must depend upon the objectives of the problem to be investigated and the time at the investigator's disposal.

In the case of an experimental design, the most acceptable and effective method of choosing a sample is by random assignment of people, to control and experimental groups, from a common population. This is also ensure that we take up the most representative cross-section of population. This is also necessary if our study intends utilizing in the data analysis inferential statistics, which require the assumption of a normal distribution from the total population.

1.1.10 Criteria for Selecting a Sampling Design :

Young has suggested three criteria which should be kept in mind while selecting or constructing a sample design.

(i) A measurable or known probability sampling technique should be used so that the risk of errors in the sample estimate can be controlled, the degree of confidence that can be placed in the published figures can be pointed out and sufficient resources can be made available to get reliable data from the sample.

(ii) Simply straight forward and workable methods adopted to available facilities and personnel should be used.

(iii) An attempt should be made to achieve maximum reliability of results for each rupee spent. Striking at an optimum balance between expenditure and a maximum of reliable information should be the guiding principle.

Generally it is advisable to conduct a pilot study to uncover potential difficulty to provide the investigating staff with training in the statistical as well as field work and thus ultimately saving time and expense.

1.1.11 Sampling Frames :

A Sampling Frame is generally the list of sampling units from which the sample can be selected at each sampling stage. Indices, maps and other population records used for the purpose are also included in this definition. Example of sampling frames are telephones directories, directories of street addresses, electoral rolls, lists of publishers of books, lists of schools and colleges in a state rating records with the local authorities. Sometimes these

lists, maps or indices are in existence and can be readily obtained. Sometimes they have to be prepared at an extra cost before sampling can be effected. Sampling frames are very important and influence every aspect of a sampling design, the population coverage, the stages of sampling the satisfaction used and the process of selection itself.

1.1.12 Errors in Sampling

The samples of behavioural research are not representative and suffer from two types of errors :

(1) Random error and (2) Systematic error. These errors can be classified further as :

(a) Sampling errors and (b) Error of Measurement Thus, it provides a four-ways calssification and has been shown in following manners :

	Random	Constant
Sampling	A	B
Measurement	C	D

Cell A refers to the unaviodable errors that occur whenever sampling is done. The sample selected at random may be high, low or average with regard to the trait measured. This error can be minimized by selecting a large sample.

Cell B refers to errors of bias in sampling, ie. sampling errors which do not cancel out, but rather lean systematically in one or the other direction of the population value. This error is due to any decision of researcher for selecting subject for the sample. The Systematic error exists, the data are of limited use as the basis for generalizing to the population, Thus, Cells A and B refer to the errors and sampling.

Cells C and D refer to errors in the process of measurement, rather than to errors in sampling. The errors in Cell C which occur from the simple fact measurement derived from any instrument of less than complete reliability are inevitable in some degree of error.The error of measurement is cancelled out by selecting a large sample. Sum of errors of measurement is always zero.

Cell D concerns with another bias-that is due to systematic errors of measurement. If, in the testing of subjects for intelligence, the test administrator allows an extra three minutes for the test, there will be probably be systematic tendency for the sample statistics to be higher than the population parameters.

The systematic errors are the bad errors in both in sampling and in measurement. The magnitude of random sampling errors as they affect the sample statistics as given below :-

$$SE_m = \frac{S}{\sqrt{N-1}}$$

If greater accuracy is required, it can be obtained by increasing the size of sample or the homogeneity of the variable under investigation or by using adequate sample design. If the results obtained are systiematecally higheir or lower than the corresponding true value, the sample is biased and the discrepancy is called an error of bias..

1.1.13 Summary

A sample is a miniature picture of the entire group of aggrigate from which it has been taken in other words a sample is a small representation of a whole. The entire group from which the sample has been taken is known as the "Population". Population may be classified is finite population and infinite population.

Reduced Cost, Greater Speed, Scope and Accuracy are the purpose of sampling. Sampling designs can be broadly classified into two categories probability sampling design. Random Sampling, Systematic Sampling, Stratified Sampling, Cluster Sampling and Multiphase Sampling are kinds of probability sampling and non-probability sampling has incidental sampling, purposive sampling, quota sampling and sequential sampling.

Sample studies are subject to sampling and non-sampling errors which are of a random or of a constant nature. The choice of sampling method depends upon many considerations unique to each individual project. These include the definition of population, available information about the structure of the population, the parameters to be estimated, the objectives of the analysis including the degree of precision required and the financial and other resources available for the project. It is advisable to keep the sample to the manageable and reasonable size.

1.1.14 Suggested Readings :

1. Aggarwal,Y.P : Sampling Methods for Social Investigations, Ambala Cantt: Bharati Vidya Mandir, 1976
2. Morris H. Hansa, William N. Hurwitaz and William G. Madow : Sample Survey Methods and Theory, Vol. I, Methods and Application, New York : John Wiley, 1953
3. Moser and G. Kalton,C.A. : Survey Methods in Social Investigations, London : The English Book Society, 1971
4. J. Francis Rummel : An introduction to Research Procedures in Education, New York : Harper and Row, 1964.
5. A. Stuart C.A. Moser and : “An Experimental Study of Quota Sampling”, Journal of the Royal Statistical Society, 1953, p. 116

6. Sharma, T.R. : Vidyak Khoj Vidhi, Patiala, Punjabi University, 1988.
7. S. Sudman : "Porobability Sampling with Quotas", Journal of the American Statistical association, 1966, 61, pp. 749-771
8. Moully, George J. : The Science of Educational Research, New Delhi : Eurasia Publishing 1964, pp. 189-190.
9. Yates, F. : Sampling methods for Census Surveys London : Griffin 1969.
10. Cochran, W.E. : Sampling Techniques, New Delhi : Wiley Eastern 1972

1.1.15 Suggested Questions :

1. What do you mean by Population, Sample, Sampling frame, finite and infinite population ? Give examples of each.
2. Why sampling is necessary in social science research ?
3. What considerations will guide you in the selection of a sampling design ? Illustrate your answer.
4. What is simple random sampling ? Describe its merits and demerits.
5. In what type of situations would you select a non-probability sampling design in preference to a probability one.

LESSON NO : 1.2 **AUTHOR : DR. S.S. GREWAL**

Structure of the lesson

1.2.1 OBJECTIVES:

After reading this lesson,the students will be able to:

1. Understand the different tools of educational research.
2. Distuinguish between the test,questionnaire and rating scale.
3. Explain the uses of test,questionnaire and rating scale.

4. Give example of each tool.

5.Describe the advantages and disadvantages of every tool.

1.2.2 INTRODUCTION

Scientific methods employ certain tools for collecting data. In physical sciences, very definite, dependable, precise and useful tools have been developed for taking readings of measurement in weight, length, volume, density, viscosity, elasticity etc. Social Sciences like Psychology, Education, Economics, Sociology are also trying to develop some reliable and valid tools for collecting data on intelligence, aptitude, attitude, perception, personality and so on. To get a correct measure of human abilities and behaviour is more difficult a task than having a measure of weight or density. Social scientists, as a result of constant labour and insight into the human behaviour have developed fairly reliable tools. Some important tools are (a) Test (b) Questionnaires and opinionnaires (c) Interviews (d) Observations, check lists (e) Inventories (f) Rating scales (g) Projective and diagnostic techniques and so on.

We shall discuss here some of the tools. It may, however, be noted that there is merit in supplementing one tool with the other to counteract bias (in the use of one tool), if any, and generate more adequate and valid data.

1.2.3 TESTS

A large variety of tests namely Intelligence Tests, Personality tests, Aptitude Tests, Achievement tests, Performance tests have been constructed and standardized by social scientists. We can define a test as an instrument designed to describe and measure a sample of certain aspects of human behaviour. Best says : "Test may be used to compare the behaviour of two or more persons at particular time, or one or more persons at different times. They yield objective and standardized description of behaviour, qualified by numerical scores. Under ideal conditions achievement or aptitude test measure the best of performance of which individuals are capable. Tests are also used to describe status, or a prevailing condition at a particular time.

In experimental studies, tests are used to equate experimental and control groups to describe the status in achievement, to measure pupil mean resulting from the application of experimental variables and to evaluate the relative effectiveness of teaching methods.

In surveys, achievements test, intelligence tests and other psychological tests are used to find out the present conditions.

1.2.3.1 MAJOR FORMS OF TESTS ARE:

i. Intelligence tests : Individual or group, verbal or non-verbal.

ii. Personality test: Rorschach Ink blot test, 16 P.F. Questionnaire, Anxiety scales etc.

iii. Achievements tests
iv. Aptitude tests

A good test should have the following qualities :

1) Validity
2) Reliability
3) Objectivity
4) Economy to be administered in a short period, less costly to prepare and easy to administer
5) Simiplicity of administration, scoring and interpretation.
6) Interesting - the tests should be interesting and enjoyable.

1.2.3.2 SOME EXAMPLES OF TESTS

1. Group Test of General Mental Ability By S.S. Jalota. It is meant for 11 to 16 years.
2. Group test of intelligence by R.K. Tandon it is meant for college educated adults.
3. Passi-Usha test Creative Problem Solving by B.K. Passsi and Usha Kumar. This test can be used with children and adults for measuring their creative problem solving ability.
4. Value test by S.N. Upadhyaya. It consist of 18 items and measures terminal and instrumental values of life.
5. Modernization Scale by R.S. Singh. It measures attitude towards social change religion, marriage, position of women and education.
6. Parenting Scale by R.L. Bhardwaj. This is intended to measure perceptions of the individual on one's own feeling as to how one is brought up by one's parents on eight modes of parenting.
7. Youth problem Inventory by M. Verma It is meant for adolescents and youth from 16 to 20 years.

1.2.3.3 USES OF TESTS

Tests of mental ability have been widely used because scores of them are correlated with school achievement. Almost all the researchers have provided that intelligence tests can be used to measure readiness for learning at different levels.

Intelligence tests are used for conducting research in different areas of human abilities. They are also used for studying mental growth and direction of individual and group curves.

They have been proved to be best predictors of future educational performance. For making decisions about future education, intelligence tests

can be used to predict the subsequent success of high school students. Teachers can use intelligence tests to make decisions for individuals students regarding their success in college or university.

Tests can be used to provide vocational guidance at different age levels in various vocations.

1.2.3.4 LIMITATIONS OF TESTS

An intelligence test permits a person to show what he can do at a certain time with certain carefully selected but small set of items taken from all the possible items which test the intelligence. Hence it is concluded that it does not tell what we really want to know.

One of the major defects of present day testing is that it is unable to get below the surface of the mind. It measures what a child knows rather than how far he can go in the pursuit and discovery of ideas. It has almost no bearing on originality or the mobilization of ideas towards a single concept or on the ability to devote his attention over a period of time to a single line of thought.

While it will be too much to claim that all standardized tests meet optimum standards of excellence, those instruments have been made as sound as possible in the light of the best that is known by experts in test construction.

1.2.4 QUESTIONNAIRES

There are times when a personal contact, as in an interview, is too expensive or time consuming. It is in those situations that questionnaires are administered by mail or otherwise. The questionnaires as a technique of data gathering has made the greatest contribution to research.

Questionnaries are widely used in research to obtain information about current conditions and practices and to make inquiries concerning attitudes and opinions. In Dictionary of Psychology, Drever (1956) defines a questionnaires as series of questions dealing with some psychological, social or educational topic to topics, sent or given to a group of individuals, with the objective of obtaining data with regard to some problems; sometimes it is employed for diagnostic purpose, or for assessing personality traits. In fact, a questionnaire consists of number of questions printed/typed in a definite order on a form (or set of forms). The form(s) are presented in a uniform manner to a number of persons who are expected to read, understand the questions and reply to them in writing in the relevant spaces provided for the purpose on the said form(s). In certain circumstances this is the most effective method of eliciting information.

1.2.4.1 TYPES OF QUESTIONNAIRES

Questionnaires are of two type : (1) Structured, or closed form (2) Unstructured,

or open form. A structured questionnaire contains the questions and alternative answers to them. The answers provided for each question are exhaustive of all possible responses and at the same time are mutually exclusive. Unstructured questionnaires do not include suggested answers.

Again, structured questionnaires are those in which there are definite and concrete questions to elicit more detailed responses. The questions are presented with exactly the same wording, and in the same order to all the respondents. The reason for standardization is to ensure that all the respondents are replying to the same set of questions. The form of the questions may be either closed or open; the important point is that they are stated in advance not constructed during the process of questioning.

