# Educational Research Evaluation and Monitoring

*Training Modules*

*Edited by*

*Dr. C. G. Venkatesha Murthy*
*Prof. D. Basavayya*

**Regional Institute of Education**
**Mysore**
**2008**

# *Foreword*

The present training modules relate to Educational Research, Evaluation and Monitoring is an outcome of the request of the Government of Karnataka and concerted efforts of faculty of our institute as well as some outside experts. This material has been prepared keeping in view the needs and requirements of the faculty of DSERT, CTEs and DIETs.

The material development activity has used different modes. Some materials are also taken from our earlier efforts of our colleagues at RIE, Bhopal too. The task of coordinating this material development has been undertaken by Dr.C.G.Venkatesha Murthy, Reader, Department of Education, of our Institute. He has been professionally supported by Prof. D.Basavayya, a retired professor of Mathematics, Prof. M.D.Usha Devi, of the Institute for Social and Economic Change, Bangalore, Dr. Vishwanathappa, and Dr. Anil Kumar of our institute. I congratulate and thank all the contributors.

The purpose of this material preparation will be served if it is put to use by the faculty members of DSERT, CTEs and DIETs. It is hoped that the present material will be found useful. Any suggestion for improvement will be appreciated.

Prof. G.T. Bhandage
*Principal*

# *Preface*

All professional organizations in education have a pleasant responsibility of updating themselves. In this sense one's clarity and dexterity about research, evaluation and monitoring can indeed enable an institution as well as the individual professionals to subject all educational phenomena and programmes to a scrutiny and get a perspective which can inform and influence not only one's functioning but also guide policy changes and perspectives. From this view point, it is important that institutions like the DSERT, the CTEs and the DIETs take proactive role in preempting some of the imminent problems and issues through research, evaluation and monitoring activities so as to enable the system to navigate smoothly and meaningfully.

In the above background, an attempt is being made by this group of the Regional Institute of Education in collaboration with external resource persons in developing the present training modules. The Coordinator is indeed very grateful to Prof. D.Basavayya, a retired Professor of Mathematics, Prof. M.D. Usha Devi of Institute for Social and Economic Change, Bangalore, Dr. G. Vishwanathappa, Reader, and Dr Anil Kumar, Lecturer, of the Department of Education, Regional Institute of Education, Mysore for their contribution and collaboration. The Coordinator is also thankful to Dr. Sabita Pattanaik, currently, Reader, Department of Education, Regional Institute of Education, Bhubaneswar, for her kind approval to use some of the write ups she has brought out in her similar effort at Bhopal. Realizing the fact that there was no point in reinventing the wheel, a couple of write ups are taken from her material with her permission, which is acknowledged sincerely. The coordinator will fail in his duty if he does not acknowledge the contribution of Mrs U. Vijayalakshmi, TGT, English, Kendriya Vidyalaya, Kathmandu, Nepal for her association with the materials by way of her language editing, which she did through e mail facility in such short notice of time.

The Coordinator is indeed thankful to Prof. G.T.Bhandage for his constant encouragement. Thanks are also due to Prof. B.S.Upadhyaya for providing logistic support. Thanks are also due to all those who are directly and indirectly responsible for the preparation of this material.

It is hoped that the present material will serve the purpose for which it is intended.

C.G.Venkatesha Murthy
*Coordinator*

# Contributors

1. Dr. D.Basavayya
   (Retd) Prof of Mathematics
   Institute of Education
   Mysore

2. Dr. M.D.Usha Devi
   Prof. and Head, CHRD,
   Institute for Social and Economic Change
   Nagarbhavi
   Bangalore

3. Dr.C.G.Venkatesha Murthy
   Department of Education
   Regional Institute of Education
   Mysore

4. Dr.G.Vishwanathappa
   Reader, Department of Education
   Regional Institute of Education
   Mysore

5. Dr. Anil Kumar
   Lecturer, Department of Education
   Regional Institute of Education
   Mysore

**Acknowledgements:** In Part A, some materials are taken from; Dr. Sabita Patnaik's *"Handbook of research methodology for the CTE and IASE faculty"*, prepared at Regional Institute of Education, Bhopal.

# Contents

# Part A

# Educational Research

# 1: Introduction to Educational Research *

**Objectives:** After going through this module, you will be able to understand;

- The meaning of educational research
- Types of research
- How to write a research proposal
- The meaning of sampling and its use in an educational research
- The meaning and method of writing a research report

## 1.1 Introduction

Teaching learning process provides ample opportunities for research. Teachers, both within and outside the classroom are faced with varied problems, which may require scientific and systematic inquiry for solving the problems in the most effective and satisfactory manner. Most of the teachers do not believe that research is and should be an integral part of their day to day duties. They usually dislike research or feel threatened by its probability because they lack technical competence to carry it out. This is because research has been usually presented as something that scientists can only do. Non-scientists may require doing it from time to time as and when they want to use scientific methods.

Research in any discipline dives deep into the ocean of knowledge to discover new relations, new theories and better understanding of various phenomena. Besides the basic role of creating a fund of knowledge, research also acts as an agent of change and improvement.

There has been a lot of discussion on reaching and utilizing the research findings at the field level. But the chasm between researchers and practitioners still remains. Nothing seems to percolate down to the classroom level, in spite of the plethora of researches going on at the university level and those done by research and training organizations.

The teacher in school, by and large, is busy, 'covering' the syllabus and improving the students' achievement in Board examinations. In these examinations degree/certificate dominated the system. Research, innovation and experimentation hardly catch the imagination of teachers. They have neither time nor enthusiasm.

However, in the National Policy on Education (NPE) 1986 a lot of importance has been give to research. Financial outlay for educational research has been substantially enhanced and frequent conduct of research seminars has become the order of the day.

---

* This material is taken from the work of Dr. Sabita Patnaik's *"Handbook of research methodology for the CTE and IASE faculty"* , prepared at Regional Institute of Education, Bhopal.

## 1.2 Types of Research in Education

Human beings desire to know more about their world that has led them from primitive superstitions to modern scientific knowledge. From mysticism, dogma and limitations of unsystematic observation based upon personal experience to the process of thinking, which itself has to develop the method. Although first applied as a method of physical science, the process of scientific inquiry has also become a prevailing method of the behavioral sciences. The research and scientific methods are mostly used synonymously and research is considered as a nonformula, systematic and intensive process of carrying on a scientific method of analysis.

The present era of a rapid change has created many social problems such on knowledge explosion, population explosion, international tension, automation, changing values of younger generation etc. In this context, the role of education in a dynamic society is to aid in solving social problems implied above and contribute towards the reconstruction of society. For this purpose, the educators have to carry out effective educative process by planning proper and required learning experiences to guide the students in positive direction and organize a better school system. As a professional, with his knowledge and skill the educator has to make valid decisions about what to do and how? Therefore, they have to know what the right answer in a particular situation is. Although there are other sources of knowledge, such as experience, authority and tradition, it is scientific knowledge about educational process that makes the most valuable contribution to decision making in education. This fund of knowledge has been made available to education as a result of scientific inquiry into educational problems. Therefore, the purpose of research is discovery of generalizations relating to various aspects of education and the educational research has to determine:

- What characteristics of education have been there in the past?
- What characteristics of educational practice and conditions are in the present and
- What effects of introducing different practices will be in the future?

The tested knowledge (through research) thus obtained contributes to problem solving or decision making i.e. what should be done in education? It may be related to objectives, if instruction or selection of instructional material or classroom management or methods of teaching or any other area.

So, research in education is a more systematic activity that is directed towards discovery and the development of an organized body of knowledge in education.

Several other meanings used to understand educational research are:

- It is a scientific approach applied to study educational problems to acquire dependable and useful information about the educative process.
- It is a total procedure employed in collecting, organizing, summarizing and interpreting the data for the purpose of arriving at dependable answers to questions about education.

Travers (1978) has defined educational research as an activity directed towards the development of an organized body of scientific knowledge about the events with which educators are concerned.

The general steps in psychological/educational researches are:

1. Selecting a problem
2. Analysis stage (Review)
3. Selecting research strategy and developing instrument
4. Collecting and interpreting data
5. Reporting the result

Educational research is a systematic activity that is directed towards discovery and the development of an organized body of knowledge in education. ***Educational research can be classified into different types based on i) purpose ii) methodology and (iii)Nature of information concerned.***

According to the purpose educational researches are: **i) Fundamental or Basic ii) Applied iii) Action Research.**

**Based on the methodology the educational researches are (i) Experiment (what will be) (ii) Ex-post facto (after the fact / incidence) (iii) Descriptive (what is) (iv) Historic (what was).** ***The descriptive researches are further classified as (i) Case Studies (ii) Surveys (iii) Correlation Studies (iv) developmental studies (v) documentary analysis and (vi) trend analysis***

## 1.2.(i) Experimental Research

Experimental research methods help to answer questions about causes of behaviour. In this type of research, the investigator deliberately controls and manipulates the conditions which determine the events in which he is interested. In other words, the experimenter makes change in the value of one variable (independent) and observes the effect of that change on another variable (dependent). For example, to see the effect of different teaching methods upon achievement in mathematics, the investigator has to manipulate method (the independent variable) i.e by using a new teaching method, to ascertain its effect upon achievement (the dependent variable). As variables are the conditions/characteristics that the experimenter manipulates, controls and observes, we can say involves:

- Manipulation of independent variable e.g. (method of teaching, instructional material etc)
- Controlling or keeping all other variables constant (e.g. age, sex, socio-economic status of students etc.
- Observing the effect of manipulation of independent variable on dependent variable (e.g. effect of method of teaching/instructional material on achievement of students in Mathematics)

In addition to independent and dependent variable, there are some other variables such as intervening and extraneous variables which have to be taken into consideration during experimental research.

- **Intervening variables**: Variables, which cannot be controlled and measured directly e.g. motivation, anxiety, fatigue etc. These variables are difficult to define in operational and observable terms but cannot be ignored as they effect upon the result of research.
- **Extraneous variables**: Extraneous variables are uncontrolled variables that are not manipulated by the experimenter i.e. time of the day, outside noise etc. They also significantly influence the outcome of the study.

In this type of research the experimenter or researcher deliberately and systematically introduces changes into natural phenomena and then observes consequences of these changes. In the simplest, in an experimental situation, the independent variable is manipulated, all other variables are held constant and the effect of manipulation of independent variable on dependent variable is observed.

i) Effect of inquiry method of training on creativity and achievement in science.
ii) Effect of activity based teaching on attitude towards mathematics

## 1.2 (ii) Ex-Post Facto Research

In case of ex-post facto research, the researcher does not have control over independent variable because their manifestations have already been achieved. Ex:

i) Comparison of performance of high IQ and low IQ children on anxiety
ii) Comparison of perception of brain damaged and non-brain damaged children.

## 1.2 (iii) Descriptive Research

Descriptive researches are designed to obtain information on concerning the current status of phenomena.

**a. Case Study**: Intensive investigation of an individual or unit. e.g.

ii) Case study of community involvement in a school inn a village of Gujarat
iii) Case study of school broadcasting in Ahmedabad
iv) Case study of an innovative institution.

**b. Survey Method**: It is a widely used method of research to gather limited data from a large number of cases at a particular time with intention of describing the nature of existing condition, identifying standards against which existing conditions can be compared or determining the relationships that exist between specific events. Thus, it is generally used for obtaining descriptive and evaluative information in education. Therefore, it may vary in their levels of complexity from those, which provide simple frequency counts to those which present relational analysis. Generally three type of information are collected in a survey. They are:

- Data concentrating on existing status

- Comparison of status and standards (to know adequacy of status)
- Finding the means of improving status.

Some examples for survey method are:

i) Study of students level of aspiration
ii) Study of parents attitude towards school
iii) Inequality in physical facilities among various types of school

**c. Correlation:** To determine the extent of relationship between two variables: ex.

i) Achievement in mathematics and chemistry
ii) Intelligence and creativity

**d. Developmental:** To get information about what children are at various stages i.e. related to intellectual, emotional and physical etc. Ex:

i) IQ scores of children from 2-8 years
ii) Development of quantitative skill for grade level at primary stage.

**e. Follow-up**: Subsequent development of subjects after a specified treatment or condition. Ex

i) Physical, mental health of gifted children in 1921 and 1936
ii) Emotional problems of school going children after earthquake

**f. Documentary analysis**: Information obtained only by examining records and documents. Ex:

i) Uses of fraction in commercial work
ii) Text book analysis for gender bias

**g. Trend Studies:** Use of survey and documentary analysis to study rate and direction of changes and use these trends to predict future status Ex:

i) increased use of biological or medical research results in +2 curriculum
ii) increased college trained labour force etc.,

## 1.2 (iv) Historical Research

Historical research includes critical investigations of events, development and experience of past, the careful weighing of evidences of validity of sources of information on past and interpretation of weighed evidence to understand present and to forecast future. Ex:

i) Biography of educationists
ii) History of primary education
iii) History of educational planning and policy formulation
iv) History of staff development for secondary schools.

## 1.3 Selection of Problems for Research in Education

Finding a problem is the first step in developing a research project. But choice of suitable problem is one of the most difficult tasks. Even for great scientists it has become a difficult task. Many a time the researchers choose poor problems for study. Beginners are likely to select problems which are too broad in scope e.g. Problems of girl students of Gujarat state. Poor selection of problems may occur due to lack of understanding of nature of research and systematic problem solving activity. It may also happen due to too much enthusiasm and desire to solve problem quickly and immediately. Therefore, the question comes about: what are the methods of finding problem for research and what are the considerations to be kept in mind in selecting a problem?