1.2.4.2 EXAMPLES OF QUESTIONNAIRES

What is your monthly income from all sources ?

a) Below Rs. 3000/- p.m.

b) Between Rs. 3000/- to 5000/- p.m.

c) Between Rs. 5000/- to 9000/- p.m.

d) Above Rs. 9000/- p.m.

On the other hand, open-ended questions are designed to permit a free response from the subject rather than one-limited to certain stated alternative. The distinguishing characteristics of open-ended questions is that they merely raise an issue but do not provide or suggest any structure for the respondent's reply. The respondent is given the opportunity to answer in his own terms and in his own frame of reference.

Example

What do you think is the cause of mass failures in the examination ?

1.2.4.3 DESIGNING OF QUESTIONNAIRES

Designing a questionnaires is a complex procedure. The nature, form and order of the questions is of great importance for meaningful results to be obtained. It involves three aspects. The researcher must first learn how to ask a question that is definite and quantifiable. Second, the format of the questionnaires must be structured so that the respondent will have no difficulty in recording his response and also he may not miss any item. The questionnaires must be composed in logical sequence of questions. Third, the instructions must be sharpened so that all ambiguities are eliminated.

Questions should be asked in a way in such a sequence that the answers can be easily organized for the purpose of data processing. Close ended questions which ask the respondent to choose from a provided list of

answers will do in most cases. These provide processed data more easily than open ended questions that permit the responding person to create an answer.

A pilot study can be most useful in devising the actual working of questions, for quite often ambiguities in the questions will be revealed by the nature of the replies which are obtained. It is difficult to phrase entirely unambiguous questions, but there are certain principles which an investigator might follow as precisely as possible. Sometimes terms or words such as 'often', 'occasionally' and 'seldom' have no agreed meaning and they will be interpreted differently by different persons. These words should be avoided.

The use of complex questions should also be avoided, since these can lead to confusion and difficulty in interpretation. General instructions should how the answers are to be given for close ended questions and whether a long or short answer is to be given for an open-ended question. Echo type, suggestive, double barreled, ambiguous, incomplete, clumsy, circumlocutionary and leading questions should be carefully avoided.

1.2.4.4 ADVANTAGES OF QUESTIONNAIRES

Questionnaire helps us to obtain maximum information about an individual at one sitting. It is most economical. It is less-time consuming. It helps us obtain the data of a large number of people in a sitting of say 10 to 15 minutes. If it is properly framed, it is quite reliable. It is easy to administer and easy to score.

Fixed, or closed questions have the advantages of being standardizable, simple to administer, quick and relatively inexpensive to analyse.

1.2.4.5 LIMITATIONS OF QUESTIONNAIRES

The method cannot be used on illiterate persons and children. There is no provision in this method for coming face to face with the respondent. Howsoever much a researcher may try to simplify and rationalise his questionnaire, it is not possible to avoid the use of each and every technical term. Some persons write so badly that even they themselves find if difficult to read their own hand. Some persons give answers which are so brief that the full meaning is incomprehensible. The main limitation is poor or inadequate return of filled in questionnaires.

1.2.5 RATINGS SCALES:

Rating scales provide a systematic procedure for obtaining and reporting the judgment of observers. Thus, rating scales involve assessments by one person of another's personality. Typically, a rating scale consists of a set of characteristics or qualities to be judged and some type of scale for indicating

the degree to which each attribute is present. The rating form itself is merely a reporting device. It can be treated numerically and as a source of data for many types of research. It has been used in research on children's learning outcomes, their development and other aspects of behaviour.

To develop a rating scale, the research worker attempts three important things: (i) identifies the variable to be measured, and (ii) place categories or units on a scale to differentiate varying degrees of that factor, and (iii) describes these units in some way. However, no rule governs the number of units that are used on a scale. Having too few categories makes it difficult to produce crude measures that have little meaning, and having too many categories make it difficult for the rater to discriminate between one unit and the next on the scale. Hopkins has suggested some guidelines for developing scales. These are as follows :

1. Do not have too many or few point categories on the scale. Too few points result in a crude measure. Too many, make discrimination, difficult and time-consuming for the observer. In any case, these categories must be three or more. Two categories make only a check list.
2. Use odd number of points on the scale.
3. Allow reactions at only the listed points. If the observer makes in between points, interpretation must be made arbitrarily by the interpreters.
4. Select meaningful description for categories, careful selection of adjectives and modifying adverbs gives direction to the observer.

1.2.5.1 TYPES OF RATING SCALES

Rating scales may take many specific forms, but the majority of them can be belonging to one of the following types:

1. Numerical rating scales

One of the simplest types of rating scales is that where the rater select one of a number of categories that are arranged in an ordered series. Typically, each of a series of numbers is given a verbal description that remains constant from one characteristic to another. Scales with five to seven categories have been most frequently used but some have used as many as eleven points. Thurston's scales are of this type.

Example :

Direction : Indicate the degree to which the teacher contributes to effective teaching by encircling the appropriate numbers. They represent the following values :

5 - outstanding, 4 - above-average, 3 - average, 2 - below average and 1 - unsatisfactory.

1. To what extent does the teacher show evidence of clear - cut and worthy professional goals ?

1 2 3 4 5

2. To what extent does the teacher attack problems in a constructive manner in his/her class room ?

1 2 3 4 5

The numerical rating scales is useful when the characteristic to be rated can be classified into a limited number of categories and when there is general agreement concerning the category represented by each number.

In general, the more clearly defined the categories, the more reliable the ratings are likely to be. Of course, how much specifications is needed depends on the fitness or distinctions required for the purpose of the study, the nature of the material etc.

2. Graphic rating scale

This is perhaps the most widely used rating scale. In this type, the rater indicates his rating by simply a mark (✓) at the appropriate point on a horizontal line that runs from one extreme on the characteristic in question to the other extreme. Typically, a set of categories identifies specific position along the line, but the rater is free to check between these points if he so desires.

Example

Directions : Indicate the degree to which pupil contributes to class discussions by placing (✓) anywhere along the horizontal line under each item.

1. To what extent does the pupil participate in discussion ?

never	seldom	occasionally	frequently	always

The scale shown in this example used in same set of categories for a characteristic and is commonly referred to as a constant alternatives scale. Where these categories vary from one characteristics to another, the scale is called a changing alternative scale.

3. Matching Scale

At each point a person, well-known to the rates is but, and the rater matches the subjects with these persons and puts him at the point where the example with whom he matches is placed.

4. Rank Order Scale

In this type of rating scale, the rater is required to rank subject/persons specifically in relation to one another. He indicates which person rates the highest and so on.

One of the problems in constructing a rating scale lies in the difficulty of conveying to the rater just what quality one wishes to be evaluates. It is likely

that a brief behavioural statement is more objective than an objective that may have no universal meaning in the abstract.

1.2.5.2 MERITS AND USES OF RATING SCALES

Rating scales can be used in the evaluation of a wide variety of learning outcomes. They are specially useful in evaluating procedures. They direct our attention to the same aspects of performance in all pupils, and they provide a common scale to record our judgments.

Where pupil performance results in some types of product, it is more desirable to judge the product rather than the procedures. The product in rating provides desirable evaluative information in many areas. It is useful in evaluating such things as hand writing drawing made, graphs, and note-books.

One of the most common uses of rating scales in the schools is for the rating of various aspects of personal social development. Most report cards have a special place of rating the pupils such attributes as interest, class room conduct and co-operation. In addition, teachers, may desire to rate each pupil on standard rating from a periodic intervals. In this case, the rating scale items should be derived directly from the specific learning outcomes. The ratings are usually made immediately after a period of directed observation. Such rating generally reflects most of the teacher's feelings and personal biases than those obtained at the end of a period of planned and directed observations.

1.2.5.3 COMMON ERRORS AND LIMITATIONS IN RATING SCALES

All techniques of rating are subject to considerable errors. These include errors due to (1) personal bias, (2) halo effect, and (3) logical error. Personal bias errors are indicated by a general tendency to rate all individuals at approximately the same position on the scale. Some raters tend to use the high end of the scale only. This is the generosity error which refers to the tendency to give the subject the benefit of any doubt. This is when the rater is not sure, he is likely to speak favourably about the person he is rating. In contrast, there is the error of severity, which is a tendency to rate all individuals too low on all characteristics. Training of the raters is probably the best way to reduce the errors. Of course, specific types of errors can be checked by taking specific steps. For example, the generosity error may be reduced by using relatively neutral descriptive terms for the scale positions rather than by using evaluative ones.

Another types of systematic error in rating people is the halo-effect. This occurs when a rater allows a generalized impression of the subject to influence his rating on every specific aspect of behaviour. This general impression carries from one item in the scale to the next. For example, some teachers are so much impressed by the academic brilliance of some students that they fail to see any fault in them in any respect of personality or behaviour.

A logical error results when two characteristics are rated as more alike, or less alike, than they actually are because of the rater's belief concerning their relationship. In rating intelligence, for example, teachers tend to over-rate the intelligence of pupils with high achievement because they logically expect the two characteristics to go together. These errors do not result from biases towards certain pupils or certain positions on the rating scale, but rather from the rater's pre-conceived notions concerning human nature. The rater assumes a more direct relationship among traits than what actually exists.

The various types of errors that appear in rating of different numbers (3-point, 5-point, 7-point, 9-point, 11-point), it is considered that a 3-point scale is too rough to bring about the real quality, 9-point or 11-point scale is considered too fine as the respondents may not be able to demarcate between the points lying too close to each other. The 5-point scale is considered as the most optimal scale for most researchers.

1.2.6 QUESTIONS FOR SELF EVALUATION

Q1) Tests,Questionnaires and rating scales are the tools of educational research. Yes/No

Q2) A good test should not have the qualities like Validity and Reliability. Yes/No

Q3) While testing it is not possible to get below the surface of the mind. Yes/No

Q4) Questionnaires as series of questions dealing with some topic, given to a group of individuals for obtaining data. Yes/No

Q5) Questionnaires are not economical and more time consuming. Yes/No

Q6) Rating scale is merely a reporting device. Yes/No

Q7) Ratings can be done by numerical rating scale, graphic rating scale,matching scale and rank order scale. Yes/No

ANSWER KEY: Yes,No,Yes,Yes,No,Yes,Yes

1.2.7 SUGGESTED QUESTIONS

1. Discuss the rationale, merits and demerits of tests as tools of educational research.
2. Illustrate the use of Rating Scales as tools of educational research, giving the limitations of Rating Scales.
3. Compare the merits of various tools of educational research in terms of their effectiveness.

1.2.8 SUGGESTED BOOKS AND WEB SOURCE:

1. Ary Donald, Jacobs, L.C. and Razavick Asghar : Introduction to Research in Education, New York : Holt; Rineharst & Winston Inc., 1972.
2. Best, J.W. : Research in Education, Prentice Hall Inc., Englewood Cliff, 1983
3. Goode, WJ & Hatt, P.K. : Methods in Social Research, New York : McGraw-Hill book Co., 1952
4. Gronlun, E. Norman : Measurement and Evaluation in Teaching New York : Mac Millian Publishing Co., Inc., 1981.
5. Lowell & Lawson, K.S. : Understanding Research in Education, Univ. of London Press Ltd., 1970
6. Meenakshi : First course in Methodology of Research, Patiala : Kalia Parkashan, 1992
7. Sharma, T.R. : Vidyak Khoj Vidhi, Patiala: Punjabi University.
8. Whitney, F.L. : The Elements of Research, New York: Prentice Hall, 1961

1.books.google.co.in, 2. www.answers.com

M.A. (EDUCATION) PART-I **PAPER-III**

(Semester-II) **Methodology of Educational Research-II**

LESSON NO : 1.3 **AUTHOR : DR. Y.P. AGGARWAL**

Structure of the lesson

1.3.1 OBJECTIVES:

After reading this lesson,the students will be able to:

1. Define the meaning of interview.
2. Understand the different types of interviews.
3. Distinguish between the observations and sociometric technique.
4. Explain the merits and demerits of each tool.

1.3.2 INTRODUCTION:

In almost every human endeavour, it is clearly visible that man is coming down from superstition and mysticism to scientific fact. Since systematic collection of data is a major characteristics of scientific methods. It is imperative for the physical sciences and the so called less precise science to develop tools and techniques of measurement and quantification which generate reliable and valid data. Where as physical sciences using laboratory techniques have acquired sufficiently well developed instruments, social sciences are trying hard in this direction. The development of highly reliable instruments or tools of data has recently become a focus of effort and attention of the sociologists, economists, psychologists, educationists and anthropologists. A more precise measurement of human behaviour in terms of perceptions, attitudes, exceptions, anticipated behaviour personality and intelligence is progressively becoming possible. Despite the fact that social human phenomena are less observable directly, the social scientists have devised methods of communicating with the sources and derive insight from it.