Before a decision is made to select a good topic, some points are to be kept in mind such as:

* The researcher should select a problem that interests him/her but the problem selected should be soluble.

- Research problem should be developed out of ongoing research problem. The relevant knowledge could be obtained from reviews and studies (exceptions: some problems confronted in classroom, school or community lead themselves for immediate investigation and solution as in an action research).

- The researcher may not have time to read extensively each topic but he /she can consult books, journals and dissertations in the library to see how much has been done. Also he/she can talk/discuss with colleagues. Discussion with colleagues also helps in asking their support and collaboration.

- The researcher should be competent to carry out the study related to the problem he/she has selected i.e.

  - He/she should have enough knowledge in the field to understand significant aspects and interpret the finding.
  - He/she should be skillful in developing, administering and interpreting the necessary data gathering devices and procedure.
  - Should have knowledge of research design and statistical procedures.
  - Should have courage and determination to pursue the study in spite of the difficulties/hazards.

- Research problem should take into consideration the resources (data accessibility, financial provision) and time available with researcher to undertake the investigation.

Search strategies for finding a research topic.

**Search Strategies**

| Personal Strategies | Interpersonal Strategies | Printed Sources | Computer Strategies |
|---|---|---|---|
| Personal experience<br>Idea generation<br>Novels/television | Interviewing professor<br>Research assistantship<br>Other researches<br>Conferences/net<br>Working | Primary journals<br>Secondary journals<br>other library<br>resources Thesis/<br>Dissertation | Literature searches<br>internet, www |

[Collected *from Leong & Austin (1196)]*

Before selection of the problem, it is better to prepare the groundwork, which helps to save time later. Sometimes pilot studies also play an important role and indicate the feasibility of research.

## 1.4 Criteria of Problems and Problem Statements

Adequate statement of the research problem is considered as one of the most important parts of research but it is difficult to state a research problem satisfactorily. The final version of the title should tell the reader what the study is about. So the researcher would be only ready to devise/write a final title when he/she is clear about the focus of the study. Keeping this in mind, the researcher has to know what the problem is. The larger part of the solution lies in knowing what it is one trying to do.

Now question comes to the mind of a beginner- what is a good problem statement? Although research problem differ greatly, and although there is non 'right" way to state a problem, certain characteristics of problems statements can be learnt and used for better research in education.

There are three criteria of good problem and problem statement.

i) The problem should express a relation between two variables (exceptions taxonomic and historical research problems)
Ex. Do the teachers' reinforcement cause improvement in pupil's achievement in Mathematics?

For the above problem, the question is whether pupil's achievement in Mathematics is related to reinforcement of teachers?

ii) The problem should be stated clearly and unambiguously in question form.

iii) The problem or problem statement should imply possibilities of empirical testing.

## 1.5 Writing a Research Proposal

A research proposal is like a blue print. A good research project is likely to result only from well-designed proposal. A researcher has certainly needs to develop a research proposal for the following reasons:

- It helps the investigator to organize his/her ideas in a form by which it would be possible for him to find out the inadequacies or mistakes.
- It provides an inventory of work to be done while undertaking a study.
- It can be considered as initial draft proposal; and can be submitted to others for comment criticism or suggestion. It can be subjected to modification in the light of the analysis of the experts.

Many institutions like UGC, NCERT, ICSSR and other funding agencies provide their own format for submitting the research proposal. But most of them suggest the basic pattern of writing the research proposal, which is as follows

1. Statement of the problem/ Title of the problem
2. Need/significance of the problem
3. Definitions, assumptions, limitations and delimitations
4. Review of related literature
5. Hypothesis/research questions
6. Methods
   a) sample
   b) procedure
   c) data analysis/statistics to be issued
7. Time schedule
8. Budget.

## 1.6 Collection of Data and Data Processing

The different kinds of data gathered should be specifically mentioned. The source of each kind of data and the tools and techniques used for collecting data need to be specified. It is appropriate to describe the tools with necessary justification for use in the study. The tools will include questionnaire, inventories, scales, tests etc. Techniques will include participant observation, case study, interview etc. Describe the manner in which the different kinds of data have been processed. If the data have been processed through computer it should be mentioned. The processed data may present in tabular form or shown with the help of bar diagrams, pie, charts, graphs and so on. Tables, which are complex and lengthy, should be placed in appendix.

## 1.7 Report Writing

Once a study is complete and the conclusive/inferences have been drawn, the researcher would like to convey the results to someone else. This is done through research report, which provides a complete and comprehensive account to research

experience of the investigator(s). The important thing is to use a style that will convey the information clearly and appropriately to the reader.

Good report writing in itself is an art. Experience shows that some reports immediately attract our attention while there are others, which perhaps wait in the racks indefinitely. Report, being the ultimate product of a research is in the hands of fellow researchers, planners and implements, requires utmost care in its preparation. There should be an attempt to faithfully communicate the process of research, its question, assumption and limitation, the design, findings and implications and also the references and suggestions in a rigorous but lucid manner. As a general rule a report should be impressive, intelligent and implementable. It should not be incomprehensible, inadequate and inconsistent. The report should cover all points, it is supposed to cover. A research report begins with the title of the research, which should be succinct and meaningful. It should tell as briefly and directly as possible the precise nature of what the research is all about. The person who is responsible for conducting the research will place his or her name as the author. When there are multiple authors, decision regarding the order of authors should be made by mutual consent. As a general rule the person who shares the major responsibility in performing the research should be, listed as the first author. This is followed by institutional affiliation, which indicated the name of the institution supporting the research study. In case the author changes his institution in between, he/she may use his/her option of listing either institution. However, a better alternative maybe to indicate the original institution along with the author's name and put the new institution in the foot note with mailing address.

The detailed report is many a time preceded by an abstract/a brief summary/an executive summary, containing the key points of the report. The main report may be sub divided into the following heads. The format suggested below is only indicative and may undergo suitable modifications depending upon the nature of the study

**(a) Statement of the Problem**

Specify clearly but briefly the problem investigated. This may be followed by establishing the need and justification of the problem along with its possible educational implications.

**(b) Objectives of the Study**

Specific objectives for carrying out the study may be spelt out here

**(c) Overview of Literature**

This section will summarize the current status of research in the area including their major findings. Up to data related studies quoted comprehensively should clearly demonstrate the relevance/insufficiency of the findings, approaches leading to further investigation into the problem at hand.

**(d) Research Questions**

Specific questions answered through the research should be sharply focused.

**(e) Hypotheses (if any)**

State the hypothesis tested clearly. Identify all variables

**(f) Design of the Study**

Mention here the plan, structure and strategy of investigation for obtaining answers to the research questions. In the light of the questions raised or hypothesis tested full information be provided regarding (1) Universe of the study, (2) Sample and sampling procedure (3) Units of observations. If the study has taken any control groups they should be specifically mentioned. An explanation of determination of size and type of sample will also be necessary. Studies not requiring a sample selection should specify their strategy appropriately and describe the rationale.

**(g) Results and Discussion**

The conclusions are to be presented concisely and related directly to the hypothesis. They announce whether the findings accept or reject the hypothesis. In fact conclusions are answers to the questions raised and suggest modifications in the existing theory. The researchers may also list unanswered questions in the process of the study, which probably require further research. In short the discussions and presentation of results should give the reader an impression of completeness and of positive gain.

**(h) Summary and Conclusion**

This selection returns to the conceptual level. It helps the readers to determine what conclusions they can actually draw based on the research presented. Hence, the summary includes a brief re-statement of the problem, a description of the procedures used and discussion of findings and conclusions of the study.

# 2. Qualitative Research*

**Objectives:** After reading this module, you will able to understand;

- The meaning and nature of qualitative research
- Ethnography as a method of qualitative research

**2.1 Introduction:** There are two important assumptions, which have affected the methods of research in social sciences, and these two assumptions provide two different ways of looking at social reality. One of them is ontological which concern the very nature or essence of the social phenomena being investigated. The second type of assumptions are epistemological that are concerned with the many houses of knowledge i.e. its nature and forms, how it can be acquired and how communicated to other human beings. Another set of assumptions are concerned to human nature i.e. relationship between human being and their environment. The above three sets of assumptions have direct implications for the methods of research in social science. Group of researchers adopting objectives (positivist) approach to social world and treating it like world of natural phenomena (hard, real, external to the individual) choose quantitative methods i.e. eliciting responses to predetermined questions, recording, measurement, describing phenomena as in surveys and but performing experiments etc., to find out regularities between selected factors in that world and the researchers adapting subjectivist (anti-positivist) approach view the social world as man created kind (so after and personal) select qualitative methods i.e. participant observations, role playing, non directive interviewing, episodes and accounts etc. in which they give importance to subjective experience of individuals in creation of social world. In case of qualitative research, the researcher tries to understand what is unique and particular to the individual rather than what is general and universal.

The approach, which is used to discover general laws are in case of quantitative research, is called as nomothetic but the approach used in qualitative research emphasizing relativistic value of social world and understanding individual behavior is called ideographic.

The qualitative research came into existence as result of some of the criticism of positivism and scientific method in which anti-positivists attacked on science's mechanistic, and reductionist view of nature which by definition excludes notions of choice, freedom, individuality and moral responsibility. They rejected the belief that human behavior is governed by general laws and characterized by underlying regularities and argue that individual's behavior can only be understood by the research by sharing his frame of reference- understanding individual's interpretations of the world around him has to come from inside not the outside. Table 1 gives a picture about two different kinds of research.

---

* This material is taken from the work of Dr. Sabita Patnaik's *"Handbook of research methodology for the CTE and IASE faculty"* , prepared at Regional Institute of Education, Bhopal

Table 1 : Two schools of social science

| Approach | Concepts | Methods |
|---|---|---|
| Positivism (normative) | Social structure social facts | Quantitative hypothesis (nomothetic) |
| Anti-positivism (Interpretive) | Social, construction, meaning | Qualitative hypothesis Generation (Idiographic) |

In order to understand the characteristic of qualitative research, some of the following versions would be quite useful.

Table 2: One version of Qualitative Research (Adapted from Bryman, 1988)

1. 'Seeing through the eyes of 'Taking the subject's perspective
2. Describing the mundane detail of everyday settings
3. Understanding actions and meanings in their social context.
4. Emphasizing time and process
5. Favoring open and relatively unstructured research designs
6. Avoiding concepts and theories at an early stage

Table 3: A Second Version of Qualitative Research (Adapted from Mammarsely, 1990)

1. The use of everyday contexts rather than experimental condition.
2. A range of sources of data collection (the main ones are observations and informal conversations')
3. A preference for 'unstructured' data collection (no prior hypothesis, no prior definitions)
4. A concern with 'micro' features of social life ('a single setting or group)
5. A concern with the meaning and function of social action
6. The assumption that quantification plays a subordinate role

Table 4: A Third version of Qualitative Research (Adapted from Hammersley, 1992)
A preference for qualitative data-use words rather than numbers. However, in principal, there is no reason to prefer any form of data.

We are not faced with a start choice of words and numbers, or even between precise and imprecise data, but rather with a range from more to less precise data. Furthermore, our decisions about what level of precision is appropriate in relation to any particular claim should depend on the nature of what we are trying to describe, on the likely accuracy of our descriptions, on our purposes, and on the resources available to us, not an ideological commitment to one methodological paradigm or another. (Hammersley, 1992).

1. A preference for naturally – occurring data rather than experiment, unstructured versions structured interviews. However, this falls because, as Hammersley says, even observation can be unrepresentative because it differs in important ways from other cases in that category'. Also no research is untouched by human-hands (cf.Hammersely and Atkinson: 1983 and the critique of naturalism.

2. A preference for meaning rather than behavior-attempting 'to document the worked from the point of view of the people studied'. However, as Hammersely rightly point out, this is a cop-out since respondents can do this for themselves; ultimately, the social scientist must analyze rather than simply let the participants speak for themselves (see Gilbert ande Mulkay: 1983)

3. A rejection of natural science as a model. However, there are many different kinds of natural science (from Botany to Theoretical Physics). Also qualitative research has a very problematic status if it totally fails to address the validity of its findings or reduced validity to participants' agreement with a set of findings.

4. A preference for inductive, hypothesis-generating research rather than hypothesis testing (Glasser and Strauss: 1967). However, hypothesis must at some point be tested, otherwise we are limited to mere speculation. As Hammersley writes: *which of these approaches is most appropriate should on our purpose, and the stage that our research has reached not on paradigmatic commitments.*

Table 5: A Fourth version of Qualitative Research (Adapted from Hammersley and Atkinson 1983.

1. A preference for 'natural' settings as the primary source of data.
2. A fidelity to the phenomena under study-this requires a cultural description of the meanings of phenomena under study –this requires a cultural description of the meanings of phenomena to participants.
3. The use of an inductivist methodology, which avoids the premature testing of hypotheses.

[All the versions are collected from Silverman (1995)]

In many of the social sciences, the quantitative researches are often given more respect as people have tendency to regard science as related to numbers and implying precision. Another reason is qualitative research take much longer, requires greater clarity of goals during design stages and cannot be analyzed by running a computer program. But qualitative research has left its mark conceptually and theoretically on social sciences by contributing to social understanding significantly.

In conclusion it can be said that in order to understand the way in which individual creates, modifies and interprets the world, the approach should be a qualitative as well as quantitative.