Tools and techniques on a table to a social scientist and researcher are like tools in carpenter's chest. Each is appropriate for a given situation. These include questionnaires, interview, observations, rating scales, attitude scales, sociometry schedules, check-lists, psychological and educational tests, and inventories. The last category comprises personality inventories, intelligence tests, attainment and aptitude tests, projective techniques, diagnostic tests and the like.

To collect data in any research, some kind of tool is pre-requisite. But no tool will help unless it is sharp, flawless and reliable. More than that it has to be capable of yielding reliable, objective and valid data. For that if the researcher has to do a lot of planning with a view to preparing it for achieving the objective of gathering scientific data.

“Many behavioural scientists feel that there is too much dependence upon single method of inquiry. Because each type of data gathering device has its own particular bias, there is merit in supplementing one with another to counteract bias and generate more adequate data. Students of research should familiarize themselves with each of these tools of research and attempt to develop skill in their use and application. Here, we shall discuss only such tools which have not been discussed in the previous lesson.

1.3.3 INTERVIEW

An interview is normally a situation of face-to-face relationship between an interviewer and a respondent. The latter may be called an interviewee or an informant. A telephonic interview, in which the interviewee is sitting at a long

distance, is also possible but may not be as effective in data gathering as a face-to-face situation. Young says, interview is not a simple two-way conversation between an interrogator and an informant. Gestures, glances, facial expression and pauses often reveal subtle feelings. "The tell-tale aspects of vocal and verbal expression like intonation, rate of speech, difficulties in enunciation, a blush, a nervous laugh, or an undue embarrassment may reveal important clues about the psychological make-up of the person.

The interaction that take place in an interview is highly complex. What the interviewers hear should be regarded by them as representing merely "the perception of the informant, filtered and modified by his cognitive and emotional reactions and reported through his personal verbal usages of the picture of the words as he sees it as he is willing to pass it on to as in his particular interview situation." To sum up an interview is

1. A face-to-face situation.
2. It involves an interviewer and a respondent or interviewee.
3. A situation is set up with a specific purpose. Casual and incidental meetings and conversations may not be termed as interview in the strict sense of the word as it is used in research unless they were planned so.

1.3.3.1 TYPES OF INTERVIEWS

Interviews can be classified as follows depending on the parameter under consideration.

(A) According to Function

(i) Clinical interview
(ii) Research interview
(iii) Admission or selection interview.

(B) According to Number of Persons Participating

(i) Individual interview
(ii) Group interview

(C) According to Length of Contact

(i) Short interview
(ii) Prolonged interview

(D) According to the Roles assumed by the interviewer and interviewee

i) Non-directive interview (uncontrolled, unguided or unstructured)
ii) Depth interview (intensive and searching hidden psychological factors)
iii) Focused interview (concentrates attention on one specific issue)
iv) Repeated interview (in order to see change or development)

1.3.3.1.1 The Clinical Interview

The Diagnostic and the Treatment interviews can also be termed as clinical interview. This personal contact is used to secure information about the client and to understand his problems, strengths and weakness, or Information regarding client's past history, family relationship, job adjustment, health adjustment, emotional adjustment and the like may be sought from him. In the case of psychotics, mentally defectives and children, unusual discrepancies/in client's reporting of information may occur. These can be made up by interviewing other persons associated with such clients.

Client's non-verbal behaviour, in the form of depressions, gait, expression, rate of speech, posture, topics avoided and word choice may provide clues to important unconscious psychological mechanisms working in the mind of the patient. Interview may be supplemented by the use of standard diagnostic tests and therapies to accomplish the two important purposes of clinical interview: diagnostic and treatment accomplish the two important purposes of clinical interview : diagnostic and treatment.

1.3.3.1.2 The Group and Individual Interview

Interviews with one interviewee at a time and more than one in group are termed as individual interviews and group interviews respectively. The optimum size that is generally accepted is 10 to 21 persons in a group interview. There is no evidence on the distinctive merits of the two types. However, the convenience, the setting and the type of information to be sought for, may be deciding factors in using a group or individual interview situation. The group interview as compared to the individual interview may have the following merits and demerits.

Merits

(i) Release of inhibitions through personal comments.
(ii) Wider range of responses based on a wider range of experience.
(iii) Recall of forgotten details through the process of group interaction.

Demerits

(i) Controversies and discussions unrelated to the subject may arise and may remain unresolved.
(ii) More Communicative members may monopolize the discussion.
(iii) Continuity of topic may be interrupted.
(iv) Certain attitudes, interests and motives may not be revealed in this quasi-public situation.

1.3.3.1.3 The Focused Interview

The focused interview concentrates attention on some particular events or experience rather than on general lines of inquiry about the event. For example, a focused interview may be held to determine the responses of individual to a

motive of a speech. The focused interview places added emphasis on the subject's verbal report of his definition of a specific situation and responses to it; the method may resemble, in some characteristics, to projective and observational methods. The interviewer's objective evaluation of a particular situation to which the subjects are exposed resembles observational techniques. However, the dimension of objective measurement puts a restriction on the use of this method in social sciences using complex social situation.

1.3.3.1.4 Depth and Non-directed Interviews

The depth interview is intensive and highly searching. It has an emphasis on the identification of such psychological factors, attitude, convictions, motives and emotions. The depth procedures enable the investigator, through depth responses, to determine the respondent's degree of detachment of personal involvement in experience, and the periphery of salient character of responses. The procedure involves i) flexibility of situation. ii) retrospective focus through past experience, elaboration, etc. iii) emphasis on effective or "feeling" aspect, and iv) a comparison between the stimulus situation and subject's own experience.

The non-directive interview is an uncontrolled and unstructured situation which provides much freedom on the part of the client to "talk about" the problem under study. Non-direction is an important technique in counselling and is often used in the focused and depth interviews. Some examples of different amount of non-direction in questioning are given below :

1) Unstructured question (Stimulus and Response free)

(i) What impressed you most in the speech ?

(ii) How do you feel about this movie ?

2) Semi structured question

(i) What did you learn from this movie that you had not known before ?

(ii) How do you feel about the reasons of Manu's leaving home in Parasad's Kamayami ?

3) Structured questions

(i) Judging from the movie you have just seen, do you think the music was good, average or poor ?

1.3.3.1.5 Research Interview

Perhaps the most rigorous and planned of all interviews is the research interview. The types of interviews discussed above may also be used for conducting research. What distinguishes a research interview from the non-research one is the great amount of rigour, pre-planning and structuring which goes into former. Since research is a systematic, objective and well planned

procedure of answering significant questions, interview as a tool of data collection also affects these characteristics. Interview has been used extensively by research worker for obtaining insight into such phenomena as personal adjustment, sex habits and practices, opinion polls etc. Adorno's world famous study of Authoritarian Personality was completed using interview as the major technique.

The distinguishing characteristics of a research interview are mentioned below :

1. Fully structured stimulus.
2. Fully structured response pattern.
3. Pre-planned and standardized instructions or directions.
4. Time, setting and other details of the interview situation fully standardized.
5. Recording based on fool proof mechanical and electronic method such as a video - tape recorder. Interview schedule with coding can also be used.
6. The interview keeps the respondent on the research problem.

Data secured from research interviews lend themselves to quantitative interpretation, since it is consistent and specific and aims at realistic and factual details which can be differentiated, tabulated and measured.

1.3.3.1.6 The Repeated Interview

When the investigator is interested in the identification of specific development and growth in social or psychological process, an interview may be repeatedly used. The study of progressive actions, factors or attitudes which may determine social behaviour be conducted by using this method. It involves a longitudinal data collection although the total span may be a few months only.

It is expensive in time, energy and money but offer a good potential for studying attitudes, behaviour, habits, which are in the process of formation.

1.3.3.2 Conduct of an interview

The process of conducting an interview involves five major steps.

(i) Preparation :- The interviewer has to decide and plan before hand, the following things: Area of information to be covered, background knowledge of the interests, beliefs, special areas of knowledge, etc. of the respondents; a definite appointment for the interview; decision about the convenient setting; try-out of the interview to detect weakness in his methods, questions, recording system; and a definite system of recording.

(ii) Rapport :- As far as possible the interviewer has to be pleasant, efficient, straight forward so that necessary psychological climate for a meaningful conversation can be formed. Although it is not very

easy but effort must be made so that hostility and suspicion should be dispelled. Confidentiality of information should be assured and the respondent be placed in a psychologically easy and no threatening situation.

(iii) Eliciting Information :- The interviewer should function as an attentive and analytical listener who can detect vagueness, contradiction or evasiveness in the respondent's answers. Alternative and more searching questions should be asked to help the interviewee to recall information, amplify statements, clarify thinking and rectify inaccurate facts. Pacing the question, sharpening the focus, follow-up of crucial clues, and cross-checking of the information provided are some important matters which should get due attention of the investigator. Termination of the interview should be done according to schedule or when the respondent becomes tired, when all pertinent information has been collected.

(iv) Recording :- A quick and accurate recording system is a highly essential requirement of any good interview. Human or mechanical-cum-electronic system can be utilized. The tape-recorder and the movie can be helpful in preserving vocal and emotional character of the replies. A note of significant emotional displays like hesitations, stammering, sudden silence or transition, quickly corrected words, and obvious omissions be taken and recorded. The decision whether the interviewee should know or not that his/her interview is being recorded or filmed has to be carefully taken depending upon the theme of the research.

(v) Interview Schedule :- The researcher should always have with him/her a list of questions which he/she intends asking the respondent. The list of questions is called "Interview Schedule". The schedule has to be prepared carefully and it must cover all aspects of the problem being looked into.

1.3.3.3 Merits and Limitations of Interview Technique :

Interview technique has been widely used and especially useful in the following situations.

(i) It is the best suited for the assessment of personal qualities.

(ii) It is of definite value in the diagnosis and treatment of emotional problems.

(iii) It is very useful in counselling.

(iv) It can be used as supplement to other method of data collection.

(v) It can be used for the verification of information collected through

other technique.

(vi) It is only method of collecting data from children, lunatics, uneducated, and the too old subjects.

Interview, as a technique of data collection, suffers from the following limitations some of which may be source of bias and error;

i) It is highly expensive in time, energy and money.

ii) The possibility of bias due to interviewer's own whims and preferences is present.

iii) The interviewer may distort information and kept back pertinent facts from the interviewee.

iv) There is threat to the ego of the interviewee in talking in a semi-public situation.

v) Memory bias often lead to wrong reporting.

vi) 'Emotional set' of the interviewee can lead to reporting of facts to arouse 'self pity' or 'self concern' and hide facts which are uncomplimentary to him or do not antagonize the interviewer.

vii) The content and the form of questions i.e. procedures decided for the interview, physical setting, mode of recording, temporary state of mind of the parties involved may lead to bias and error in interview data.

viii) Other potential sources of error and bias in interview include firmly fixed attitudes, personality characteristics, motives and goals related to group membership and loyalties.

All in all, the researcher, through an interview, may stimulate subject to greater insight into his own experiences, and thereby explore significant areas not anticipated in original plan of investigation. Interview is also appropriate when dealing with young children, illiterates, those with language difficulties and those of limited intelligence.

1.3.4 OBSERVATION

Observation may be defined as systematic viewing, coupled with consideration of the seen phenomenon. Observation is concerned with the overt behaviour of persons in appropriate situations, sometimes under conditions of normal living, and at other times with some special set of factors operating. As a general rule, an observer is not concerned with what a respondent places on paper or with what he says in an interview situation. Persons may furnish wrong information in questionnaires or interviews, but a direct observation of overt behaviour can reveal what the subject actually does. An observant investigator may find very useful clues to the significant aspects of behaviour of personality at the time the subject is filling up a questionnaire or participating in an interview or taking standardized tests. The direct examination of material objects or

specimens is relatively simple and may consist of classifying, measuring or counting. But when the process involves the study of human subjects in action, it is much more complex.

School plants may be observed and studied with respect to material of construction used, number and size of room for various purposes, amount of furniture and other equipment, presence and absence of drinking water and health facilities, etc. and their adequacy as determined by comparison with reasonable or fixed standards.

Examples of observation of human behaviour in various settings like sports activities, debates, classroom, hostel life, recess period, dinning hall, etc. can be multiplied.

A very precaution that an observer has to take while using observation is that he has to guard against the human tendency of serving what he wants to observe. He has to be objective, free bias etc. while observing behaviour.

1.3.4.1 TYPES OF OBSERVATIONS

A technique of observation may be classified as follows :

A. Place or Position of the Observer

i) Participant observation

ii) Non-Participant observation

B. Degree of structuredness in the observation

i) Controlled observation (structured observation) and

ii) Non-Controlled observation (Non-structured observation).

Different combinations of the two major categories can generate situations like the following :

i) Controlled and Participant

ii) Controlled and Non-Participant

iii) Non-Controlled and Participant and

iv) Non-Controlled and Non-Participant.

Participant and Non-Participant Observation

The concepts of participant and Non-participant observation were perhaps for the first time used in Edward Lindeman's publication Social Discovery (1924). It was taken up as a more effective method of study by other investigators like Anderson in his research, *"The Habo."*

The researcher may act as a member of the group (Participant observer) or as an outsider (non-participant observer). A participant observer generally lives with, or shares in the life and activity of the group under study, distinguishing himself to be accepted as a member of the group. By establishing a favourite rapport with the group, the researcher is able to make observations, as would have been possible in a series of interviews or short observation under non-participant conditions. Two non-participant observers may be branded as an

"outsider" and thus, may lead to reactions somewhat different from what the observer would do if "outsider" was not present. A rapport and acceptance of the non-participant observer with the observed, however, may enable him to collect information.

"Passive Participant" observation over long periods of time can also be used with the hope of obtaining reliable data. In Roethisberger and Dickson's study *'Management and the Worker'*, researchers functioned as passive members of work groups within the industrial plant. They had their desks placed at the back of workroom and kept their presence as unconspicuous as possible while keeping a record of the verbal and overt behaviour of other workers. Physically and personally they were part of the group, but socially they were isolated. An "active participant" observer may function as an active member of the group just like a member of the football team or a teacher employed temporarily in a school with the normal duties but at the same time having the objective to observer the teacher or students to study patterns of human relationship in the school. Both these types have their own merits and limitations.

A non-participant observer, however, may use devices which can collect data with his being present in the group or without others knowing that he was present. The use of one-way screens, sound recorders, photographic equipment and other mechanical, electrical and electronic devices may be of great help in such situations. Child guidance clinics often use a one-way minor cabin in the playroom of the clinic.

Merits and Limitations

The advantages and the limitations of the participant and the non-participant observations depend largely on the situation. The participant observer can get deeper insights into the life of the group being observed by becoming a regular and accepted member of group. However, with the passage of time, he may become less observant and blinded to the particularities which he is supposed to observe. He may thus miss significant incidents and information. By becoming a part and parcel of the group he may lose his objectivity, neutrality and accuracy. He may lose his status as a scientific observer. Some investigators feel that the distortion caused by the non-participant observer may not be significant and after short period of adjustment, the observer may resume the usual behaviour. Vales found that no difference would take place in the observations made by the observer in the open and those behind the one way screens.

1. Controlled and Non-Controlled Observation

All scientists have an ideal the objectification of their methods and techniques. They aspire to observe, record and interpret events in such a fashion that independent observers can verify their findings. Hence, the factor of

control in the observational and recording procedures becomes an important dimension to classify observation as controlled and uncontrolled. The terms "structured" and "unstructured" observation can also be used in place of the controlled and uncontrolled. In controlled observation, only areas considered significant fall within the scope of functioning of the observer. He is expected to make a precise investigation of one or a few restricted aspects of the situation with a view to derive a more rigorous generalization. The "structuring" involves a statement of the standard observation situation including a complete description of the total structure of the situations for the reaching individual group, what the observer does, the questions he asks or the stimulus he presents, the instructions he gives and the comments and the interactions that take place. The standardization of recording procedures is also necessary.

In the uncontrolled observation the observer does not strictly depend upon the pre-planned procedures and scope but rather shifts according to the demands of the situation. As the investigation proceeds and the phenomenon is seen with greater clarity, the investigator may record important clues to various aspects of the problem in hand, Though there is no strict pre-planning and restrictions in the controlled observations, yet it cannot be called as haphazard or casual. It does not have definite aims. It may be more useful in digging out the hidden and less visible aspects of the observer's personality, his motives, attitudes etc. The non-structuring or recording procedures may later create difficulties for organization and analysis of data. Although observations may be placed on a continuum or structuredness, from structured to extremely structured, it may become difficult to differentiate clearly between the two, as some degree of 'structuring' is always involved in all observations.

2. Time Sampling

Time sampling is another important type of observational technique. The total time of observation is divided into a number of systematically spaced, definite, time intervals, say of five minutes. The behaviour is objectively defined and frequency of its occurrence or non-occurrence is recorded. This is structured or controlled observation. Its use is extensively made in the classroom interaction-analysis to record the indirect and direct influences in "teacher talk", "student talk", "student initiated talk", "silence" and "confusion" etc. A child's behaviour in a school can be studied in intervals of 10 minutes after every one hour during his attendance in the school on a particular day. It saves time and provides a better sample of behaviour over a longer period.

3. Critical-Incident Technique

Flangan devised a technique by which the observation lists the presence or absence of some incidents or important occurrences during his observation.

This critical incident defines the trait or traits under study. For example, "a good worker is prompt; a poor worker is tardy". A good worker takes advice and suggestions with appreciation; a poor worker resents criticism and suggestions. An incident is an observable human activity sufficiently complete in itself to permit inferences and predictions about the performing act. To be considered critical, an incident must occur in a situation where the purpose or intent of the acts seems fairly clear to the observer and where its consequences are sufficiently definite to leave little doubt concerning its effect. Critical incident technique has been successfully used in education in the observation of the job behaviour and situation involving human interaction.

4. Recording observation

Human memory is not as dependable as the mechanical one. Hence, to minimise distortion due to forgetting, there is need for immediate recording. On the other hand spontaneous recording through slow techniques may lead the observation to miss significant details. Taking longhand notes is generally inadvisable. However, using a comprehensive check list with simple categories of 'Yes', 'No' or a code letter or number may be a great help to the observer. A number of other devices have been invented for a more precise recording. Thermometer, stethoscope, audiometer, stopwatch, light meter, binoculars, cameras and tape-recorders may be used to make more defined observation. They can be replayed in slow motion to facilitate analysis. A number of paper-pencil devices like check lists, observation schedules, anecdotal records, etc. are also frequently used to aid proper recording. Mechanical devices are no doubt foolproof in recording, yet the analysis of huge amount of data and information including that which is superfluous, generated by them, will be very difficult and time-consuming.

5. Criteria for an observation

Observation as a scientific tool as compared to the most primitive, subjective and impressionistic technique, must have the following characteristics. There criteria will increase the reliability and validity of observation data.

i) Scientific observation is systematic and not a haphazard or opportunistic. The scientific observer must be an expert who knows precisely what he is looking for in the total situation. He should be able to isolate those aspects of the situation which are related with the hypothesis. The observation is oriented to what narrow bands leading to greater dependability of observation.

ii) It is, as far as possible, objective and free from bias. It has hypothesis to guide its course. Though this aspect is highly debatable, yet the usefulness of guiding post has been recognized by many seasonal researchers. As a precaution, however, the observer must maintain his neutrality. Not only

must he consider hypothesis as something to be tested rather than proved he must withdraw from observing those situations in which he cannot be impartial and objective.

iii) The observer must be in a good physical and mental health and possess adequacy of sense organs; he should be able to see and hear clearly.

iv) As far as possible, scientific observation should be quantitative. Many socio-psychological phenomenon may be difficult to qualify, yet the aim should be to derive some means of quantifying observations so that they become more precise and prone to analysis.

v) The observation should have usability, validity and reliability. It would require dependence in repeated observations. A number of observations rather than a single one, and introductions of more than one observer can increase the reliability of the results. People observation is always more reliable than single or one man's observation. Scientific observation is verifiable hence it must be recorded accurately in a manner that leads to verification later on.

vi) The training of observers and a careful planning of the whole process are two other important pre-requisites of good observational study. An inadequately planned observation may turn out to be repetition of the fable, “the blind man looking at the elephant.”

vii) The scientific observation should be based on adequate and representative sample of observations. It would call for deriving a representative sample of observations from a representative sample of subjects as defined in the study.

viii) It uses instruments and devices of a reasonable level of scientific sophistication designed to promote precise observations. A list of such tools is given in the section on recording of the observation.

1.3.5 INVENTORY :

Inventories are research based assessment instrument that probe student's understanding of particular concept. It is a criteria-referenced test designed to determine whether a student has an accurate working knowledge of a specific set of concepts.

Historically, inventories have been in the form of multiple-choice tests in order to aid interpretability and facilitate administration in large classes. Unlike a typical, teacher authored multiple choice test, questions and responses choices on inventories are the subject of extensive research. The aim of research include ascerting (a) the range of what individual thinks a particular question is asking and (b) the most common responses to the questions. Inventories are evaluated to ensure test reliability and validity. In its final form, each question includes one correct answer and several distractes.

Inventories are intented to measure the effectiveness of your teaching, by assessing, on average, what your students learned in the course, and it follow that the results of the class as a whole are more important than

individual students score.

Good Inventories are different from typical exams, in that their developments involves extensive research and development by experts in subject matter to ensure that the questions represents concepts that faculty think are important.

1.3.6 QUESTIONS FOR SELF EVALUATION :

Q1) Interview and Observation are the tools of educational research but socio-metric device is not. Yes/No

Q2) An Interview is a face to face situation which is setup with a specific purpose. Yes/No

Q3) The main limitation of interview is that it is very expensive in time,money,and energy. Yes/No

Q4) In participant observation ,the researcher act as a member of the group and in non-participant observation,he act as an outsider. Yes/No

Q5) In participant observation, the resercher can get deeper insights into the life of the group. Yes/No

ANSWER KEY: No,Yes,Yes,Yes,Yes

1.3.7 SUGGESTED QUESTIONS:

1. What are the uses and applications of Interview technique in Educational Research ?
2. Compare interview techniques with Questionnaire as a tool of data collection.
3. What are the uses of 'observation' in research ? How can observation he made scientific ?

1.3.8 SUGGESTED BOOKS AND WEB SOURCE:

1. Carter V. Good, Introduction to Educational Research, New York : Appleton Century Crofts, 1963.
2. D.B. Ven Dalen, Understanding Educational Research, New York : McGraw Hill, 1996.
3. George, J. Mouly, The Science of Educational Research, New Delhi: Eurasia Publishing Co, 1964.
4. Best, John W. , Research Education New Delhi : Prentice Hall of India, 1978.
5. L. Festinger and D. Katz, Research Methods in the Behavioural Sciences, New Delhi: Amerind Publishing, 1970.
6. Rummel,J.F. An introduction to Research in Education, New York : Harper and Row, 1964.
7. Koul,Lokesh Methodology of Educational Research, New Delhi : Vikas Publications, 1984.
8. Pauline V. Young, Scientific Social Survey and Research, New Delhi : Prentice Hall of India, 1988.

1. upscportal.com,2.thinkexist.com

LESSON NO : 1.4 **AUTHOR : DR. Y.P. AGGARWAL**

RELIABILITY AND VALIDITY OF RESEARCH TOOLS

Structure of the lesson

1.4.1 OBJECTIVES:

After reading this lesson,the students will be able to:

1. Distinguish between the Reliability and Validity.
2. Understand the methods for measuring Reliability and Validity.
3. Explain the factors affecting reliability and Validity.

1.4.2 INTRODUCTION

Suppose we say that a worker is reliable. What does this convey ? It might be supposed that the worker reports at same time everyday; appears in the same condition and performs consistently. This similarity of behaviour from time to time would be comparable to one of the tester's approaches to reliability. “Trustworthiness” is another word used by the common man to

signify the term reliability of human being. Reliability pertains to a class of test characteristics. These are stability, equivalence and internal consistency. These are time associated and form-associated bases of reliability. Reliability is not concerned with appropriateness of measurements, an issue that falls within the scope of the meaning of another test characteristic called validity. The main concern of reliability is an accurate repeatability of scores over time and parallel forms of a test. Reliability of physical measurements using steel tape etc. is very high and these instruments give consistent results with high accuracy. However, psychological tests and instruments are less reliable as many measurement errors are bound to creep in. The Co-efficient of reliability of a test forms the basic index reliability reported in the literature. It is based on self-correlation of a test.

1.4.3 RELIABILITY

1.4.3.1 CHARACTERISTICS OF RELIABILITY:

1. It is consistency of a test scores.
2. It is the measure of variable error.
3. It is the function of a test length.
4. It refers to the stability of a test for a certain population.
5. It is the coefficient of stability
6. It is the coefficient of internal consistency.
7. It is the self correlation.
8. It is the reproducibility of the scores.
9. It is the important characteristics of measuring instrument.
10. It refers to the accuracy or precision of a measuring instrument.
11. It is the temporal stability or consistency of a measuring instrument.
12. It does not ensure the validity or truthfulness or purposiveness of a test.