## 2.2 Ethnography

In psychological or educational researches, researcher mostly uses a nomethetic methodology in which questionnaire survey is used to collect a large quantity of data. A questionnaire on 'attitude towards environment' channel responses along a predetermined route with very little opportunity for the individual to say anything about their concern towards environmental problems in terms their behavior, the way they think about issues related to environment etc. This type of problem encourages the use of ethnography.

Ethnography is not a single method. It includes direct observation, interviewing, discourse analysis, diary techniques and even questionnaire survey. The purpose of ethnography is cultural description. Spradley (1979) defined ethnography as 'culture studying culture'. It consists of a body of knowledge that includes research techniques, ethnographic theory, and hundreds of cultural descriptions. It seeks to build a systematic understanding of all human cultures from perspectives of those who have learned. So ethnographic research emphasizes on the actors understanding and theorizing about their action i.e. the views is not the outsider looking in but the insider looking around.

For better understanding of ethnography Werner and Schoepfle (1987) proposed that 'Ethnography is description' and that 'description must closely resemble the original cultural reality. The resemblance must be good enough that natives are able to recognize in it familiar features of their own culture'. Goetz and Le Compte (1984) have defined ethnography as the 'analytic descriptions or reconstitution of intact cultural scenes and groups. Ethnograhics recreate for the reader the shared beliefs, practices, arte-facts, folk knowledge and behaviour of some groups of people'.

According to Goetz and Le compte (1984) "ethnography is a process in which phenomenological data elicited i.e aiming to represent world view of those individuals or groups under investigation – representation of the world is structured by the participants, not by the researcher."

Ethnographic techniques are employed in natural settings. The researcher finds out how individuals and groups behave in their own real world setting unmanipulated by the researcher.

- Ethnographic research attempts to present the totality of the phenomenon under investigation. The context is as important as the action. The temporal and the environmental factors and the social/cultural and economic context are not noise but fundamental contributory explanatory variables.

Ethnographic field strategies can be divided into four steps (Berg, 1989)

i) Accusing a field setting – Getting in
ii) Becoming invisible
iii) Watching, listening and learning
iv) Disengaging and getting out.

**Some details about ethnographic strategies**

| Accusing a field setting – getting in | Becoming invisible | Watching, listening & Learning | Disengaging and getting out |
|---|---|---|---|
| • Attitude of Ethnographer<br>• Subjective motivational factor<br>• External motivating factors<br>• Gaining entry<br>• Developing research<br>• Bargain<br>Guides and informatics | • Disattending – Erosion of visibility by time<br>• Disattending – Erosion of Visibility by Display of no Symbolic detachment<br>*<br>Disattending- Erosion of visibility by display of symbolic attatchment. | •<br>What to watch & listen for<br><br>-Taking in physical setting<br>- Developing<br>- Relationship with inhabitants .<br>- Tracking, Observing, eaves dropping<br>- Locating subgroups and stars<br>- Field notes<br>- Erosion of memories<br>- Clues and strategies for recalling data<br>- Analyzing the ethnographic data | * Physical removal |

For the research topic 'Analytic study of delinquent children of slum areas' ethnographic strategies can be applied by the researcher. `

# 3. Quantitative Research

## Objectives:

After reading this module, you will be able to:

1. Define quantitative research
2. Identify features commonly associated with quantitative research rather than with qualitative research
3. Give an outline of the most prominent forms of quantitative research
4. Write down key terms associated with quantitative research
5. List out criteria for selection of the tools and techniques required for any proposed quantitative research.

**3.1 Introduction:** Educational research is a systematic activity that is directed towards discovery and the development of an organized body of knowledge in education. *Educational research can be classified into different types based on (i) purpose (ii) methodology and (iii)Nature of information concerned.*

According to the purpose it serves, educational researches are: (i) Fundamental or Basic (ii) Applied and (iii) Action Research.

Based on the methodology the educational researches are (i) experiment (what will be) (ii) Ex-post facto (after the fact / incidence) (iii) Descriptive (what is) and (iv) Historic (what was). *The descriptive researches are further classified as (i) Case Studies (ii) Surveys (iii) Correlation Studies (iv) developmental studies (v) documentary analysis and (vi) trend analysis.*

Educational researches are classified as quantitative or qualitative research if the concerned information respectively is in numerical form or it is not.

Quantitative research is a formal, objective, systematic process in which numerical data are utilised to obtain information about the world.

Quantitative research is used to measure how many people feel, think or act in a particular way. Qualitative research is used to help us understand how people feel and why they feel as they do. It is concerned with collecting in-depth information asking questions such as why do you say that? Qualitative research is used to explore and understand people's beliefs, experiences, attitudes, behaviour and interactions. It generates non- numerical data, e.g. a student's liking of his / her teacher rather than quantifying his liking. In Psychological studies, qualitative techniques have been commonly used in research, documenting the attitudes and behaviours of the people.

Quantitative research involves analysis of numerical data. Qualitative research involves analysis of data such as words or text (ex:. from interviews), photographs or pictures (ex video), sound recordings, or objects (ex. an artifact).

There are five major and important steps involved in any research process:

1. **Defining the Problem**
2. **Research Design**
3. **Data Collection**
4. **Analysis**
5. **Report Writing & presentation.**

The brief discussion on each of these steps is:

1. Problem audit and problem definition - What is the problem? What are the various aspects of the problem? What information is needed?
2. Conceptualization and operationalization - How exactly do we define the concepts involved? How do we translate these concepts into observable and measurable behaviours?
3. Hypothesis specification - What claim(s) do we want to test?
4. Research design specification - What type of methodology to use? - examples: questionnaire, survey
5. Question specification - What questions to ask? In what order?
6. Scale specification - How will preferences be rated?
7. Sampling design specification - What is the total population? What sample size is necessary for this population (if required)? What sampling method to use?- examples: (i)**Probability Sampling:-** (cluster sampling, stratified sampling, simple random sampling, multistage sampling, systematic sampling) & (ii)**Nonprobability sampling:-** (Convenience Sampling,Judgement Sampling, Purposive Sampling, Quota Sampling, Snowball Sampling, etc. )
8. Data collection - Use mail, telephone, internet, etc.
9. Codification and re-specification - Make adjustments to the raw data so it is compatible with statistical techniques and with the objectives of the research - examples: assigning numbers, consistency checks, substitutions, deletions, weighting, dummy variables, scale transformations, scale standardization
10. Statistical analysis - Perform various descriptive and inferential techniques (see below) on the raw data. Make inferences from the sample to the whole population. Test the results for statistical significance.
11. Interpret and integrate findings - What do the results mean? What conclusions can be drawn? How do these findings relate to similar research?
12. Write the research report - Report usually has headings such as: 1) executive summary; 2) objectives; 3) methodology; 4) main findings; 5) detailed charts and diagrams. Present the report to the client in a 10 minute presentation. Be prepared for questions.

In quantitative research the aim normally is to determine the relationship between one thing (an independent variable) and another (a dependent or outcome variable) in a population. Quantitative research designs are either descriptive (subjects usually measured once) or experimental (subjects measured before and after a treatment). A descriptive study establishes only associations between variables. An experiment establishes causality.

For an accurate estimate of the relationship between variables, a descriptive study usually needs a sample of hundreds or even thousands of subjects; an experiment, especially a crossover, may need only tens of subjects. The estimate of the relationship is less likely to be biased if you have a high participation rate in a sample selected randomly from a population. In experiments, bias is also less likely if subjects are randomly assigned to treatments, and if subjects and researchers are blind to the identity of the treatments.

In all studies, subject characteristics can affect the relationship you are investigating. Limit their effect either by using a less heterogeneous sample of subjects or preferably by measuring the characteristics and including them in the analysis. In an experiment, try to measure variables that might explain the mechanism of the treatment. In an unblinded experiment, such variables can help define the magnitude of any placebo effect.

Quantitative research is all about quantifying relationships between variables. Variables are things like weight, performance, time, and treatment. You measure variables on a sample of subjects, which can be tissues, cells, animals, or humans. You express the relationship between variable using effect statistics, such as correlations, relative frequencies, or differences between means etc.

## 3.2 Major Types of Quantitative Research

Studies aimed at quantifying relationships are of two types: **descriptive** and **experimental**. Descriptive studies are also called **observational**, because you observe the subjects without otherwise intervening. In a descriptive study, no attempt is made to change behavior or conditions--we measure things as they are. In an experimental study we take measurements, try some sort of intervention, and then take measurements again to see what happened. The simplest descriptive study is a **case**, which reports data on only one subject; examples are a study of an outstanding athlete or of a dysfunctional institution. Descriptive studies of a few cases are called **case series**. In **cross-sectional** studies variables of interest in a sample of subjects are assayed once and the relationships between them are determined. In **prospective** or cohort studies, some variables are assayed at the start of a study (e.g., dietary habits), then after a period of time the outcomes are determined (e.g., incidence of heart disease). Another label for this kind of study is **longitudinal**, although this term also applies to experiments. **Case-control** studies compare **cases** (subjects with a particular attribute, such as attitude or ability) with **controls** (subjects without the attribute); comparison is made of the **exposure** to something suspected of causing the cases, for example volume of high intensity training,

or number of alcoholic drinks consumed per day. Case-control studies are also called **retrospective**, because they focus on conditions in the past that might have caused subjects to become cases rather than controls.

A common case-control design in the exercise science literature is a comparison of the behavioral, psychological or anthropometric characteristics of elite and sub-elite athletes: you are interested in what the elite athletes have been exposed to that makes them better than the sub-elites. Another type of study compares athletes with sedentary people on some outcome such as an injury, disease, or disease risk factor. Here you know the difference in exposure (training vs no training), so these studies are really cohort or prospective, even though the exposure data are gathered retrospectively at only one time point. The technical name for these studies is **historical cohort**.

Experimental studies are also known as **longitudinal** or **repeated-measures** studies, for obvious reasons. They are also referred to as **interventions**, because you do more than just observe the subjects.

In the simplest experiment, a **time series**, one or more measurements are taken on all subjects before and after a treatment. A special case of the time series is the so-called **single-subject design**, in which measurements are taken repeatedly (e.g., 10 times) before and after an intervention on one or a few subjects.

Time series suffer from a major problem: any change you see could be due to something other than the treatment. For example, subjects might do better on the second test because of their experience of the first test, or they might change their diet between tests because of a change in weather, and diet could affect their performance of the test. The **crossover** design is one solution to this problem. Normally the subjects are given two treatments, one being the real treatment, the other a control or reference treatment. Half the subjects receive the real treatment first, the other half the control first. After a period of time sufficient to allow any treatment effect to wash out, the treatments are crossed over. Any effect of retesting or of anything that happened between the tests can then be subtracted out by an appropriate analysis. **Multiple crossover** designs involving several treatments are also possible.

If the treatment effect is unlikely to wash out between measurements, a **control group** has to be used. In these designs, all subjects are measured, but only some of them--the **experimental group**--then receive the treatment. All subjects are then measured again, and the change in the experimental group is compared with the change in the control group.

If the subjects are assigned randomly to experimental and control groups or treatments, the design is known as a **randomized controlled trial**. Random assignment minimizes the chance that either group is not typical of the population. If the subjects are **blind** (or **masked**) to the identity of the treatment, the design is a **single-blind** controlled trial. The control or reference treatment in such a study is called a **placebo**: the name physicians use for inactive pills or treatments that are given to patients in the guise of

effective treatments. Blinding of subjects eliminates the **placebo effect**, whereby people react differently to a treatment if they think it is in some way special. In a **double-blind** study, the experimenter also does not know which treatment the subjects receive until all measurements are taken. Blinding of the experimenter is important to stop him or her treating subjects in one group differently from those in another. In the best studies even the data are analyzed blind, to prevent conscious or unconscious fudging or prejudiced interpretation.

Ethical considerations or lack of cooperation (compliance) by the subjects sometimes prevent experiments from being performed. For example, a randomized controlled trial of the effects of physical activity on heart disease may not have been performed yet, because it is unethical and unrealistic to randomize people to 10 years of exercise or sloth. But there have been many short-term studies of the effects of physical activity on disease risk factors (e.g., blood pressure).

Both designs, quantitative and qualitative are said to be systematic. In fact having a system or following a process is a defining principle of research.

Broadly speaking, quantitative research is thought to be objective whereas qualitative research often involves a subjective element. It is thought that in gaining, analysing and interpreting quantitative data, the researcher can remain detached and objective. Often this is not possible with qualitative research where the researcher may actually be involved in the situation of the research.

Consider a study being undertaken into waiting times in the Accident & Emergency (A&E) Department of a hospital.

A quantitative study, measuring how long people wait, can be purely objective. However if the researcher wanted to discover how patients felt about their waiting time, they would have to come into contact with the patients and make judgments about the way they answered their questions. If the researcher asked the patient "how are you feeling having waited an hour to be seen by the doctor?" they would almost certainly register the patients' non-verbal behaviour as well as document the response; in this way the researcher is adding a subjective element to the study.

Quantitative research is inclined to be deductive. In other words it tests theory. This is in contrast to most qualitative research, which tends to be inductive. In other words it generates theory. Using the A&E waiting time example again, the quantitative approach might test the hypothesis that "Patients attending this A&E department do not wait for more than one hour to be seen by a doctor".

A qualitative approach which explores the feelings of patients who wait an excessive time to be seen by the doctor might generate the theory that "patients who experience an excessive wait to be seen by the doctor, experience an enlargement of the symptoms that brought them to the department".

However, qualitative studies tend to produce results that are less easy to generalise. This has to do with the problem of the sample used at the time. We all know, for example, that our feelings about waiting can change dependent on our particular set of circumstances. Even if the researcher encountered the same group of clients on another day, they may find different results. Generally, it is difficult to generalise with qualitative results.

Lastly here, the most obvious difference between quantitative research and qualitative research is that quantitative research uses data that are structured in the form of numbers or that can be immediately transported into numbers. If the data cannot be structured in the form of numbers, they are considered qualitative. (Note that qualitative data can sometimes be handled in such a way as to produce quantitative data. e.g. the researcher exploring feelings of patients can analyse the responses in clusters that are negative or positive so as to produce a figure/percentage of negative patient and positive patient feelings).

**Predispositions of Quantitative and Qualitative Modes of Inquiry**

**Quantitative Mode**

**Assumptions**

- Social facts have an objective reality
- Primacy of method
- Variables can be identified and relationships measured
- Ethic (outside's point of view)

**Purpose**

- Generalizability
- Prediction
- Causal explanations

**Approach**

- Begins with hypotheses and theories
- Manipulation and control
- Uses formal instruments
- Experimentation
- Deductive

**Qualitative mode**

**Assumptions**

- Reality is socially constructed
- Primacy of subject matter
- Variables are complex, interwoven, and difficult to measure
- Emic (insider's point of view)

**Purpose**

- Contextualization
- Interpretation
- Understanding actors' perspectives

**Approach**

- Ends with hypotheses and grounded theory
- Emergence and portrayal
- Researcher as instrument
- Naturalistic
- Inductive

| | |
|---|---|
| • Component analysis<br>• Seeks consensus, the norm<br>• Reduces data to numerical indices<br>• Abstract language in write-up | • Searches for patterns<br>• Seeks pluralism, complexity<br>• Makes minor use of numerical indices<br>• Descriptive write-up |
| **Researcher Role** | **Researcher Role** |
| • Detachment and impartiality<br>• Objective portrayal | • Personal involvement and partiality<br>• Empathic understanding |

Although some social science researchers (Lincoln & Guba, 1985; Schwandt, 1989) perceive qualitative and quantitative approaches as incompatible, others (Patton, 1990; Reichardt & Cook, 1979) believe that the skilled researcher can successfully combine approaches. The argument usually becomes muddled because one party argues from the underlying philosophical nature of each paradigm, and the other focuses on the apparent compatibility of the research methods, enjoying the rewards of both numbers and words. Because the positivist and the interpretivist paradigms rest on different assumptions about the nature of the world, they require different instruments and procedures to find the type of data desired. This does not mean, however, that the positivist never uses interviews nor that the interpretivist never uses a survey. They may, but such methods are supplementary, not dominant.... Different approaches allow us to know and understand different things about the world.... Nonetheless, people tend to adhere to the methodology that is most consonant with their socialized worldview.

The two techniques are not mutually exclusive and a **qualitative research** consultant may well suggest that both techniques are used in conjunction with one-another.

## 3.3 Quality of Designs

The various designs, mentioned above differ in the quality of evidence they provide for a cause-and-effect relationship between variables. Cases and case series are the weakest. A well-designed cross-sectional or case-control study can provide good evidence for the **absence** of a relationship. But if such a study does reveal a relationship, it generally represents only suggestive evidence of a causal connection. A cross-sectional or case-control study is therefore a good starting point to decide whether it is worth proceeding to better designs. Prospective studies are more difficult and time-consuming to perform, but they produce more convincing conclusions about cause and effect. Experimental studies provide the best evidence about how something affects something else, and double-blind randomized controlled trials are the best experiments.

**Confounding** is a potential problem in descriptive studies that try to establish cause and effect. Confounding occurs when part or all of a significant association between two variables arises through both being causally associated with a third variable. For example, in a population study you could easily show a negative association between habitual activity and most forms of degenerative disease. But older people are less active, and older people are more diseased, so you're bound to find an association between activity and disease without one necessarily causing the other. To get over this problem you have to **control for potential confounding factors**. For example, you make sure all your subjects are the same age, or you include age in the analysis to try to remove its effect on the relationship between the other two variables.

There is no universal standard for categorising research designs and different authors may change names of designs in their discussions of them. Thus what is shown here is intended more to be informative than exhaustive. This lack of universalism also causes problems when critiquing research, as many published studies do not identify the design used. Selecting an appropriate design for a study involves following a logical thought process. A calculating mind is required to explore all possible consequences of using a particular design in a study.

## 3.4 Descriptive Design

Descriptive designs are designed to gain more information about a particular characteristic within a particular field of study. A descriptive study may be used to, develop theory, identify problems with current practice, justify current practice, make judgments or identify what others in similar situations may be doing. There is no manipulation of variables and no attempt to establish causality.

Correlational studies as already mentioned are not universally accepted as a form of quantitative research. However they do crop up in the literature so we will briefly discuss them here. As already noted they are also known as ex post facto studies. This literally means "from after the fact".

The term is used to identify that the research in question has been conducted after the variations in the independent variable has occurred naturally. The basic purpose of this form of study is to determine the relationship between variables. However the significant difference from experimental and quasi-experimental design is that causality cannot be established due to lack of manipulation of independent variables.

***"Correlation does not prove Causation"***

Examples include many studies of lung cancer. The researcher begins with a sample of those who have already developed the disease and a sample of those who have not. The researcher then looks for differences between the two groups in antecedents, behaviours or conditions such as smoking habits.

| Research with Subjects (Quantitative) | Research with Informants (Qualitative) |
|---|---|
| 1. What do I know about a problem that will allow me to formulate and test a hypothesis? | 1. What do my informants know about their culture that I can discover? |
| 2. What concepts can I use to test this hypothesis? | 2. What concepts do my informants use to classify their experiences? |
| 3. How can I operationally define these concepts? | 3. How do my informants define these concepts? |
| 4. What scientific theory can explain the data? | 4. What folk theory do my informants use to explain their experience? |
| 5. How can I interpret the results and report them in the language of my colleagues? | 5. How can I translate the cultural knowledge of my informants into a cultural description my colleagues will understand? |

References:


1. Daniel Muijs : *Doing Quantitative Research in Education with SPSS*
2. James Neil : *Qualitative versus Quantitative Research*
3. John W Best : *Research in education*
4. Nigel Fielding and Margrit Schreir : *Introduction: On the Comparability Between Qualitative and Quantitative Research Methods.*

# 4. Sampling

**Objectives:**
After going through this module, you will be able to understand:

- The meaning of survey
- The objectives and scope of surveys
- The concepts of population and sample
- Sampling issues in educational research
- Different types of sampling methods

**4.1 Introduction:** Since the beginning of the twentieth century, the different activities of human beings have undergone substantial changes due to rapid developments in the fields of science and technology. To satisfy the human needs, importance is given to achieve specialization in mass production and utilization goods and services of a given type with a view to getting the maximum possible benefit per unit cost. For this, a careful planning is required. In any planning there is the need for various type of quantified information to be collected and analyzed in an objective manner, and presented suitably so as to serve as a sound basis for taking policy decisions in different fields of human activity. To execute its various responsibilities, a variety of information is required about different aspects of education. Such information can be obtained through carefully planned and organized surveys.

The data one needs and resources that would be required for proper planning and execution of projects and for assessing their effectiveness may be classified into two types ***(i) survey data*** comprising of data already in existence, which are collected and recorded by observation or enquiry and ***(ii) experimental data*** which can only be obtained through well designed and controlled statistical experiments. In our discussion data refers to survey data.

**4.2 Survey:** ***An examination of an aggregate of units is known as Survey.***
The working out of the plan for organization of a survey consists of making certain decisions with respect to:

- What the survey is about and the type of data that are needed?
- Why the survey is being made?
- Where the needed data can be found?
- Where or in what areas, the survey will be carried on?
- When the survey is to be started and when to be concluded?
- How much material or how many cases will be surveyed?
- What basis of selection of data will be used?
- What techniques of gathering data will be adopted?

Thus, the considerations which enter into making the decisions regarding what, where, when, how much, by what means, constitute a plan of survey or survey design. The important steps involved in an organization of a survey are:

1. Specification of the objectives and scope of the survey
2. Identification of the population to be covered in the survey
3. Determination of the data to be collected (the frame)
4. Selection of statistical units
5. If the survey is sample survey then
    a. choice of sampling unit
    b. identification of the sample frames
    c. determination of the sample size
    d. selection of the sample
6. Specification of degree of precision or accuracy
7. Identification of the sources for of information
8. Determination of types of enquiry
9. Specification of time reference and reference period
10. Selection and preparation of data collection tools
11. Training of the concerned personnel (Interviewers and Supervisors) regarding data collection
12. Organization of pilot study
13. Modification of data collection, tools if necessary, based on the pilot study.
14. Collection of data
15. Scrutiny of data
16. Dealing with non-response
17. Tabulation and analysis of data
18. Interpretation of the results
19. Report writing

### 4.2.1 Objectives and Scope of the Surveys

The purpose for which the survey is to be conducted must be clearly and precisely laid down. This will indicate the type of information, which is needed, and the uses to which information obtained will be put. The various uses to which the results of a survey are usually put are

a) **Fact finding**: The facts might relate to knowing the number of institutions, enrollment, teaching personnel, inspectorate, educational cost etc.
b) **Comparative survey**: It may be desired to make a comparative study of some aspects of education between different regions of a country or between different countries.
c) **Evaluative surveys**: It is intended to evaluate existing status of any aspect of education in the background of agreed criteria, it will be necessary to have evaluative surveys.

After the specification of the purpose of the survey, the scope of the survey must also be spelt out clearly and unambiguously. The scope of the survey usually fixes the limits of the survey. The nature of the information to be collected, the geographical area to be surveyed, types of survey and the time limit for the completion of the survey are to

be clearly indicated. The objectives differ with the nature of survey. Sometimes hypotheses may be formulated and tested.

### 4.2.2 Population

The aggregate of all units of a given type under consideration is called population. A population may consist of persons, objects, attributes, qualities, behaviours of people, answers to various items of a test, cities, families, opinions of the electorate and the like. A population is a well-defined group of any of these. Defining a population means fixing the limits in terms of one or more of its various aspects.

These aspects may be regarding geographical area, age, sex, socio-economic status, physical attributes, psycho-social behaviours etc., In a survey, it is necessary to specify what group of units is to be covered by the survey and to assemble pertinent information and materials about the population and its subgroups that will be useful in drawing and getting in touch with the sample. These steps are particularly important since the general character of the population and many of its particular characteristics will limit the opportunities available for other aspects of the survey design. The costs and efficiencies off different procedures vary from population to population. The actual measurement and observations and the materials available for use in estimation and analysis are also affected by the specification of the population. Hence, not only it is necessary to specify precisely the population studied but it is also necessary to assemble and review information about the population, as carefully as other preparatory material on which the survey must rest.

### 4.2.3 Data to be collected

Once the population is identified, it is necessary to specify the type of information required –whether the required data to be collected confidentially or publicly; officially or personally; in adhoc way or regular way. The amount of data required from each unit is also clearly to be specified by identification of frames.

### 4.2.4 Statistical Units

Before a survey is conducted the units that are counted must be strictly defined. A Statistical unit is a unit of measurement applied to the data in any particular problem or investigation. For example, we investigate 'accidents', 'income', 'employment', 'wages' etc. these are statistical units. Suppose an enquiry is to be conducted about the wages of workers in some industry. Wage is a common terms which may mean money wage or real wage or piece wage or time wage. Wage of skilled worker, wage of unskilled worker, etc. one who conducts the enquiry must state clearly the sense in which he proposes to use the term 'wage'. Two kinds of statistical units are; 1***) units of collection*** and 2) ***units of analysis and interpretations***. Units of collection are those in terms of which measurements are made or data collected. These units are again of two types i.e **1) simple and 2) composite**. A simple unit expresses a single condition without any qualification. For example, a worker, a building etc. A composite unit is one which is formed by adding a qualifying word to a simple unit. For example, an industrial worker. Units of analysis and interpretation are those which make comparison possible and easy.

These units include; 1) rates 2) ratios and 3) coefficients. Specification of statistical units determines the selection of other aspects of the survey.

### 4.2.5 Census survey or sample survey

One way of obtaining the required information at regional and country level is to collect the data for each unit (person, school, school building, etc) belongings to the population or universe, which is the aggregate of all units of a given type under consideration. This procedure of obtaining information from all the units of population is called complete enumeration survey or census survey.

The effort, money and time required for carrying out census survey will generally be extremely large. However, if the information is required for each and every unit in the domain of the study, a census survey is necessary. For various reasons, census survey may be incomplete, inadequate or in accurate. Census surveys are unattainable in practice and complete accuracy is an illusory concept. Therefore, instead of obtaining data from all the units of the population, a proper method of obtaining data from some few selected units of the population may be used to get the required results. This type of enumeration of a subgroup of the population units is termed as **sampling.** Sometimes sample surveys are more efficient for the following reasons:

- Reduced cost: If data are secured from only a small fraction of the aggregate, expenditure may be expected to be smaller than if a complete census is attempted.
- Greater Speed: For the above stated reason, the data can be collected and summarized more quickly with a sample. This is very useful when the information is urgently needed.
- Greater Scope: In certain types of enquiry, highly trained personnel or specialized equipment, limited in availability, must be used in obtaining the data. In such cases, census, survey may be impracticable. Thus surveys, which rely on sampling, have more scope and flexibility as to the types of information that can be obtained. However, if the information is wanted for many subdivisions or segments of the population, the complete enumeration offers best solution.
- Greater Accuracy: Because, personnel of higher quality can be employed and can be given intensive training, a sample survey may actually produce more accurate results than the kind of census survey that is feasible to take.
- Practicability: Many populations about which inferences are to be made are quite large. For example, consider the population of class I students in Karnataka whose number is approximately in lakhs. The education department is interested in finding out as to what proportion of children enters the school at the specific age of enrollment. But the big size of the population makes it physically impossible to conduct a census. In such a case, selecting a representative sample may be the only way to get the information required.