1.4.3.2 Method of Estimation of Reliability

There are basically four different procedures generally used for computing reliability.

1. Test-retest method (Repetition of the same test or measure over time).
2. Alternate or parallel form method.
3. Split half technique.
4. Rational equivalence method.

All these methods furnish estimates of the reproducibility of test scores and one method may be preferred to another according to the demands of the structure. These methods are described below.

1. Test retest Method

This method involves (i) repetition of a test on the same group

immediately or after a lapse of time, and (ii) computation of correlations between the first and the second set of scores. The correlation co-efficient thus obtained indicates the extent or magnitude of the agreement between the two sets of scores and is often called the co-efficient of stability. Although test-retest is sometimes the only available method, this procedure is open to several serious objections. Immediate repetition of a test may involve (i) immediate memory effects, (ii) Practice effect, (iii) confidence effects, induced by familiarity of contents. Intervals of six months or long in young children may show "maturity effect". The factors of intervening learning and unlearning may lead to lowering of self-correlation. It may not be possible to control conditions on the second testing.

Memory, practice and other carry-over effects may be offset by increasing the time interval between the two testing.

Example : Product Moment or Rank Difference Correlation may be drawn between the first (test) scores and second (retest) scores of the same group laid in two parallel column.

2. Alternate or parallel form method :

This method involves the administration of equivalent or parallel forms of the test instead of repetition of a single test. This avoids to a great extent the disadvantages of the test-retest method involving short or long intervals of time. The two equivalent forms are so constructed as to make them similar (but not identical), in content, mental process involved, number of items and the difficulty levels of the items. The subjects take one form of the test and then as soon as possible, the other form. The correlation co-efficient between the two sets of scores determine the agreement and is generally co-efficient of equivalents. Alternate forms should be drawn very carefully be matching test material for content, difficulty and form. An interval of at least two to four weeks should be allowed between the two administration of the test, to offset the carry-over effects due to familiarity of content. The method has some other limitations also. The second form of a test may not be available. A second testing may place heavier demands on the time of resources of the researcher and the subjects. The psychometricians have devised another technique called the split-half method which takes care of these defects.

Correlation drawn between the scores yielded by form A and those yielded by form B (both tests having been administered at the same time on the same group) illustrates the use in this method.

3. The Split-half method :

The most used procedure for establishing reliability from a single testing consists of dividing a particular test into two presumably equivalent halves. The test is divided into two halves only for the purpose of scoring and not for

administration. It means that single test is given at a single sitting and with a single time limit. However, two correlations between these two sets of scores provides a measure of the accuracy with which the test is halved in the odd-even split technique. The odd-numbered items 1, 3, 5, 7 etc. and the even numbered items 2, 4, 6, 8 etc. form two different sets of items for scoring. This procedure it better than others like first half and the second half and spilt into blocks etc. and ensures balancing of factors in items form, content covered, and difficulty level in the test having 60 or more items.

The computed correlation, in this technique, is between two half-length tests. This value is not directly applicable to the full length test which is the actual instrument prepared for use. Hence Spearman Brown Prophecy Formula is used to estimate the reliability of the full length test from the self correlation of the half tests.

$$r11 = \frac{2 \text{ x } r_{½½}}{1 + r_{½½}}$$

(Spearman Brown Prophecy Formula for estimating reliability from two comparable halves of a test where r_{11} = reliability co-efficients of the whole test.

$r_{½½}$ = reliability coefficient of the half test found experimentally.

Example

For example, if correlation between the two halves of a test is 0.60 then reliability coefficient of the full test will be

$$r_{11} = \frac{2 \times 0.60}{1 + 0.60} = \frac{(1.20)}{1.60} = 0.75$$

The Split - half method is regarded as the best by many psychometricians. The main argument in favour of this method is the single short approach, where there is no place for errors due to repetitions and lapse of time. Convenience and saving of time and expense are some other important considerations.

However, this method suffers from some limitations.

(i) It is not useful for the speed tests.

(ii) Splitting of the test can be done in several ways, thus, leading to the non-comparability of reliability induced from these.

(iii) It is not so useful for tests having lesser number of items i.e. less than twenty.

(iv) Chance errors may affect scores on two halves equally and thus lead to higher correlations.

4. Rational "Equivalence" Method :

The three methods described above suffer from some limitations. The method of rational equivalence is an approach to avoid some of these objectives. This method is based upon inter-correlations of the items in the test and the correlation of the item with the rest as a whole. Several formulas for calculating reliability by this method have been suggested, However, the formula, given below, is the most popular:

$$r_{11} = \frac{n}{(n-1)} \times \frac{\sigma^2 t - \sum pq}{\sigma^2 t} \quad \text{.....2}$$

(Kuder-Richardson formula (2) based on difficulty and the inter-correlation of test items

r_{11} = reliability coefficient of the whole test

n = number of the items in the test

σ = SD of the test scores

p = the proportion of the group answering a test item correctly.

q = (1 - p) = the proportion of the group answering a test item incorrectly.

Another formula which is simple approximation of the above formula is as follows.

$$r_{11} = \frac{n\sigma^2 t - M(n-M)}{\sigma^2 t\,(n-1)} \quad \text{.....3}$$

(Approximation to formula 2)

Where r_{11} = reliability of the whole test

n = number of items in the test

σ = SD of the test scores.

M = mean of the test score.

Formula (3) saves labour as it is based on mean, SD and number of items correlation coefficient between two halves is not required.

These formulas are based on the assumption of equal difficulty of items which may not be generally met in power tests. The formula provides a minimum estimate of reliability thus giving an underestimation.

Example

If a test of 60 items has been given to a group of 85 boys, and σt = 8.50 and ∑ pq = 12.43, By substituting calculated values in formula 2 we have

$$r_{11} = \frac{60}{59} \times \frac{72.25 - 12.43}{72.25} = 0.842$$

1.4.3.3 FACTORS AFFECTING RELIABILITY:

1. Length of Test

Other things equal, the reliability of a test is a function of its length. Longer tests tend to be more reliable than shorter tests. Logically the more samples we take of a given area of knowledge, skill behaviour and the like, the more reliable will be our appraisal of the area. Lengthening the test means, adding more items, or having several applications of the test or use of parallel form. The Spearman Brown Prophecy Formula for calculating the reliability of a test with increased length is.

$$r_{nn} = \frac{nr_{11}}{1 + (n-1)r_{11}}$$

Where r_{nn} = the correlation between n forms of a test and n alternate forms.
Where r_{11} = the reliability coefficient of Test 1

Example

Suppose, in a test of 100 items, the reliability coefficient is 0.70. What will be the effect upon the test reliability of tripling the length of the test ?

Substituting x in r_{11} = 0.70 and n = 3, we have:

$$r_{3\,11} = \frac{3 \times .70}{1 + (3-1).70}$$

$$= \frac{2.10}{2.40} = .87$$

Tripling the tests length increases the reliability coefficient from .70 to .87

Spearman Brown formula shows that after we have reached a high degree of reliability, additional items do not improve the reliability enough to justify the extra time and efforts required for building items for testing pupils.

2. Range of Talent

The range of talent, achievement or availability of pupils on whom the reliability is based has a direct effect on the reliability coefficient. The greater the variability, the higher the reliability.

3. Testing Conditions

The conditions in which the test is administered and scored may effect reliability on either side. Other such factors are : mental set of pupils, level of their motivation, speed of subjects, emotional stability, distraction, accidents and cheating by pupils.

1.4.4 VALIDITY:

Validity of a test of an evaluation device can be defined as the degree to which the test measures what it is intended to measure. A test which is meant to measure achievement in Mathematics is not meant to measure achievement in History or any other subject. In physical science, validity of measuring instruments like scales, thermometers, chronoscope and ammeters is found by comparing their measurement with standard measures. The validity of a psychological test is also found by comparison with standard (and sometimes arbitrary) measure. However, the validity of a physical instrument can be estimated accurately while it can never be done in the case of a mental test.

Validity is a relative term and has reference to a particular purpose or situation. The question "Is the test valid?" can be answer by replying as "Valid' for what ?" Hence there are different types of validity meant for different purposes.

1.4.4.1 Types of Validity

1. Content Validity

Content validity is evaluated by showing well the content of the test samples, the class of situations or subject matter are from which conclusions are to be drawn. It is based on a comparison of the analysis of test content with the analysis of the course content and the instructional objectives. It is to be seen that how well the former represents the latter. The analysis is done essentially through logical, rational and judgement process. That is why, sometimes the content validity may be referred to as rational or logical validity. Content validity is important primarily for measures of achievement. The test maker first determines the widely accepted goals of instructions in the subject and then prepares a blueprint for the test. Test content is drawn from the course content and weighted according to weightage. Content validity of a test involves the objectives of the course and the course content. An appraisal of the content validity of a test involves a careful and detailed examination of the actual test tasks.

2. Face Validity :

Face validity has something to do with the mere appearance of a test. A test is said to have validity when by appearance it "looks like". The

judgemental process is used to determine answers to such questions as whether test content "appears to" correspond to that of the course. A test of mathematics should have numerical question; and a test of history questions about kings, movement, wars, etc. The relevance of the test items to specific situations, groups, language group, etc. is also a matter of concern. However, no single numerical index of the face validity of a test can be calculated.

3. Concurrent Validity

Concurrent validity is evaluated by showing how well the test scores correspond to an already accepted measure of performance of status made at the same time. For example scores on a test of knowledge of basic concepts in Geography can be validated against school grades, teacher's rating, etc. A newly constructed test of intelligence may be validated by finding its correlation with another already existing well-accepted test in this area. In these cases, a correlation between the two sets of measures is calculated as an index of validity. The main problem is to set up a criterion which is independent and reliable.

4. Criterion related Validity

In a discussion of concurrent validity as given above, the test is validated against a criterion at the same point of time. However, we may be interested in using a test to predict some future outcome. A test of aptitude for teaching may be used to admit students to teacher training colleges and expected to predict success at the job as teacher. A scholastic aptitude test may be used to predict how likely will the high school student be successful at a college. A clerical aptitude test may be used to predict success on the job as clerks. We are thus interested in success or performance in the future. This process is also called Predictive Validity. A correlation coefficient between the test scores and the criterion scores is calculated. The higher, the correlation, the better the test as a predictor. However, the problem of selecting "an appropriate criterion" is very ticklish. The main problem arises when the criteria of success on the job are to be determined; the records are not available; time interval between competition of training and placement working on the job for a period is not long enough to allow proper evaluation of success and the like. The success or failure of worker may depend on conditions external to his own personality and skill. Rating of success by superiors may be influenced by many factors other than the proficiency of the worker being rated. However, the following are desired in a criterion measure :

(i) Relevance
(ii) Freedom from bias (providing equal opportunity to all to perform well).
(iii) Reliability
(iv) Availability.

5. Construct Validity

Sometimes questions like the following are asked. "What does this test mean or signify ?" What does the score tell us about the individual ? "Does it correspond to some meaningful trait or construct that will help us in understanding him ?" These questions are related with the construct validity of the test. The term construct is used in psychology to refer to something that is not observable, but it is literally constructed" by the investigator to summarize or account for the regularities or relationships that he observes in behaviour. Thus, most names of traits refer to constructs. Intelligence, sociability, extroversion, aggressiveness, and achievement and verbal reasoning are some examples of constructs. Tests of these functions are valid in so far as they behave in the way such trait should reasonably be expected to behave. A "theory" about a trait will lead to predictions of two types :

(i) A theory makes prediction about differences in the groups which are supposed to possess the trust in high degree and those in low or non-existent degree just as delinquents and non-delinquents, intellectually superior and intellectually inferior groups.

(ii) A theory may predict modification of a human characteristic as a result of experimental conditions or treatments.

For any test that presumes to measure that quality, we can formulate a network of theory leading to definite predictions as explained above. In so far as they are verified the validity of the test as a measure of the trait or construct is supported. In so far the predictions fail to be verified, we are led to question about the validity of our test or our theorizing or both. Evidence of construct validity is partly rational and partly emotional and judgment and these evidences join together in the validation enterprise.

6. Factorial Validity

Factorial Validity is, in a way an extension of the construct validity discussed above. The inter-correlations of a large number of tests are examined and if possible accounted for in terms of much smaller number of more general 'factors' or trait categories. Sometimes 3 or 4 factors may account for the inter-correlations among coefficients of 16 to 20 tests. The factorial validity of a test is defined by its correlation with a factor, called factor loading. A word comprehension test may correlate 0.82 with the verbal extracted from a test battery. This will be the test's factorial validity.

1.4.4.2 Factors Affecting Validity

The following are some of the factors which affect test validity. The test users should recognize factors that tend to make their tests invalid.

(i) Cultural factors such as socio-economic status, social class

structure, differential sex roles affect performance on various tests.

(ii) Response sets or test taking habits, of the examine may differentially affect the validity estimates.

(iii) An increase in the number of test items may boost up the reliability but may bring down the validity.

(iv) Difficult and ambiguous directions to pupils may render the test a measure of something different from what the author of this test intended.

1.4.5 RELATION BETWEEN RELIABILITY AND VALIDITY:

Reliability and Validity are the two important aspects of the same quality of test, called "test efficiency". Reliability is concerned with the consistency of test scores and does not go beyond the test. Validity on the other hand, implies evaluations in terms of outside and independent-criterion. A test, to be reliable, need not be practically valid while the test to be valid must be reliable. A clock which gains twenty minutes a day is a perfectly reliable instrument as it will repeat the same gain everyday. However judged against a standard time piece, the clock is not valid.

1.4.6 QUESTIONS FOR SELF EVALUATION:

Q1) Reliability means consistency of a test scores. Yes/No

Q2) Test-retest method,split half method and parellel form method are used for computing reliability. Yes/No

Q3) Reliability of the test depends upon the length of the test. Yes/No

Q4) Validity means the degree to which the test measures what it is intended to measure. Yes/No

Q5) Content validity , facee and concurrent Validity are the different types of reliability. Yes/No

Q6) Socio-economic status,differential sex roles,test taking habits,wrong directions affect the performance of various tests. Yes/No

ANSWER KEY: Yes,Yes,Yes,Yes,No,Yes

1.4.7 SUGGESTED QUESTIONS:

1. What do you mean by Reliability ? How is it established ?
2. What will be the effect of lengthening of a test on reliability ?
3. A test of 60 items has a reliability coefficient of 0.60 What will be the reliability coefficient of the test if :
   (i) The number of items is increased to 120
   (ii) The number of items is increased to 180. (Answer = 0.67 and 0.82)
4. Estimate the reliability of a test of 100 items if σ_t - 10.00-; $\sum pq = 20$?

Which method of reliability is applicable here ? (Answer = 0.81)

5. What does 'Validity' stand for ? Explain various types of Validity that a test can have.

1.4.8 SUGGESTED BOOKS AND WEB SOURCE:

1. Anastasi, Anne : Psychological Testing, New York: Macmillian, 1968.
2. Garrett, H.E. : Statistics in Psychology, and Education Bombay: Vakils, Feffer and Simons 1973.
3. Guilford, J.P. : Fundamental Statistics in Psychology and Education, New York : McGraw Hill, 1965.
4. Nunnally, J.C. : Psychometric Theory, New Delhi : Tata McGraw Hill, 1911.
5. Remmers, H.H. Gauge N.L. Rummel, J. Francis : A practical Introduction to Measurement and Evaluation Delhi : University Book Stall, 1967.
6. Stanlsey, John C. : Measurement in Today's Schools, Englewood cliffs : Prentice Hall. 1994.

1.www.comnet.com. 2,www.yoursdictionary.com

LESSON NO : 1.5 **AUTHOR : DR. S.S. GREWAL**

Structure of the lesson

1.5.1 OBJECTIVES:

After reading this lesson,the students will be able to:

1. Understand the different methods of research.
2. Describe the steps of Historical research.
3. Describe the different studies in descriptive research.

1.5.2 INTRODUCTION

Depending upon the nature of inquiry or problem to be investigated into, the social scientists employ different methods of research. For example, if some researcher is looking into historical development of Higher Education in India, he will employ a technique called Historical method of research which involves collection of form evidence, internal and external criticism of evidence gathered and finding sources, trends and so on. And if he is interested in some status as 'Opinion of school principals on 10+2+3 formula", or "Open Book examination system", he will employ a technique called survey method, and if he wants to compare the effectiveness of Alphabet method of teaching language to that of the Sentences Story method, he will have to employ experimental technique. In social

sciences, usually four such methods are employed, they are :

1) Historical method
2) Normative Survey or Descriptive method
3) Experimental Method
4) Genetic Method

Some authors also mention 'Case-study and Observation' as methods of research. But some writers feel that these are not the methods but tools of research. Here we shall discuss only three methods of research.

1.5.3 HISTORICAL METHOD OF RESEARCH

Historical research is an attempt to collect facts about past events and arrive at conclusions about educational matters. The historical researcher must systematically and objectively locate, evaluate, and interpret evidence available for understanding the past. From this evidence he hopes to show what may be contributed to by past experience to greater understanding of present situations and of what might happen in the future.

Etymologically, history has been derived from the term 'historia' which means systematic search for knowledge. In this manner, the subject matter of history forms the basis for research. History is the complete record of man including all his accomplishments.

The historical method of research utilizes acceptable procedures to determine the accuracy of statements about the past, to establish relationship, and possibly the connection between cause and effect. Good,[1] says : Viewed as research, history may be defined as an integrated narrative or description of past events or facts, written in spirit of critical inquiry, to find the whole and report it.

The process of educational research about the past generates facts and, in turn, conclusions, using techniques to interpret the remains of the past. It also evaluates and verifies statements of others. The historian cannot use direct observation or create his/her own data. The main limitations in historical research, we have no choice over somes. Whatsoever is available and comes in our way get it and interpret. The data is seldom complete and we have to rely upon incomplete data most of the time.

1.5.3.1 STEPS IN HISTORICAL RESEARCH:

Historical inquiry uses a process similar to that of other scientific inquiry. The historical researcher's activities can be identified with the following steps :

1. Identification and defining of the problem.
2. The formulation of hypotheses.
3. Collection of evidence from primary and Secondary Sources.
4. Criticism of historical data.

5. Organization of facts into logical cause and effect sequences.
6. Interpretation of facts.
7. Research conclusions.
8. Report writing.

1. Identification and Defining of the Problem

In historical research, it is important that the student carefully defines his problem and appraises its appropriateness. Historically based problem can be delimited or isolated along any of the four dimensions : time, geographic area, persons involved and activities. The scope of the study may be expanded or reduced by control on one or more of these categories. When a too broad question is asked; the researcher is likely to be forced to decrease the scope. Many problems are not adaptable to historical research methods and cannot be adequately treated using this approach, whereas some other problems have little or no chance of producing significant result either because of the lack of pertinent data or because the problem is trivial one.

2. The Formulation of Hypotheses.

A fact that many students fail to realize that historical research usually requires the setting up of specific and testable hypotheses. Without such hypotheses historical research often becomes little more than an aimless gathering of facts. When the researcher has some knowledge to provide a historical content for a proposed answer to the question being studied a hypothesis should be developed. The value of the historical hypothesis rests in its facility to provide something to refuse or support, and its contribution makes the drawing of conclusions more objective.

3. Collection of evidence

All the information recorded in the past plus everything else which remains of the past is available for historical research. The task of the historical researcher is to gather appropriate information, place it in the context of the historical period under study, and interpret its relationship to the logically conceived hypothesis in regard to the support of the hypothesis.

Historical sources may usually be classified as either documents or relics. Documents include a wide range of written and printed materials, recorded for the purpose of transmitting information. Relics include physical objects related to the period being studied. In education, these might include such materials as school furniture or various physical objects used in teaching, such as flash cards, multiplication tables, and so on.

Each of these sources is available to the researcher in different ways : the documents for their objectivity and directness in reporting and the relics for their transmission or information in interpretation. Documents present data directly in a straight forward manner that requires no interpretation. Relics

have given scientists an opportunity to reconstruct past civilizations with accuracy, although many people of the past left no documents. For example, the daily activities of inhabitants can be reconstructed through the utensils they used to obtain and prepare food in. Also games and skill indicate something about the leisure time. Other glimpses into early cultures are provided by what they left by way of relics.

The use of information as a fact depends on the source of the information and how well it withstands a test of its authenticity and trustworthiness. In historical research, primary sources are generally defined as those documents in which the individual observing the event being described was present. The best evidence about the past is provided by primary sources 1) the relics that remains of the past 2) documents, 3) reports of eye witness and happenings of the past. Primary sources are considered as original having only the observer or eye-witness coming between the information and the user of the information. Also useful to the historian are secondary sources : accounts of the events that were not actually experienced by the reporting authority. Secondary sources are those in which the person describing the event was not present but has obtained his description from someone else who may or may not have directly observed that event. If a newspaper reporter attends a city council meeting and writes a report of the meeting in an article, the article is a primary source. If he interviews those who participated in the meeting and writes an article, the article is a secondary source since the information is interpreted and reported in the mind of a non-observer. Sources are consciously or unconsciously left behind. The personal diary or description of a quarrel between two kings if recorded by a historian is a consciously left source. A king's crown or gown is an unconsciously left source.

4. Criticism of historical data

A necessary part of the process of historical research is the subjection of the material to criticism. Criticism is a scientific investigation of collected information in regard to matters of origin and validity. It is the process of establishing the authenticity of source or authorship. They are known as external criticism and internal criticism respectively.

i) External Criticism : External criticism concerns with the authenticity of the information, is it genuine ? The focus of external criticism is not on the meaning attached to the evidence, but rather on establishing, why, where, when how and/or by whom the document or relic was created. The purpose of external criticism, lies not merely in detecting frauds, but it also lies in establishing the source. If a historian comes across a letter claimed to have been written by Ghalib, he must first make sure that it was written by Ghalib and that the whole of it is genuine.

ii) Internal Criticism : Internal criticism concerns with the meaning and trustworthiness of evidence. What meaning can be attached to the isolated facts ? The focus of internal criticism is not the object or evidence itself but the meaning attached to it. After the process of external criticism has established the authenticity of information, the credibility of the information is checked by internal criticism of the contents.

The general reputation, integrity and competence of author are of particular importance in determining the trustworthiness of the content of historical document. The historian must also attempt to apprehend the motives of the writer in order to make sure that he was not biased. The circumstances under which the document was written is an equally important consideration in deciding the correctness of the narration.

The validity of a historical fact can sometimes be verified by comparing it with the statement of others. Thus, it is an exacting task, calling for a high level of scholarship. Criticism is both negative and positive. When we proceed to prove that the data are false, the criticism is called negative and when we try to prove that data are genuine, the criticism is called positive.

5. Organization of facts into results

The several facts that have been validated for a study must be placed into some logical arrangement to allow for the collective meaning to the results of the study. The notes and establishment fact should already be on bibliography cards, only one fact per card.

6. Interpretation of Facts

Since facts by themselves have little meaning after they are determined and organized, they must be interpreted. The interpretation of the facts is the analysis of the data. The actual analysis is a study of the facts in regard to how they support or reject the hypothesis, the interpretations drawn from the evidence are based on what happened in the past but, more importantly, also include why and how it happened. The historian must deal with the representativeness of his limited data in the interpreting ideas and events i.e. are they unique or are they representative ? The historian must be on his guard in his use of analogy as a source of hypothesis or as a fame of reference for interpretation. His endeavour is to relate facts in the cause and effect sequence and make this effort synchronize with the hypotheses.

7. Research Conclusions

The outputs of the analysis of the fact are the research study's conclusion. Conclusions are drawn to show the reader what the study has revealed. The historian's goal is not only to establish facts to determine trends which the data may suggest, but also generalizations which can be derived from the data.

8. Report Writing

It needs the highest level of scholarship to write the historical report. The discontinuous and incomplete nature of historical evidence places a burden on the ingenuity and insight of the researcher in providing the required continuity. The historian is expected to fill the gaps, if any, in his or her knowledge and to create anything that the story needs, using historical knowledge as a guide. He/ she must deal not only with the established facts but, at the same time, must build continuity through insight. Producing a report of historical research requires the highest level of writing expertise.

The research hypothesis can give direction to what is relevant to the study. It also furnishes a basis for analysis and for conclusions, provides an agent for the synthesis that follows the criticism and validation of the evidence, and serve as a vehicle to build continuity into the writing. Historical researcher can provide a perspective to many educational problems in which we must make constant efforts to take decisions.

Historical research describes "what was". The process involves investigating, recording, analyzing and interpreting events of the past for discovering generalizations that are helpful in understanding the present and in predicting the future. The focus of historical research is on change, growth and development of individuals, groups practices, movements, institutions or ideas. Historical research attempts to go back to primary, original, or first hand sources of informations.