- Accessibility: There are some populations that are so difficult to get access to all units. For example – people in prison, crashed airplanes in the deep sea, satellites in space, etc. The inaccessibility may be economic or time related. Therefore, in case of inaccessible populations, only sample study is possible.

Based on the resources, the choice of census survey or sample survey is to be determined. If the survey is a sample survey then a) choice of sampling unit b) identification of sampling frames, c) determination of the sample size and d) selection of the sample should be specified.

**(i). Sampling Unit:** A decision has to be taken concerning a sampling unit before selecting a sample. Sampling units are usually of the following:

- Geographical units such as state, district,
- Social units such as family, club, school, etc.
- Individuals
- Construction units such as house, flat, etc

The investigator (researcher) must decide as to which one of such units he has to select for his purpose.

**(ii). Sampling Frame**: All rigorous sampling demands a sub division of the material to be sampled into units termed sampling units, which form the basis of the actual sampling procedure, for using methods in the collection of data, it is essential to have a frame at all the sampling units belonging to the population to be studied with their proper identification particulars and such a frame is termed the sampling frame. Similarly, in area, sampling maps will serve as sampling frames. Sometimes these frames are in existence and can be readily obtained. Sometimes these have to be prepared at an extra cost before sampling. Sampling frames are very important and influence every aspect of a survey design, particularly sample design. Construction of a frame is often one of the major practical problems faced by the investigator. The readily available frames in general are inadequate, incomplete or partly illegible or many contain an unknown amount of duplication. The problem of inadequacy of a frame arises when it does not cover the whole of the population to be surveyed. A frame may be inadequate for one purpose and quite adequate for another. A frame is incomplete when some of the population units are supposed to be on it or in fact not on it. Since these elements will have no chance of being selected, the sample will be unrepresentative of the population to that extent. Therefore, construction of sampling frame is very essential in sample surveys.

### 4.2.6 Errors in Surveys

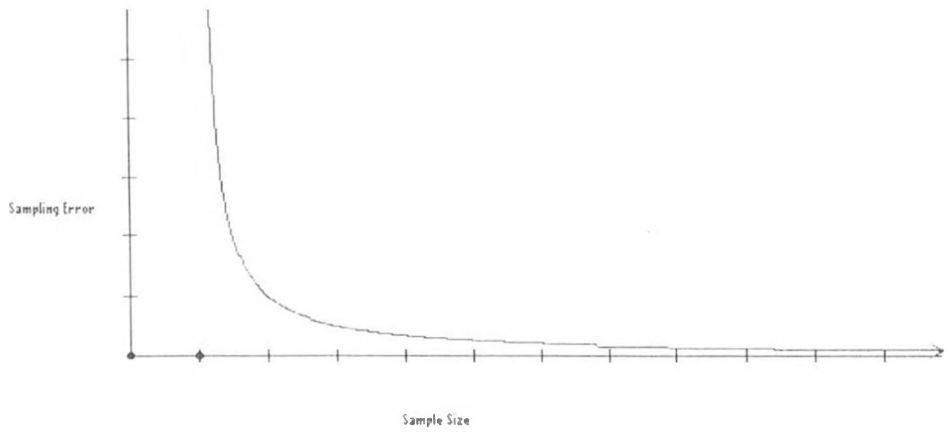
The deviations of the survey results from the true values are termed as errors. The different factors for these errors in surveys are:

a. Data specification being inadequate and inconsistent with respect to the objectives of the study.

b. Omission or duplication of units due to imprecise definition of the boundaries of area units, incomplete or wrong identification particulars of units or faulty methods of enumeration.
c. Inaccurate or inappropriate methods of interview, observation or measurement with inadequate or ambiguous schedules, definitions or instructions.
(Errors due to the above three factors are known as **specification errors**).
d. Lack of trained and experienced investigators
e. Difficulties involved in actual data collection arising from recall errors and other types of errors on the part of respondents.
f. Lack of adequate inspection and supervision of primary staff,
(Errors due to the above three factors are known as **ascertainment errors**)
g. Inadequate scrutiny of the basic data
h. Errors in data processing operations such as coding, tabulation, etc,. and
i. Errors committed during presentation and printing of tabulated results, graphs, etc.
(Errors due to the above three factors are known as **tabulation errors**).

The ascertainment errors and tabulation errors can be further classified as **coverage errors and contents errors**.

The above-mentioned errors do occur both in census surveys and sample surveys and these errors are termed as **Non-sampling errors**. In addition to these errors, in sample surveys, another type of errors occurs because of partial information based on the sample selected. These errors are called as sampling errors and these are inevitable. The sampling error usually decreases with increase in sampler size as shown in the following figure.




It is observed from the above figure that considerable greater effort is needed after a certain stage to decrease the sampling error than in the initial instances. From this point of view, there is a strong case for resorting to a sample survey to provide values within permissible margins of error instead of complete enumeration survey. As regards the non-sampling than in the case of sample survey, since it is possible to reduce the non – sampling error to a greater extent by using better organization and suitably trained personnel at the field and tabulation stages in the sample survey than in the census

survey. The behaviour of the non-sampling error with increase in sample size is likely to be opposite of that of sampling error. That is, non-sampling error is likely to increase with increase in sample size. In many situations, it is quite possible that the non-sampling errors taken together in a sample survey, and naturally in such situations, the latter is to be preferred to the former. A sample survey may also become a necessity in dealing with characteristics where serious non-sampling errors are expected, when special precautionary measures cannot be taken during collection and tabulation of data.

### 4.2.7 Estimates

A sample surveys has two important objectives: i) estimation of certain parameters of a population and ii) testing of certain hypotheses. The procedure of obtaining a value for the parameter on the basis of a sample observation is known as estimate of that parameter. For example, to find the populations mean we can use the sample mean. Here the mean of the sample selected is the estimate of the population mean. The difference between the estimate and the parameter value is termed as **sampling** error. Suppose't' is an estimate and $\theta$ is the parameter. In this case the sampling error is expressed as the difference of t and $\theta$ . From sample to sample't' takes different values. If $\theta$ is equal to the average value of t, then the average of squares of deviations of t values (based on all possible similar samples) about $\theta$ is known as the sampling variance of t. The positive square root of sampling variance is known as standard error of that estimate and usually denoted by $\sigma$ . If t is the mean of a simple random sample with replacement then the standard error (s.e) will be equal to $\sqrt{\frac{\sigma^2}{n}}$ where $\sigma^2$ is the population variance and n is the sample size. Similarly, the standard error of 't' in case of random sampling without replacement will be

$$\sqrt{\frac{N-n}{N-1}}\frac{\sigma}{\sqrt{n}}$$

The measure of standard error is based on the deviations of the estimate and parameter; sampling error will be expressed in terms of standard error. It is observed that standard error decreases with increase of sample size. To decide the sampling method, the measure of standard error of the estimate concerned is essential as mentioned earlier. The method, which will have smaller standard error, will be preferred normally.

### 4.2.8 Confidence Limits

The standard error gives an idea of the frequency with which errors of a given magnitude may be expected to occur if repeated random samples of the same size are drawn from the population. From knowledge of the standard error of the estimate and with the help of the normal probabilities, we can locate the actual unknown value of the parameter within certain limits with a known probability. Take the example of estimating the population mean. We know that the mean of a random sample will be approximately normally distributed if the size of the sample is not too small and if the population from which it is drawn is not very different from the normal distribution.
We may, therefore, expect that

$$\left|\overline{y_n} - \overline{y_N}\right| \le \frac{\sigma}{\sqrt{n}}$$

on an average in 68 out of 100 occasions, and

$$\left|\overline{y_n} - \overline{y_N}\right| \le 2\frac{\sigma}{\sqrt{n}}$$

on an average with a frequency of about 95 out of 100. (Here $\overline{y_n}$ is the sample mean, $\overline{y_N}$ is the population mean, $\sigma^2$ is the population variance).
In general we expect the interval

$$(\overline{y_n} - z_\alpha \frac{\sigma}{\sqrt{n}} \ , \ \overline{y_n} + z_\alpha \frac{\sigma}{\sqrt{n}})$$

to contain population mean $\overline{y_N}$ with certain probability. Here the probability depends on the value of $z_\alpha$. This type of interval which is likely to contain the population value with certain probability is known as confidence interval for that population value based on the sample drawn. Concept of confidence interval is very useful in fixing the sample size in probability sampling as described latter.

**4.3 Sample Size:** How large should the sample be? This question can be answered according to a prescribed list of sampling ideals. For most practical research projects, it is often difficult to satisfy these sampling ideals by offering a predetermined set of blanket generalization covering all aspects of how one should sample. The limitations of money, facilities, and staff are enough to jar any investigator into the position of doing the best with what he has, even though it is far from what he would want to do ideally.

An adequate sample is one that contains enough cases as to unsure reliable results. Hence, planning in advance for the size of the sample is very important. Several rules of thumb exist for estimating how large a sample should be. The most common is to obtain one-tenth of the population he studies in this sample. But there are other better methods, one being the following

The determination of sample size depends upon the desired precision and cost. Normally one of these will be fixed to find the suitable sample size to draw the sample units. The precision is surely specified in terms of the margin of errors permissible in the estimate and the coefficient of confidence with which one wants sure that the estimate is within the permissible margin of error. Thus, if the error permissible is the estimate of the population value of the mean is say D and the degree of assurance is then clearly we need to know the size of the sample , n so that,

$$P\{\left|\overline{y_n} - \overline{y_N}\right| \le D\} = \alpha$$

where $\overline{y_n}$ is the sample mean based on n observations and $\overline{y_N}$ is the population mean.

The above statement can be written as

$$P\{\frac{|\overline{y_n} - \overline{y_N}|}{S.E\ of\ \overline{y_n}} \leq \frac{D}{S.E.of\ \overline{y_n}}\} = \alpha$$

(Here S.E of $\overline{y_n}$ means standard error of $\overline{y_n}$ )

But $\frac{|\overline{y_n} - \overline{y_N}|}{S.E\ of\ \overline{y_n}}$ is a standard normal variant and therefore, for any given $\alpha$ we can obtain normal value, say $z_\alpha$ from the normal tables satisfying the above equation. Hence,

$$z_\alpha = \frac{D}{S.E\ of\ \overline{y_n}}$$

We know that in case of simple random sampling with replacement the standard error of $\overline{y_n}$ is equal to $\sigma / \sqrt{n}$ and therefore,

$$z_\alpha = \frac{D}{\sigma / \sqrt{n}}$$

or, $$n = \frac{z_\alpha^2 \sigma^2}{D^2}$$

Here the value of $\sigma^2$ will be obtained with previous knowledge or through pilot study. In this way, we can find the sample size, so as to keep the error in the desired limits. So to apply the formula for estimating sample size we must know the following things:

a. The level of confidence or the level of significance.
b. The tolerance error we will be willing to permit in the sample size estimate.
c. Some estimate of the sample standard derivation, a measure of variability or dispersion.

Suppose the cost is fixed at C. Now we can write $C=C_0+ n C_1$ where $C_0$ is the overhead cost in conducting the sample survey and $C_1$ is the average cost to deal with the sample units. Therefore,

$$n = \frac{C - C_0}{C_1}$$

$C_0$ and $C_1$ can be estimated with the past experience or through pilot survey. Here $C_1$ contains the cost of collection, tabulation, loss due to the decisions based on error etc.

Therefore, while deciding the sample size, generally we should try to minimize both sample error and non-sampling errors (errors occurring during the collection, analyzing the data etc.)

## 4.4 Sampling Method

The choice of the method of selection of a sample is generally based on the sampling error. *Sometimes, the sample will be selected on the prior knowledge and past experience and at other times based on the theory of probability. The sample selected on the basis of probability theory is known as probability samples. For example, ( i) random sampling, (with or without replacement, (ii) stratified random sampling, (iii) systematic random sampling,( iv) cluster sampling,( v) multistage sampling, etc.*

The different common non-random sampling procedures are: *(i) systematic sampling, (ii) quota sampling,(iii) purposive sampling, (iv) judgment sampling, (v) convenience sampling or chunk sampling, (vi) accidental or incidental sampling (vii) snowball sampling.*

Nature of the survey, size of the population, size of the sample, availability of funds, time and other resources would influence the selection of a particular method. However, the selection of probability sample once again depends upon the smaller sampling error. For the benefit of the readers, brief outline of each of the above sampling method is given in the following paragraphs.

### 4.4. 1 Probability Sampling

A method of selecting sample units such that each possible sample has a fixed and determine probability of selection is known as random selection. In random selection, if each possible sample has equal probability of selection then the method is known as simple random sampling and the sample so obtained is simple random sample.