1.5.4 DESCRIPTIVE METHOD OF RESEARCH

A descriptive research study is designed to obtain information about the current or present status of a phenomenon. The descriptive study is directed towards determining the nature of a situation as it exists at the time when the study is being conducted. There is no manipulation, control, or administration of variables or treatments as is done in experimental research. The purpose of descriptive research is to describe "what exists" with respect to the variables or conditions in a situation. Generally, descriptive research is not directed towards hypothesis testing. A school principal may be interested in knowing the number of students likely to be enrolled in class XI during the next session so that effective use of school laboratories and other facilities may be planned. He is not interested in studying class XI enrolment as a variable in relation to some other variable.

1.5.4.1 STUDIES IN DESCRIPTIVE RESEARCH:

1) Case Study 2) Surveys 3) Development Studies 4) Follow - up studies 5) Documentary analysis 6) Trend Analysis and 7) Correlational studies. Sometimes these studies may be used for hypothesis testing, still they are generally classified as descriptive research.

1. **Case study or Case History :** A typical case study is an intensive or in depth study of one single individual or one small unit such as a school, group, or a family. In case the researcher gathers detailed information about individual's past experiences, present status, his environment and how these factors interrelate with one another. The case studies of Sigmund Freud of individuals with personality problems led to the development of theory and practice of psychoanalysis. Case studies of Jean Piaget regarding maturation and development of intellect and thinking are very famous and they led to the formulation of a Piagetion theory of thinking process. Cyril Burt's collection of 1000-case histories throw a flood light on delinquent behaviour.

   Case studies often helps in gaining valuable insights into the dynamics of behaviour. We may gain access into hidden information which may otherwise not be apparently available or evident.

   The main limitation of case study is that sometimes it becomes difficult to confirm or refute observation and findings regarding individual cases. Then, there may be difficulties in generalizing the findings.

2. **Surveys :** Surveys gather limited information from a large no. of individuals. A typical survey question may be : how many pucca rooms do the primary schools have on the average in Punjab ? Surveys can be used not only for describing existing conditions but also comparing these conditions with a pre-determined criteria for evaluating the effectiveness of programmes. Quite often there is a feeling that survey studies are poorly conducted with regard to samples taken and the questions asked to the respondents. Respondents tend to become casual to the questions asked or information sought. But this is a commentary on the poor, management, rather on the method of research. Pre poll and Past poll surveys are very popular in these days.

3. **Development studies :** Since education deals with children and human beings, it is important to study their physical, intellectual, emotional and social growth and development. School curriculum, teaching methods, school plants etc. must be based on the developmental characteristics and needs of the children.

   Developmental studies may be carried out through longitudinal method or cross-" sectional methods. In the longitudinal method the same group of individuals is studied over a span of time, e.g. physical growth of a group of children may be studied every six months over a period of five years. The limitation of the longitudinal method is that sometimes it becomes difficult to maintain contact with the subjects over an extended period of time. Another difficulty is that no new variables can be introduced once the

study has been carried out for some time. Similarly maturity of subjects or familiarity with the task creates problems. In the cross-sectional method, individuals of various age groups are studied at the same time. For example, we may study physical development of children of age group 5, 6 and 7 years instead of studying the same group of individuals when they are of ages 5 years, 6 years, and then 7 years. Disadvantage of the cross-sectional method is that chance, difference in samples rather than true differences may distort the results.

4. **Follow-up Studies :** Follow-up Studies are similar to longitudinal studies. L.M. Terman's famous study of the gifted for a period of 40 years was based on this approach. The subjects were studied in 1921-22, six years afterwards and then again in 1936, 1940 and 1945. Gifted children matured into gifted and successful individuals.

   The limitations of the follow up studies are similar to those of longitudinal studies. It is difficult to maintain contact with the subjects over an extended period of time.

5. **Documentary Analysis :** It involves analysis of documents and records. E.L. Thorndike used documentary analysis to identify most commonly used words in the English language. The identified words were incorporated in the language text books. Documentary analysis is also referred to as content analysis. It can also be used to study psychological and sociological variables.

6. **Trend analysis :** In order to plan for the future, school need to project the demands that will be made on their services in future. Through surveys at repeated intervals, rate and the direction of change is studied and these trends are used to predict future status. For example, a school in the lower middle class locality may have served 300 students in 1980, 500 in 1990 and may be serving 600 students in the year 2000. The trend analysis of prices and population can be helpful in educational planning.

7. **Correlation Analysis :** "This technique is used for determining the relationship between variables. For example, in a simple two variable study it may be hypothesized that there is a relationship between performance in Mathematics and Physics at the Senior Secondary level. This relationship can be studied through determining a correlation between the performance of a group of senior secondary students on two subjects Mathematics and Physics.

   A Correlation coefficient indicates the association or relationship between two variables. The magnitude tells us about the degree of relationship and the sign of the correlation tells us about the direction of the relationship. A positive relationship tells us that the change in one variable is accompanied by a change in the other variable, in the same direction. A

negative correlation indicates that change in one variable is accompanied by change in other variable in the opposite direction. The first variable may be caused by the second or vice versa. There is also a possibility that relationship between the both is caused by third variable. A correlation co-efficient can have any value between +1 to -1. In correlational studies it can be assumed that if a relationship exists, it will be evident in samples of 50 to 100 cases. Typically, correlational studies do not need large samples. The basic method of determining correlation co-efficient between two variables is the Pearson's Product Moment Method. Correlation techniques may be employed to either hypothesis generating or hypothesis testing studies. In hypothesis generating studies, measures of a large number of variables are taken and correlation co-efficient are worked out of these measures to find out which of the variables are related. In a hypothesis testing study, the investigator has a basis for expecting a relationship between the variables. For example, it may be hypothesized that there is a relationship between intelligence and academic achievement.

The steps of descriptive research are selection of the problem statement & definition of the problem, selection or development of tools, selection of sample for collection of data analysis and interpretation of data and writing of the research report.

1.5.5 QUESTIONS FOR SELF EVALUATION:

Q1) Historical research is an attempt to collect facts about past events and arrive at conclusions about educational matters. Yes/No

Q2) Historically based problem can be delimited to time, geographic area, persons involved and activities. Yes/No

Q3) A discriptive research study is designed to obtain information about the current or present status of a phenomenon. Yes/No

Q4) Case study, Surveys, Documentary analysis and Trend analysis are the studies used for testing hypothesis in descriptive research. Yes/No

Q5) Experiment is a series of observations carried under controlled conditions for testing a hypothesis or to establish cause and effect relationship. Yes/No

Q6) In the parallel group design,instead of repeating the experiment on the same group,the experiment is conducted with the help of two groups. Yes/No

ANSWER KEY: Yes,Yes,Yes,Yes,Yes,Yes

1.5.6 SUGGESTED QUESTIONS:

1. Take an imaginary research problem which will require the use of historical method. Discuss important steps involved in Historical Method.
2. Show your acquaintance with normative Survey Technique. Give examples in support of your answer.
3. "The Parallel group method is an improvement over the simple group method of experimentation", Discuss.
4. The experimental method is the only scientific method of research. Explain

1.5.7 SUGGESTED BOOKS AND WEB SOURCE:

1. Best. John W., Research in Education, New Delhi: Prentice Hall of India Pvt. Ltd. 1977.
2. Koul, Lokesh, Methodology of Educational Research, New Delhi: Vikas Publishing House Pvt. Ltd. 1984.
3. Meenakshi, First Course in Methodology of Research, Patiala : Kalia Parkashan, 1992.
4. Sharma, T.R. Vidyak Khoj Vidhi, Patiala: Punjabi University.
5. Travers, Robert M.W., An Introduction of Educational Research, New York: Mac-Millian Publishing, Co., Inc. 1976.
6. Van Dalen, Deobold B., Understanding Educational Research, New York : Mc-Graw Hill Book Co. 1973.

1. www.noteaccess.com, 2. www.esortment.com

LESSON NO : 1.6 **AUTHOR : DR. S.K. BAWA**

Structure of the lesson

1.6.1 Objectives :

After studying this lesson the students will be able to :-

(i) Explain the experimental design.
(ii) understand the significance of a research proposal.
(iii) to know the steps of writing a proposal.
(iv) to be able to prepare time and financial budget for a research project.

1.6.2 Introduction :

In this chapter we will discuss experimental research method i.e. experimental research and Research proposal and report. The most difficult phase of research project is preparing a research proposal. It provides a basis for evaluation of research and gives direction to conduct research in a scientific and objective manner. It also provides a systematic plan of procedure to be followed by the researcher. A good research work requires a well-designed proposal which is carefully planned and systematically carried out. In nut shell, a research proposal

deals with problem or topic that is to be investigated. It has a variety of formats which vary in their length. A research proposal includes an introductory section, objectives, hypotheses, assumptions, method of research, tools, justification and implication of study, a chapter scheme to be followed for writing a report and references. It is written in present or future tense end submitted for approval before starting the actual research work. It is a planning phase of a research work which is produced in the written form to judge its worth.

1.6.3 EXPERIMENTAL METHOD OF RESEARCH

Experiment is a series of observations carried under controlled conditions for testing a hypothesis or to establish cause and effect relationship.

The most directive way to establish cause and effect relationship is by a carefully devised experiment. The investigator attempts to find out what is related to what, and by control of the variables, what cause is associated with what effect. Despite it scientific rigour, experimentation is only one aspect of the scientific method. Scientific methods involves a great number of activities of which experimentation is simply one important form.

From the operational point of view, it is variation in the independent variable in order to study the changes in the dependent variable under controlled conditions. The experiment is generally regarded as the most sophisticated research method for testing hypothesis. Experimental method has four main components :

1. Control : The method depends upon the assumption that if two situations are identical in every way and if one variable is added to or subtracted from one and not the other, the difference occurring in the two situations is the consequence of the variable added or the variable subtracted. This is called Mill's law or law of single variable. In an experiment, only one variable is allowed to operate, while all other interfering variables are controlled.

2. Manipulation : The variable whose impact is to be seen on other variables is called independent variable and the variables on which effect is to be seen are called dependent variables. Such variables which are not under study but interfere with the variables unnecessarily are called interfering variables. Their effect is controlled eliminated or atleast minimized. The independent variable is manipulated to study its effect on dependent variable.

The method depends upon comparing the experimental group with the control group. Of course both these groups have to be equal in every way This is done by using the following techniques :

a. Random assignment
b. Ramdomized matching
c. Homogeneous selection
d. Analysis of covariance.

Random Assignment

Subjects are assigned to experimental group and control group at random.

Randomized Matching

An alternative procedure for assigning subjects to groups is to match individual subjects with as many extraneous variables as one can think of and then assign subjects of each set to the different groups at random.

Homogenous selection

Another method that can be used to make groups comparable on an extraneous variables involves selecting all the subjects only at a particular level of that variable e.g. if age is an extraneous variable, all subjects may be taken of one age level only. But this will restrict the generalization of the result of the experiment to this age level only.

Analysis of Covariance

Another form of control is the statistical method called Analysis of Covariance. This method can be used when it is not possible to assign subjects at random or with the use of matching procedures. The allowance for the initial difference in an extraneous variable, a record of which is kept, is made statistically at the end of the experiment, while analyzing the results.

A distinguished characteristics of experimental research is the manipulation of independent variables. Some examples of independent variables in educational research are methods of teaching, intelligence, sex, social class, attitudes, types of motivation, personal types, class room environment etc. Although they may function as independent variables, obviously not all of them can be manipulated by the experiment. Their categories which are already present in the society have to be utilised in the experiment, which in these circumstances comes to be known as 'ex-post facto research'. Variables, like methods of teaching, can be manipulated by the researcher but those like sex and personality types cannot be. He/she has to take them from the society ready made.

3. Observation : In an experiment, the researcher is interested in the effect of the manipulation of the independent variables on a response variable. Observations of dependent variables are made with respect to manipulation of independent variables. The dependent variables are generally interested in explaining or predicting achievement.

4. Replication : In experimental research findings be replicated to be sure about the research. The experiment is repeated and findings are compared.

1.6.3.1 STEPS FOR THE EXPERIMENT METHOD:

1. Selecting and delimiting the problem
2. Reviewing the literature
3. Drawing up the experimental design
4. Defining the population and sample

5. Selection of the tools
6. Carrying out the experimentation
7. Measuring the outcomes
8. Analyzing and interpreting the outcomes
9. Drawing up the conclusions
10. Reporting the results

1.6.3.2 EXPERIMENTAL DESIGNS:

Experimental design refers to the conceptual frame work within which an experiment is conducted. Experimental Designs may be classified in various ways. According to the traditional mode of classification, there are three basic experimental designs, namely, the single group design, the parallel group design and rotation group design.

1. The single group design

The single group design usually involves three steps. The first step is the administration of a pre-test for measuring a dependent variable. The second step is the application of the experimental treatment (independent variable) to the subjects, and the final step is the administration of a post-test measuring again the dependent variable. The difference in means of the pre and post-test scores is supposed to indicate the difference by the application of the independent variable.

Let us assume that a teacher wants to evaluate the effectiveness of a new technique of teaching on fourth grade social studies students. At the beginning of the year, he administered a standardized test that appeared to be a good measure of the achievement of objectives of fourth grade social studies. He introduced the new technique, and at the end of the year, administered the standardized test second time. He compared the scores from the first and second administrations of the test in order to determine what difference the new technique of teaching had made.

The major limitation of the single group design is that no control group is used, the experimenter cannot assume that the change between the pre-test and post-test was brought about only by the experimental treatment. There is always the possibility that some extraneous variables brought about all or part of the change. Thus, this design is lacking in internal validity.

Advantages

1. It is more useful than others.
2. The planning and activation of this is easy and simple.
3. Equalization of group of changing is not needed.
4. It is applicable in classroom.
5. It works as a stimulus for better teaching.

Limitations.

1. Influence of growth and maturity.
2. Influence of transfer or training.

3. Influence of experiment and knowledge.
4. Influence of skill factor and practice.

2. The parallel group design

In this design, instead of repeating the experiment on the same group, the experiment is conducted with the help of two groups. One of the groups is called the experimental group and the other is known as the control group. These groups are made equivalent either by randomization, or by matching of pairs or by matching for means and standard deviations on some variable(s) related to the criterion variable (i.e. the dependent variable). The working of the experiment can be understood by the aid to the following paradigm.

	Experimental group	**Control group**
Step I	Pre-test	Pre-test
Step II	Experimental factor	Control factor
Step III	Final test/Post-test	Final Test/Post-test
Step IV	Comparison of gains in both the groups is done	

Educational researchers are often faced with situations in which it is not possible to select subjects at random from the population of interest. In such cases, one must begin with the available subjects and randomly assign the conditions. It restricts external validity of the study. The weakness can be partially overcome if the study is replicated with different groups and in different settings. If the results of the replication agree with the original findings, they provide evidence of external validity.

Advantages

1. It is free from the weakness of one group method.
2. Both the control and the experimental factor or factors are applied simultaneously on different groups so that the errors due to the maturation and practice effect do not arise.

Limitations

1. It suffers from the difficulty of equating groups and controlling significant variables.
2. There are several administrative problems for selecting subjects from schools.
3. The sampling error cannot be minimised by this procedure.

3. The rotation group design

When the experimental and the control groups are only approximately equivalent the rotation group design may be used to iron out the difference in the two units of teaching material of the same subject of equal length, complexity and difficulty and two separate tests for each of the units and these two equal length and with items of equal difficulty may be taken for an educational experiment to test the relative efficacy of two methods of teaching by this method.

	Experimental Treatments	
	Method A	Method B
Unit X	Pre test (X1)	Pre test (X1)
	Group 1	Group 2
	Post-test (X2)	Post Test (X2)
Unit Y	Pre test (Y1)	Pre test (Y1)
	Group 2	Group 1
	Post-test (Y2)	Post Test (Y2)

Means of gains in the two above cells = M_A Means of gains in the two above cells = M_B

In an educational situation, it may not be possible to divide the classes and equalize two groups and the above design may prove helpful in rotating out differences by subjecting both the groups to the methods alternatively and taking the means of the gains under the method by adding the total of the gains of both the groups under each of the methods.

The gains are computed by subtracting the scores of the test for the units administered before teaching (pre-test) from the scores on the same tests after teaching (post-test). The relative efficiency of the two methods may be judged by comparing the means M_A and M_B. The short comings of this design are :

1. It may not be easy to locate equivalent of units, material and to construct equivalent tests on them.
2. There may be carry over effect from one experimental treatment to another.
3. It is boring to the students to undergo so many tests.

4. The Factorial design

A factorial design is one in which two or more variables are manipulated simultaneously in order to study the independent effect of each variable on the dependent variable as well the effects due to interaction among the several variables.

Factorial designs have been developed at varying levels of complexity. The simplest factorial Design is the 2 by 2. In this design, each of the independent variables has two values.

Advantages

1. This design is highly applicable in educational experiments because in educational situations the more chance of contamination effect are there because subjects are human beings.
2. Other experimental designs can control only basic types of errors but not the contamination effect, which can be eliminated by this design only.

3. In this design 'G' error has greater scope which can be eliminated by repeating the experiment and more over the group variation is eliminated at analysis stage.

5. Analysis of Variance design

In this case, two or more variables are manipulated so as to test their effect on the dependent variable by using the statistical technique of analysis of variance. Lindquist has given six basic designs. The designs are more scientific and modern.

A critical point in conducting scientific experiments is that they must be well-planned in advance, and then each step of the process should be followed as accurately as possible.

A general open minded attitude towards problems be adopted and given a concrete shape in terms of a carefully controlled experiment. This is very much possible in both natural and social sciences.

Advantage :

More variables can be studied at a time. This design can also find the effect of interaction.

Drawbacks of Experimental Method

1. It is time consuming and costly method.
2. Subject react to the whole situation and not only to the stimulus.
3. Experimental method is used when independent variable can be manipulated. But many independent variables such as intelligence and home environment cannot be manipulated.
4. Every teacher does not have competency/expertese to conduct experiment.

1.6.4 Need of writing a research proposal :

Preparation of research proposal is needed because of the following reasons.

* To get the direction of conducting the research.
* To provide basis for evaluation of research work.
* To know the feasibility and significance of the problem to be researched.
* To get the vision of completion of research work in terms of time, money and resources.
* To gain confidence and reassurance of courage and determination of conducting research.
* To submit it to the funding agency for carrying out the work.
* To identify the level of research.
* To know the problems to be coped up during research process.
* It provides a systematic plan of procedure to the researcher to follow.

1.6.5 Types of Research Proposal :

On the basis of the type of research study and the purposes, the research

proposal can be divided into three categories.

(i) Master's or doctoral level research : Such type of proposal is submitted to get master's or doctoral degree. It is the general requirement of the university. It is just the beginning of research. The proposal is evaluated by the committee of experts to determine the feasibility of the study.

(ii) Research Projects : Such type of research is conducted by the university or college and the proposal is submitted to the research organizations for perusal, such as U.G.C., ICSSR, NCERT, RCI and such other organizations. There are funding agencies and provide funds to conduct research on thrust areas. The proposal submitted to them is evaluated by their experts and then projects are allotted to the investigators.

(iii) Research proposals for private and government funding agencies : There are private and govt funding agencies who give financial assistance to the researchers. A panel of experts evaluate the research proposal and on their recommendation, financial assistance is provided.

1.6.6 Preparation of Research Proposal :

The research proposal starts with the introduction of the title, its theoretical base and significance.

(i) The Title : The title is usually a declarative statement but may be in question form which is focused on a stated goal and gives direction to the research process. The title should not be burdened by superfluous words, not lengthy or too short. It is to be worded in such a way that it gives sufficient information about the nature of the study.

(ii) Statement of the Problem : It differs from the title as it is an attempt to focus on a clear goal and gives direction to the research process. It suggests answers or conclusion and provide a focus for research activity. No doubt, it is an expansion of the title but gives the detailed information about the vision of the research.

(iii) Review of Related Literature : The conceptual and research literature is reviewed and presented in the proposal to show the roots of the problem in the existing literature. It gives the evidence that researcher is familiar with the known, still unknown and delated research problems. It helps to avoid duplication and formulate hypotheses. It gives direction to select appropriate procedure for work. The studies which are relevant, competently executed and clearly reported should be included. It is valuable guide to define the problem and recognize its significance.

(iv) Significance of the Problem : The research must include in the proposal the worth of spending time, money and effort for carrying out the work. Its implications must be included in the proposal. The influence of results on theory or practice must be indicated or it should justify the gaps which exist in the knowledge provided by the previous studies and the way by which

the present study will fill those gaps. It should also be related to the existing social issues and to the priority areas of educational research.

(v) Definitions, assumptions, limitations and delimitations : The proposal defines the terms to be used in the study theoretically as well as operationally to avoid misinterpretation. The beliefs of the researcher should be given in the form of the assumption and limitation and the condition which are beyond the control of the researcher and may place restrictions on the conclusions of the study and their application to other situations. Delimitations are the boundaries of the study which are set keeping in mind the type of study, its purpose, financial and time constraints.

(vi) The Hypotheses : The hypotheses are formulated to give direction to the data gathering procedure and it also clarifies the nature of the problems. A good hypothesis is formulated on the basis of the implication of related literature. It should be reasonable, consistent with known facts or theories, and stated in simple terms.

(vii) Method and Procedure : It includes the details of the population and sample of the study. The size of sample, area, sampling technique are selected on the basis of type and purpose of research. The procedure gives the outline of the research plan. The method of research to be followed, how and what is to be done. Specification of the tools to be used in the study. Above all, the statistical procedure to be followed.

(viii) References : In the end, there is a list of books, journals and other documents that one has used is selecting problem and may use for carrying out the research.

(ix) Time and Budget Schedule : A realistic time schedule is prepared by dividing the study into phases and assigning dates for completion of each phase. It helps the researcher to use the time constructively, efficiently and systematically. Similarly, the proposals which are submitted to funding agencies, a budget schedule is also prepared giving estimate of expenditure in data collection, processing, purchasing equipment, tools, books and contingency etc.

(x) Chapter Scheme : A tentative scheme of chapters to be included in the research report is also listed the end of the research proposal.

1.6.7 Format of the Research Proposal :

The format of the research proposal is as given below :

1. Title
2. Introduction
3. Significance of the problem
4. Statement
5. Definitions of the terms used
6. Objectives
7. Hypotheses

8. Method of Procedure
   (a) sample and sampling procedure
   (b) tools
   (c) research design
   (d) statistical strategy
9. Delimitations
10. Time and Budget Schedule
11. Chapter Scheme
12. References

1.6.8 Format of the Research Report : For the purposes of presentation of a research report as following :

(a) Preliminary Section
   (i) Title page
   (ii) Preface,including acknowlegement
   (iii) Table of Contents
   (iv) List of Tables
   (v) List of figures, maps of illustration.

(b) Introduction
   (i) Statement and definition of problem.
   (ii) Analysis of previous research
   (iii) Relation of present problem to theoretical position of the previous research.
   (iv) Statement of specific research objectives.
   (v) Significance of the problem.
   (vi) Delimitation of the study
   (vii) Assumptions underlying research hypotheses
   (viii) Statement of research question
   (ix) Operational definitation of important terms.

(c) Design of the study
   (i) Procedures employed
   (ii) Sources of data
   (iii) Data gathering instruments
   (iv) Sampling and methods of gathering data.

(d) Analysis and interpretction of the data
   (i) Text
   (ii) Tables
   (iii) Figures

(e) Summay and Conclusion
   (i) Brief restatement of problem and procedures
   (ii) Principal findings and Conclusions with their Practical implications.
   (iii) Suggestions for further Research

Reference Section

(i) Bibliogaphy
(ii) Appendix
(iii) Index

1.6.9 Summary :

A research proposal is like a blue print of a building plan. It is carefully planned and systematically carried out. The research proposal includes an introductory section, objectives, hypotheses, assumptions, method of research, tools, justification and implication of study, a chapter scheme to be followed for writing a report and references. It is written in present or future tense and submitted for approval before starting the actual research work. It is a planning phase of research work. It is needed to get the direction of conducting the reseach, to provide basis for evaluation of research work, to know the feasibility and significance of the problem to be researched, to get the vision of completion of research work in terms of time, money and resources, to gain confidence and reassurance of courage and determination of conducting research, to identify the level of research and to know the problem to be coped up during research process.

1.6.10 Suggested Questions :

1. "The Parallel group method is an improvement over the simple group method of experimentation", Discuss.
2. The experimental method is the only scientific method of research. Explain
3. What is research proposal?
4. Explain the need of preparing a research proposal.
5. Discuss the types of research proposals.
6. Briefly write the format of a research proposal.

1.6.11 References :

Neuman, W.L : *Social Research Method,* Boston : Allyn & Bacon Publishers(1991).

Best, J.W. : *Research in Education,* New Delhi : Prentice Hall of India(1992.

Koul, Lokesh : *Methodology of Educational Research,* New Delhi : Vikas Publishing House Pvt. Ltd(1998).

Cohen L.: *Research Methods in Education,* London : Routledge Palmer(2002)