**(i) Simple Random Sampling:** A simple way of obtaining a simple random sample is to draw the units one by one from the population units with equal probability of selection assigned to each unit of the population units with equal probability of selection assigned to each unit of the population at the first and each subsequent draw. The successive draws may be made with or without replacing the units selected in the preceding draws. The former procedure is called the simple random sampling with replacement and the latter is random sampling without replacement. Suppose 'n' units are to be selected from 'N' units of the population. In case of with replacement, the probability of drawing each unit of the population is equal to the reciprocal of the number of units in the population ie., 1/N. Similarly in case of without replacement, all the units in the population will have the probabilities of first, second,...Nth draw are 1/N, 1/(N-1), ....., 1 respectively. Random sampling will be done either by (i) a lottery method (or fishbowl method), (ii) a table of random numbers and (iii) computer-determined random sampling method. Because of their easiness and convenience, second and third methods

will be preferred. Generally, simple random sampling procedures will be preferred because;

1. The amount of sampling error associated with any given sample drawn can easily be computed
2. The investigator (or researcher) does not need to know the true composition of the population before hand
3. This method is used in conjunction with all other probability sampling plans and
4. Of all probability sampling plans, this method is the easiest to apply.

**(ii) Stratified Random Sampling**

To get a representative sample, the entire population will be divided into homogeneous groups or classes based on certain characteristics before selecting a random sample from each of these subgroups. These sub-groups are called strata and the combined sample obtained by this method is known as stratified random sample. A stratified sample is thus equivalent to a set of random samples of a number of such populations, each representing a single type or stratum.

A stratified random sampling may be either proportionate or disproportionate. In a proportionate stratified sampling, the number of items drawn from each stratum is proportionate to the size of the stratum (number of units in that stratum). On the other hand, if an unequal number of units drawn from each stratum regardless of how the stratum represented in the population, then such a sample obtained is known as disproportionate stratified sampling.

The primary use of disproportionate stratified random sampling is when elements exist few in number among certain strata. If there is a strong likelihood that one or more strata will not be included in a simple random sample of elements from strata will not be drawn. Then the researcher should take steps to see that elements those particular strata will be included. This detracts from the randomness of the draw, however, and to that extent is less a probability sample than before. The researcher exercises his judgment as to whether the particular strata is important enough so that steps should be taken to ensure that elements from it are included in subsequent sample. When the population under investigation can be stratified according to some characteristic with which the investigator has some familiarity, it is usually amenable to disproportionate stratified random sampling. However, while using disproportionate stratified random sampling, the researcher should keep in mind the following:
In a given stratum, one should take a larger sample if

i) The stratum is larger
ii) The stratum is more variable internally
iii) Sampling is cheaper in stratum (cost per sampling is less)

But generally, the relative variability and the relative costs in the strata are known before hand. However, these difficulties are overcome by using the information based on similar previous surveys or pilot surveys.

In most of the practical situations strata are the geographical regions or regions formed for administrative convenience or private and public institutions, etc. Whenever strata are not readily available, they can be formed such that each stratum is sufficiently homogeneous with respect to some characteristics like age, sex, and profession. There is no general rule to decide the number of strata.

When compared to simple random sampling procedure, stratified random sampling procedure ensures greater reliability and validity of the survey results.

**(iii) Systematic Random Sampling**

A systematic sample is one in which every kth unit (say 1 in 20 units is selected in a list representing a population. The number k is called the sampling interval. The units of population are arranged in systematic order on the basis of its important characteristics. The first number is chosen at random from the first k items. The value of k is approximately equal to the ratio of the number of units in the population and the size of the sample required (N/n). This is an easy procedure for an inexperienced investigator to follow. A systematic sample is usually more evenly spread over the entire population and thus can provide more information about the population than an equivalent amount of data contained in a simple random sample.

When k exceeds N/n, the number of observations (units) selected by the above procedure is less than n. In this case, we choose a random start from 1 to N and select units corresponding to this random start and thereafter every kth unit in a cyclical manner till a sample of n units is obtained. This procedure of drawing a sample is known as Circular Systematic Sampling

The Systematic sampling is generally used in those cases where a complete list of the population from which the sample is to be drawn is available. The operational simplicity of systematic random sampling over simple random sampling is of considerable importance in large scale sampling. Systematic sampling is not advisable when data is periodic. If the starting value is not chosen randomly then we call the procedure systematic sampling and this will be considered as non-probability sample.

**(iv) Cluster Sampling**

Cluster sampling involves division of the population units under consideration into groups or clusters that serve as primary units. A selection of the cluster is then made to make up the sample. Thus in cluster sampling unit contains groups of units instead of individual units in the population. For example, for the purpose of selecting 10 percent sample from all primary school children in Delhi, the investigator may list up all the primary schools and select randomly a 10 percent of schools. All the children in these selected schools are the units for enquiry.

Some times each cluster is further subdivided into sub-clusters and from each cluster some sub-clusters are selected to form sampling units. This procedure is known as two stage cluster sampling.

If the population units are geographical areas, cluster sampling is more appropriate and accordingly cluster sampling is also known as area sampling.

Cluster sampling is much easier to apply when large populations are studied or when large geographical areas must be covered. The cost of area of cluster sampling is much less compared with other sampling methods. In a multistage cluster sampling, it is possible to employ different forms of sampling in several successive stages. There is a larger sampling error associated with cluster sampling.

Public opinion polls are frequently conducted using cluster sampling. When lists of specific individuals are unobtainable or inaccessible, cluster sampling is recommended.

**(v) Multistage Sampling**

The above methods we have covered are the simplest probability sampling procedure. In most real applied social research, we would use sampling methods that are considerably more complex than these sample variations. The most important principle here is that we can combine the simple methods described earlier in a variety of useful ways that help us address our sampling needs in the most efficient and effective manner possible. Such combine sampling methods are known as multistage sampling. For example, in cluster sampling procedure, we may select the units from each cluster by stratified sampling procedure.

### 4.4.2 Non-Probability Sampling

**(i) Quota Sampling**

Quota sampling is defined as obtaining a desired number of units by selecting those most accessible to the investigator and those that possess certain characteristics. This sampling method is considerably less costly than most other sampling methods and this method is satisfactory when quick, crude results will satisfy the research objectives of the investigator. But generalization to population is not possible based on quota sample.

**(ii) Purposive sampling**

The researcher purposefully selects and also purposefully leaves units while selecting the sample units based on the objectives and other reasons. This method is suitable when there are only a small number of sampling units in the population. It is most suitable in urgent surveys.

**(iii) Judgment Sampling**

Judgment sample is one, which is selected mainly on the basis of expert's opinions. For example, if a sample of five students is to be selected from a class of forty for sampling to know study habits of students. The investigator would select five students with the help of class teacher. In certain situations these methods are preferred over simple random sampling. The purposive and quota sampling procedures form part of judgment sampling. Judgment samples are less costly and more readily accessible. They are also guaranteed inclusion of relevant units.

**(iv) Convenience Sampling**

Selecting 'convenient' population units is called convenience sampling. A sample obtained from readily available lists such as automobile registrations, telephone directories, etc. is a convenience sample. The results obtained by this method can hardly be representative of the population. They are generally biased and unsatisfactory. However, convenience sampling is often used for pilot studies.

**(v) Accidental or incidental Sampling**

Accidental or incidental samples are identical to quota sample with the following exceptions. Where as quota samples attempt to include units possessing apparent characteristics. Accidental samples make no such attempt. The investigator is guided mainly by convenience and economy. For example, the investigator may consider the first 50 persons who are willing to be interviewed or to provide the kind of information that he is seeking, he meets on any one of the pedestrian paths of a street. Where too much accuracy is not needed or where pre-occupation with tentative clues to hypothesis formulation (in exploratory surveys), accidental sampling is quite useful.

**(vi) Snowball Sampling**

In snowball sampling, we begin by identifying someone who meets the criteria for inclusion in our study. We then ask them to recommend others who they may know who also meet the criteria. Although this method would hardly lead to representative samples, there are times when it may be the best method available. Snowball sampling is especially useful when you are trying to reach populations that are inaccessible or hard to find. For instance, if we are studying the homeless, we are likely to be able to find good lists of homeless people within a specific geographical area. However, if we go to that area and identify one or two, we may find that they know very well whom the other home-less people in their vicinity are and how we can find them.

No doubt, non-probability sampling method has several weaknesses and many experts do not recommend it. Even then, in certain situations it becomes essential to follow this few sampling and in some cases it is preferred over probability sampling. Following are a few examples of such cases.

i) when the sampling frame is either not available or incomplete (ex. Selection of unrecognized schools)
ii) when the sample size is very small
iii) when in a pilot survey tryout of a questionnaire or a set of tools is required
iv) When the field work has to be done quickly in order to reduce memory errors

v) when reactions of an affected group of persons (from floods, riots, etc) are required

vi) to assess reaction of audience to some film or exposure audio-visual programme, a sample is needed.

# 5. Tools and Techniques of Data Collection

**Objectives:** After going through this module, you will be able to:

(a) Understand the concept of scale and its types
(b) Understand different kinds of tools and techniques that can be used in research
(c) Understand qualities of a good research tool

**5.1 Introduction**: Knowledge originates in information that can be received directly from observation or indirectly from reports of observations. Anything we hear, read, smell or otherwise experience could become part of our knowledge. If it is remembered, it does become knowledge. But if it is only remembered, without being thought about, it remains mere information, the most elementary and least useful form of our knowledge. If, on the other hand, information becomes the subject of our reflective thought, if we ask ourselves, "what does it mean?" "How do we know?" "Why is it so?" "We may come to understand the information. It can be integrated into a system of relations among concepts and ideas, all of which constitute a structure of knowledge. This process of encoding is essential to enable later retrieval; observations that are not encoded in some way cannot be recalled.

In our daily life, we only distinguish the character and qualities of things, and on the basis of this, we reject or chose certain things. In day-to-day affairs, it is necessary to make judgment upon qualities, which cannot ordinarily be differentiated from one another. Propositions affirming qualitative differences are the first fruits of enquiry in science (Crhen and Nagel). But it is not enough to know only the qualitative differences in daily life or in sciences. It is more necessary to know the magnitude of such differences precisely. This leads to the theory of measurement, dealing with the principles employed in representing quantities and relationships in terms of numerical values and correlations.

**Measurement will be performed at four levels as given below.**

**(i) Nominal scale:** A nominal scale describes differences between things by assigning them to categories - such as Professors, Readers, Lecturers – and subsets such as males or females. Nominal data are counted data: Each individual can be a member of only one set, and all other members of the set at the same defined characteristic. Such categories as nationality, gender, socioeconomic status, race, occupation or religion prove examples. Nominal scales are non-orderable, but in sometimes this simple enumeration or counting are the only feasible method of quantification and may provide an acceptable basis for statistical analysis.

**(ii) Ordinal scale:** Sometimes it is possible to indicate not only that things differ but also that they differ in amount of degree. Ordinal scale permits the ranking of items or individuals from highest to lowest. The criterion for the highest and lowest ordering is expressed as relative position or rank in a group. Ordinal measures have no absolute values and the real differences between adjacent ranks may not be equal. The ranking spaces them equally, though that may not actually be equally spaced.

**(iii) Interval scale:** An arbitrary scale based on equal units of measurement indicates how much a given characteristic is present. The difference in amount of the characteristic possessed by persons with scores of 90 and 95 is assumed to be equivalent to that between persons with scores of 60 and 65. The interval scale represents a decided advantage over nominal and ordinal scales because it indicates the relative amount of trait or characteristic. Its primary limitation is the lack of a true zero. It does not have the capacity to measure the complete absence of the trait, and a measure of 90 does not mean that the person has twice as much of the trait of someone with a score of 45. Psychological tests and inventories are interval scales and have this limitation, although they can be added, subtracted, multiplied, and divided. A good example is the scale by which we usually represent temperature. One unit (degree) increase in temperature is defined by a particular change in volume of mercury in a thermometer. Consequently, the difference between any two temperatures may be measured in units, or degrees.

**(iv) Ratio Scale: A** ratio scale is used when not only the order and interval size is important, but also the ratio between the two measurements is meaningful. The only distinction between the ratio scale and the interval scale is that the ratio scale has a natural measurement that is called zero, while the zero measurement is defined arbitrarily in the interval scale. The numerals of the ratio scale have the qualities of real numbers and can be added, subtracted, multiplied, and divided and expressed in ratio relationships. For example, 5 grams is one half of 10 grams; 15 grams are 3 times of 5 grams.

In behavioural research, many of the qualities or variables of interest are abstractions and cannot be observed directly. It is necessary to define them in terms of observable acts, from which the existence and amount of the variables are inferred. This operational definition tells what the researcher must do to measure the variable. For example, intelligence is an abstract quality that cannot be observed directly. Intelligence may be defined operationally as scores achieved on a particular intelligence test. Progress is being made in developing more valid operational definitions and better observation techniques. The quantitative approach is not only useful but may be considered indispensable in most types of research.

**5.2 Data collection tools:** To carry out any type of research, data must be gathered for which tools or techniques are employed. The nature, purpose and scope of a research study greatly determine the choice of method to be adopted for collecting data. Availability of time and finance also governs the choice of method. The word 'tool' literally means implement or appliance for mechanical operation. But in educational evaluation, it is a means of collecting evidence for students' performance. All the 'means', which are used to measure the students' behaviours precisely and accurately, are

called the tools of evaluation. The tools of evaluation can be variously used according to the needs and purposes of the evaluator and the kinds of behaviors to be evaluated.
***The major data gathering tools of research may be classified broadly into the following categories*:**

1. **Tests**
2. **Inquiry forms**
3. **Observation**
4. **Interview**
5. **Sociometric Techniques**

**1. Tests**

A test is an instrument of systematic procedure for measuring a sample of behaviour. Tests designed by test specialists and administered, scored and interpreted under standard conditions are called standard tests. The tests constructed by teachers are known as informal or teacher- made tests. The teacher- made and standard tests may be classified as objective type and essay type. Objective type is further classified into:

a. Supply type
   (i) Short answer
   (ii) Completion

b. Selective type
   (i.) True / False or alternative response
   (ii.) Matching
   (iii.) Multiple Choice
c. Essay type is further classified into:
   (i) Extended response type
   (ii) Restricted response type.

The objective in standardizing a test is to obtain scores in some objective form, which can be used as a basis upon which to assess each individual respondent's performance that is to estimate the goodness or badness of his performance when compared with other respondents who took the test. The five different forms of standardized objective tests, which normally a teacher is concerned, are:

(i) Tests of general intelligence or mental ability
(ii) Tests of attainment or achievement in school subjects
(iii) Tests for the diagnosis of backwardness in a school subjects
(iv) Tests of special ability and
(v) Tests of prognosis in respect of future progress in school subjects.

Intelligence test is the tests by which we can obtain reasonably accurate measure of child's power to think accurately or his ability to learn. An attainment or achievement tests is a test, which measures acquired knowledge in a particular field. It provides objective assessment of the levels reached by pupils in various school subjects. The diagnostic test is constructed not to assess levels but to reveal difficulties in school

subjects. Tests of special ability aim at revealing the degree of development and also, to a lesser extent the latent power of individuals in specific field. Prognostic test in a school subject is designed before hand the degree of success that pupils are likely to obtain in studying that subject.

A knowledge of standardized tests should be a part of an adequately trained at teachers' classroom equipment. Those who use tests should make sure of the purpose for which it is prepared. In the absence of standardized tests, teacher-made tests will serve the purpose. The teacher-made test is an indicator of what an individual teacher considers most important. The teacher's tasks are indicative of his or her basic educational objectives.

Teacher-made tests can be classified into oral tests, written tests and performance tests. Excepting the oral tests, the other teacher-made tests can be standardized by using standardization procedures like reliability and validity.

**Oral examination**: The oral method of examination is the oldest form of testing. It serves as the good evaluation technique, and as an informal means of appraising the progress of students. Oral tests can serve as an excellent means for the following purposes.

(i) To diagnose students difficulties
(ii) To determine how well an individual student has integrated his or her knowledge
(iii) To determine how well a student can apply the knowledge to various situations
(iv) To evaluate certain performances
(v) To measure the achievements in case of young children who have not yet learnt how to read or write full sentences.

Advance planning for construction of oral test produces better results.

**Written tests**: The main considerations of these tests are:

(i) Determining the purpose of testing
(ii) Developing the test specifications
(iii) Selecting appropriate item types
(iv) Preparing relevant test items

The classroom tests can be used to determine:

(i) Whether pupils have the pre-requisite skills needed for the construction (to determine readiness) or
(ii) To what extent pupils have already achieved the objectives of the planned instruction (to determine pupil placement) or
(iii) To monitor learning progress (formative tests) or

(iv) To measure the extent to which the intended learning outcomes have been achieved (summative testing).

To provide assurance that classrooms test will measure a representative sample of instructionally relevant tasks, some type of test specifications can be used. One device that has been widely used to this purpose is a two-way chart, called a table specification or test blueprint. This table contains the various instructional objectives and outline of course content along with the columns and arrows in general.

The items used in classroom tests are divided into two categories (i) objective items and (ii) essay questions.

The construction of items for a classroom tests should be preceded by a series of preliminary steps as mentioned earlier and finally the test items should be constructed in accordance with the decisions in the preceding steps. A classroom test is most likely to provide a valid measure of the instructional objectives, if the test items are designed to measure the performance defined by the specific learning outcomes. A test, no matter how extensive, is almost always a sample of the many possible test items that could be included. The difficulty of the items to be included in a classroom test depends largely on whether the test is being designed to describe the specific learning tasks pupils can perform (criterion referenced) or to rank the pupils in order of their achievement (norm referenced). In item construction, care must be taken to eliminate any extraneous factors that might prevent pupils from responding. Test items should be constructed so that pupils obtain the correct answer only if they have attained the desired learning outcome.

**General suggestions for writing test items:**

(a) Use your test specifications as a guide to item writing.
(b) Write more test items that needed.
(c) Write the test items well in advance of the testing date.
(d) Write each test items so that it calls for the performance described in the intended learning outcomes.
(e) Write each test items so that the task to be performed is clearly defined.
(f) Write each test items at an appropriate reading level.
(g) Write each test item so that it does not provide help in answering other items in the test.
(h) Write each test item so that the answer is one that would be added upon by the experts.
(i) Write each test item so that it is at the proper level of difficulty.
(j) Whenever a test item is revised, recheck its relevance.

**Achievement Tests:** Achievement tests attempt to measure what an individual has learned – his or her present level of performance. Achievement test scores are used in placing, advancing or retaining students at particular grades or levels. Some achievement tests are aimed to measure the pupils' level of achievement in various content and skill areas by comparing their test performance with the preference of other

pupils in some general reference group (eg. A nation wide sample of pupils at the same grade or level). These tests are known as norm – referenced tests. Some other achievements aimed at to measure the basic skill areas, especially reading and mathematics and these are known as criterion-referenced achievement tests.

Standard achievement tests are familiar in educational testing.

**Characteristics of Standard Achievement Tests**

i) The test items are of a high technical quality.
ii) Directions in administering and scoring are so precisely stated that the procedures are standard for different users of the test.
iii) Norms based on representative groups of individuals are provided as aids in interpreting the test scores.
iv) Equivalent and compatible forms of the test are usually provided as well as information connecting the degree to which the forms are comparable.
v) A test manual and other necessary material are included as guides.

The different achievement tests in specific areas are:
a) Content-oriented tests
b) Reading tests
c) Readiness tests

Test publishers are now making available customarized achievement tests, for individualized instruction and mastery learning. Some publishers also provide item banks and software programs that enable schools to produce their own customized tests with micro-computers.

**Aptitude Tests**

Aptitude tests are designed to predict future performance in some activity. These used in schools range from the traditional scholastic aptitude tests to the more comprehensive differential aptitude tests. Aptitude tests attempt to predict an individual's capacity to acquire improved performance with additional training. These inferred measurements have been applied to mechanical and manipulative skills, musical and artistic pursuits and many professional areas involving many types of predicted ability. In music, for example, ability to remember and discriminate between differences in pitch, rhythm pattern, intensity and timber seems to be closely related to future levels of development in musicianship.

To the extent that aptitude tests measure past learning, they are similar to achievement tests. To the extent that they measure non-deliberate or unplanned learning, they are different.

**Projective Technique**

A projective instrument enables subjects to project their internal feelings, attitudes, needs, values or wishes to an external object. Thus the subjects may

unconsciously reveal themselves as they react to the external object. The use of projective method is particularly helpful in counteracting the tendency of subjects to try to appear in their best light, to respond as they believe they should. Projection may be accomplished through a number of techniques : (1) Association- where the respondent is asked to indicate what he or she sees, feels or thinks when presented with a picture, cartoon, word or phrase. (2) Completion- where the respondent is asked to complete an incomplete sentence or task. (3) Role-playing- where subjects are asked to improvise or act out a situation in which they have been assigned various roles. (4) Creative or Constructive- where subjects are permitted to model clay, finger print, play with dolls, play with toys or draw or write imaginative stories about assigned situations.

Many learning outcomes pertaining to knowledge, understanding and application can be measured by paper-and-pencil tests. But outcomes like:

i) Skill (speaking, writing, listening and reading)
ii) Performing laboratory experiments
iii) Drawing
iv) Playing a musical instrument
v) Dancing
vi) Gymnastics
vii) Work skills
viii) Study skills
ix) Social skills
x) Work habits (effectiveness in planning, use of time, use of equipment, use of resources, demonstration of such traits as initiative, creativity, persistence and dependability).
xi) Social attitudes (concern for the welfare of others, respect for laws, respect for the property of others, sensitivity to social issues, concern for social institutions, desire to work toward social improvement).
xii) Scientific attitudes (open mindedness, willingness to suspend judgment, sensitivity to cause-effect relations, and an inquiry mind).
xiii) Interests (expressed feelings towards various educational, mechanical, aesthetic, scientific, social, recreational, vocational activities).
xiv) Appreciations (feeling of satisfaction and enjoyment expressed towards nature, music, art, literature, physical skill, outstanding social contributions).
xv) Adjustments (relationship to peer, reaction to praise and criticism, reaction to authority, emotional stability, social adaptability)

are difficult to evaluate with the usual paper-and-pencil test. Those outcomes are generally be evaluated by (i) observing pupils as they perform and describing or judging that behaviour, (ii) observing and judging the quality of the product resulting from their performance, (iii) asking their peers about them and (iv) questioning them directly.

The observational techniques, which are useful to a teacher, are (i) anecdotal records, (ii) rating scales and (iii) checklists.

Anecdotal records are factual descriptions of the meaningful incidents and events that the teacher has observed in pupils' lives. Each incident should be written down shortly after it happens. A good anecdotal record keeps the objective description of an incident separate from an interpretation of the behaviour's meaning. The use of anecdotal records has frequently been limited to the area of social adjustment. It is not possible to observe and report on all aspects of pupil's behaviour, no matter how useful such records might be.

In contrast with the unstructured descriptions of behaviour gathered in anecdotal records, rating scales provide a systematic procedure for reporting observers' judgments. 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. Rating scale will direct observation towards specific aspects of behaviour, it will provide a common frame of reference for comparing all pupils on the same set of characteristics; and it will provide a convenient method for recording the observers' judgments. Rating scales are of two types – numerical and descriptive graphic. Rating scales are especially useful in evaluating procedures because they focus on the same aspects of performance in all pupils and have a common scale on which to record our judgments. If the rating form has been prepared in terms of specific learning outcomes, it also serves as an excellent teaching device.

A checklist is similar in appearance and use to the rating scale. The basic difference between them is in the type of judgment needed. Checklist calls for a simple yes-no judgment. The checklist is especially useful at the primary level, where much of the classroom evaluation depends on observation rather than testing.

Sometimes peer judgment will be the valuable tool in evaluation.

**Interview Method**

The interview method is a verbal method of securing data especially in the field research. An interview is a face-to-face conversational exchange where one person elicits information from the other. The person who is interviewing is called an interviewer (or investigator) and the person who is giving interview is called interviewee or respondent or informant. If the interviewer (researcher) directly contacts the informant to collect information, then the method is known as direct personal interview. In some situations, the interviews are conducted by some specially trained persons, at other times, the interviewer instead of directly approaching the concerned informants, interviews several third persons who are directly in touch with the information sought. For example, in a study regarding the study habits in a school, the students may not give proper correct information about their own study habits. In such cases, the desired information can be obtained from other students who may know the habits of the concerned. Such of these interviews are known as indirectly personal interviews. Personal interviews are also conducted through telephones.

The objectives of an interview are (i) to exchange of ideas and experiences and (ii) to elicit information pertaining to a very wide range of data.

Several factors have a direct bearing on the usefulness of interviews in achieving the scientific objectives for which they are employed. Three of them are (i) the qualities of the interviewer, (ii) the qualities of the interviewee and (iii) the nature of the problem of study. Interviewer's characteristics such as sex, age, religion, social class, patience, manner of dress and speech are important in the interview.

The various types of interviews are:

i) Formal or informal (based on formality)
ii) Personal or group (based on the number of interviewees)
iii) Diagnostic or treatment or research or to fulfill curiosity (based on purpose).
iv) Short or prolonged (based on period of contact)
v) Quantitative and qualitative or mixed (based on subject matter)
vi) Non-directive or focused or repeated (based on role)

In general, interview method is considered as a powerful tool for data collection in majority of the surveys, because of the following reasons:

i) The interview method usually yields a high percentage of returns (responses), as most people are willing to cooperate.
ii) The information secured through interviews is likely to be more correct than that secured by other techniques since the interviewer can clear up seemingly inaccurate answers by explaining the questions to the informants.
iii) The interviewer can collect supplementary information about the informant's personal characteristics and environment which is variable in interpreting results and evaluating the representative of persons surveyed.
iv) Scoring and test devices can be used, the interviewer acting as experimenter.
v) Visual material to which the informant is to react can be presented.
vi) Return visits to complete items on the schedule or to correct mistakes can usually be made without annoying the informant.
vii) The interviewer may catch the informant off guard and thus secure more spontaneous reactions than would be the case otherwise.
viii) The interviewer can usually check which person or persons answer the questions. Group discussion can be held with the personal interview method, if desired.
ix) The personal interview may allow the informant to become oriented to the topic under investigation.
x) Questions about which the informant is likely to be sensitive can be carefully sandwiched in by the interviewer. A delicate situation can usually be handled more effectively by a personal interview than by other survey techniques. For example, questions like "Do you avoid tax payment?", "Does your wife love you as much as you would like?"
xi) In short span of time, more and more information can be elicited by interviews.
xii) The language can be adopted to the ability or educational level of the person interviewed. Therefore, it is comparatively easy to avoid misinterpretations or misleading questions.

Following are the limitations of the interview method:

i) The transportation costs and the time required to cover addressees in a large area may make the personal interview method unfeasible.
ii) Unless the interviewers are properly trained and supervised, the data recorded may be inaccurate and incomplete.
iii) In some situations, interviewer's personal presence may lead to wrong information. In such cases, telephone interview may be used.

In interview method, the researchers should be careful about the following to get the reliable results:

i) The way the interviewer presents the purpose and scope of the study
ii) The way the questions are asked
iii) The way inadequate answers are probed
iv) The way answers are recorded and
v) The way the interpersonal aspects of the interview are handled.

This can be achieved by proper selection and training of interviewers. The training should cover the topics like:

a) Procedures for contacting respondents and introducing the study
b) The conventions that are used in the design of the questionnaire with respect to wording and skip instructions so that interviewers can ask the questions in a consistent and standardized way
c) Procedures for recording inadequate answers in a non-directive way
d) Procedures for recording answers to open-ended and closed questions
e) Rules and guidelines for handling the inter-personal aspects of the interview in an unbiased way
f) Specific purposes of the survey including the sponsorship, the general research goals and anticipated uses of the research
g) The specific approach that is used to sampling (in sample surveys), again to provide a basis for answering respondent questions
h) Details regarding the purposes of specific questions
i) The specific steps which are taken with respect to confidentiality and the kinds of assurances that are appropriate to respondents
j) Dealing with non-response cases

**Questionnaire Method**

The most common problem regarding observation and interviewing is the personal bias of the observer and or interviewer. The observer may misinterpret what he sees or he may fail to see something that is important to the group he is observing. The interviewer may misreport what he hears in an interview. Also, it is likely that the observer will influence the individuals when he is observing.

Next to the interview method, the schedule and the questionnaire methods are the popular and widely used ones in collecting data from primary sources. The difference between

the schedule and the questionnaire lies in the method of filling the form containing a set of questions. A questionnaire is self-administered whereas a schedule is not. Questionnaire or schedule deals with a variety of questions for the purpose of having a reliable and valuable data. The questions are generally aimed at (a) ascertaining facts, (b) ascertaining belief about what the facts are, (c) ascertaining feelings, (d) discovering standards of action, (e) exploring present or past behaviours and (f) knowing the conscious reasons for beliefs, feelings, policies or behaviours.

Questionnaires may be classified into structured and non-structured questionnaires. The structured questionnaire contains definite and direct questions and non-structured questionnaire is often called interview guide, which is used for depth and non-directive interview.

There are basically two methods for administering questionnaires – the mailed questionnaire and the face-to-face questionnaire administration. In the mailed questionnaire method, the previously designed questionnaire along with the instruction to fill is mailed to the respondents and the filled in questionnaire are obtained back after some period and if necessary after some reminders. The face-to-face method of questionnaire administration requires that predetermined subjects be given questionnaires to complete in the presence of the investigator or enumerator. The major benefit of a mailed questionnaire is economy, privacy and degree of anonymity. But the drawback of this method is the rate of non-response and the reliability of the respondent. The major advantages of administering a questionnaire in a face-to-face situation are (a) a high rate response, (b) reliability of the respondent, (c) on the spot scrutiny and modification for the unanswered questions, etc. But this method is very time consuming.

Responses to questionnaires may be (i) fixed or closed, (ii) open-end or (iii) a combination of closed and open end. Fixed response questionnaires consist of items with a fixed number of choices. In these, the questions provide for making a 'Yes' or 'No', a short response or checking an item out of a list of given responses. It restricts the choice of response for the respondent. He has simply to select a response out of supplied responses and does not frame his response in his own way. It is easy to fill out, takes less time, keeps the respondent on the subject, is relatively more objective, more acceptable and convenient to the respondent and is fairly easy to tabulate and analyze. Open-end questionnaires are characterized by questions that require short or lengthy replies by respondents. This form provides for greater depth of response. This form provides for greater depth of response. Sometimes the responses in this case are difficult to interpret, tabulate and summarize in the report. Practically many questionnaires are made up of items with both fixed and open-end responses.

Based on their uses, questionnaires are classified as (a) rating questionnaire, (b) observation questionnaire, (c) interview questionnaire and (d) document questionnaire.

In questionnaire, the quality of the response mainly depends upon the quality of the questionnaire. Designing a good questionnaire involves selecting and framing the questions needed to meet the survey objectives, testing them to make sure that they can

be asked and answered as planned, then putting them into a form to maximize the ease with which respondents and interviewers can do their jobs.

The following questions should be kept in mind while preparing questionnaire for data collection.

a) What is the definition of the population about which we seek information?
b) What is the socio-economic and or educational level of the intended target (respondents) of our enquiry?
c) What kinds of facts do we wish to learn about them?
d) How accessible are the respondents?
e) How will the questionnaire be administered?
f) What kinds of response patterns will be used?
g) How long should we make the questionnaire?
h) How much control can we exert over ensuring their response to our questionnaire?

Questionnaire is to be designed at a particular level of readability commensurate with that of the respondents. The shorter the questionnaire, the greater the response rate. Extremely lengthy questionnaires can cause some respondents to become test-weary. Tired respondents may become somewhat careless in the answers they provide.

Physical appearance of the questionnaire, contents of the questionnaire, subject matter, quality and size of the paper, type of printing, arrangements of items on the questionnaire are to be considered carefully in order to get a good response.

The language should be simple; words of precise meaning should be used; an expression of appreciation for the efforts put forth by the respondent should be made; questions as reliable as possible should be made. Multiple questions should be asked with different question forms that measure the same subjective state; to make provision to combine the answers into a scale.

There are four practical standards that all questions should meet.

i) Is this a question that can be asked exactly the way it is written?
ii) Is this a question that will mean the same thing to every one?
iii) Is this a question that people can answer?
iv) Is this a question that people will be willing to answer, given the data collection procedures?

Once a set of questions is ready, they need to be put into a form to facilitate interviewer or respondents. Questions requiring a lot of thinking or which may be sensitive are often reserved for the middle or later part of the questionnaire.

Quality of the questionnaire takes care of the quality of the response.

### Observation

When it is important to see behaviour in its natural situation, to know the features of conduct in different situations, some form of observation becomes essential as the primary method of acquiring information. Observation is like other instruments of data collection in that it demands mastering certain skills if it is to be used effectively. A major purpose of observation is to capture human conduct as it actually happens, to permit us to view behaviour in process. Observation is a careful and systematic watching of facts as they occur in nature with regard to the mutual cause-effect relations. In every set of observations, there are two components – the object or what is observed and the subject or the observer. An observer may be a visiting stranger, a participant observer, an attentive listener or an eager leaner.

Observation involves: (i) sensation, (ii) attention and (iii) perception. Sensation merely reports facts as observed but perception enables the mind to recognize the facts. Observation is at once a physical as well as mental activity. Observation is also selective and purposeful in the sense that one does not observe anything and everything.

Observation must be, carefully planned, systematic and perceptive. For this, the investigation must determine in advance what to observe or what types of phenomena merit recording. The formulation of the problem and the specific points needing investigation dictate what should be observed and provide some controls over the observer.) By thinking through the problem and determining what to observe, by devising a simple method of recording his observations and by deliberately concentrating upon the specific aspects which interest him, the investigator will obtain more reliable and authentic data than he will by observing in a haphazard and hurried manner. In other words, observers should be objective. They should recognize their likely biases and strive to eliminate their influence upon what they see and report. Whenever qualitative values are to be observed, each should be rated and recorded independently of others by use of a well-defined rating scale. To conserve time and to standardize the procedures for multiple or independent observations, it is highly important to determine the informational and statistical units to be used in recording observations. To help speeding up the recording process, an observation schedule or checklist can be designed. Observers should observe the facts and make interpretation at a later time. Instruments such as camera, stopwatch, audiometer, audio and video tape recorders, maps, sociometric scales, etc. can be used for more precise observations.

The data collected by observation are more accurate and reliable because they are based on first-hand perception by the eyes.

### Sociometry

Sociometry is a technique for describing the social relationships among individuals in a group. It attempts to describe attractions or repulsions between individuals by asking them to indicate whom they would choose or reject in various

situations. Sociometric choices may be represented graphically on a chart known as a **sociogram.** There are many versions of the socio pattern. A process of description closely related to sociometry is the **'guess-who' technique** developed by Hartshorne and May (1929). In this process, children are asked to name the individuals who fit certain verbal descriptions like 'This one is always happy', 'This one will always help you'. Items of this type of yield interesting and significant peer judgments and are useful in the study of individual roles.

Another approach to the description and measurement of social relationships is the **social-distance scale,** developed by Bogardus (1933). This device attempts to measure to what degree an individual or group of individuals is accepted or rejected by another individual or group by specifying as 'complete acceptance', 'partial acceptance' or rejection.

## 5.3 Qualities of a Good Research Tool

Research is always dependent on the measurement. The important qualities of a good research tool are:

i) Validity
ii) Reliability
iii) Usability
iv) Ease of administration
v) Time required for administration
vi) Ease of interpretation and application
vii) Availability of equivalent or comparable forms
viii) Cost of testing

**(i) Validity**

A tool is valid if it measures what it claims to measure. Validity can also be thought of utility. There are several types of validity, and different types of tests and uses of tests need different types of validity.

(a) **Content Validity** refers to the degree to which the test actually measures, or is specially related to, the traits for which it was designed. It shows how adequately the test samples the universe of knowledge and skills that a student is expected to master. Content validity is based upon careful examination of course textbooks, syllabi, objectives and the judgments of subject matter specialists. Panel of experts for its adequacy often assesses content validity without using any numerical measure. Achievement tests require this validity.

(b) **Construct validity** is the degree to which scores on a test can be accounted for by the explanatory constructs of a sound theory. How well test performance can be interpreted, as a meaningful measure of some characteristic or quality is known as the construct validity. To verify this validity establish the meaning of the scores on the test by controlling (or examining) the development of the test, evaluating the relationships of the scores with other relevant measures, and experimentally determining what factors influence test performance.

(c) **Criterion-related validity:** Whenever test scores are to be used to predict future performance on some valued measure other than the test itself (called a criterion), we are concerned with criterion-related evidence. For example, reading readiness test scores might be used to predict pupils' future achievement in reading. Criterion-related validation may be defined as the process of determining the extent to which test performance is related to some other valued measure of performance. This validity actually refers to two different types of validity with different time frames – (i) predictive validity refers to the usefulness of a test in predicting some future performance, such as the usefulness of the high school scholastic aptitude test in predicting college grade – point averages and (ii) concurrent validity refers to the usefulness of a test in closely relating to other measures, such as present academic grades, teaching ratings or scores on another test of known validity.

Tests are often validated by comparing their results with a test of known validity.

**(ii) Reliability**

A tool is reliable to the extent that it measures whatever it is measuring consistently. The reliability or stability of a test is usually expressed as a correlation coefficient.

i) Stability over time (test-retest): The scores on a test will be highly correlated with scores on a second administration of the test to the same subjects at a later date if the test has good test – retest reliability.

ii) Stability over item samples (equivalent or parallel forms): Some tests have two or more forms that may be used interchangeably. In these cases, the scores on a test will be highly correlated with scores on an alternative form of the test if the test has this type of reliability.

iii) Stability of items (internal consistency): Scores on certain tests will be highly correlated with scores on other test items. There are two methods of measuring for internal consistency – split halves and Kuder-Richardson formula.

Split halves: This can be accomplished in two different ways. Scores on the odd-numbered items can be correlated with the scores on the even-numbered items.

Second, on some but not most tests, the scores on the first half of the test can be correlated with scores on the second half of the test. Generally, the longer a test is the more is internal consistently. Correlation coefficient can be modified by using the Spearman – Brown formula.

Kuder-Richardson formula: This formula is a mathematical test that results in the average correlations of all possible split half correlations.

iv) Stability over scores (inter-scorer reliability): This can be determined by having two persons independently score the same set of the papers and then calculating a correlation between their scores, determined by the scores.

v) Standard error of measurement: If it were possible to test a pupil over and over again on the same test, we would find that the scores would vary somewhat. The amount of variation in the test scores would be directly related to the reliability of the testing procedures. Low reliability would be indicated by large variations in the pupil's test scores and high reliability would be indicated by little variation from one testing to another. Although it is impractical to administer a test, many times to the same pupils, it is possible to estimate the amount of variation to be expected in test scores. This estimate is called the standard error of measurement and is used to measure the reliability of the test.

Factors influencing reliability measures are: (a) length of test (b) spread of scores (c) difficulty of test (d) objectivity

**(iii) Usability:** The use of the tests should be practicable.

**(iv) Ease of administration:**

If the tests are to be administered by teachers or others with limited training, ease of administration is an essentially important quality to seek in a test. For this purpose, the directions should be simple and clear, the subtests should be relatively few and the timing of the test should not be too difficult.

**(v) Time required for Administration**

We always favour the short test, other things being equal. But in this case, other things are seldom equal, because reliability is directly related to the test's length. A safe procedure is to allot as much time as is necessary to obtain valid and reliable results and no more. Somewhere between 20 and 60 minutes of testing time for each test is probable a fairly good guide.

**(vi) Ease of Interpretation and Application**

The success or failure of a testing programme is determined by the use made of the test results. If they are interpreted correctly and applied effectively, they will contribute to more intelligent decision. When the test results are to be presented to pupils or parents, ease of interpretation and application are especially important.

**(vii) Availability of Equivalent or Comparable Forms**

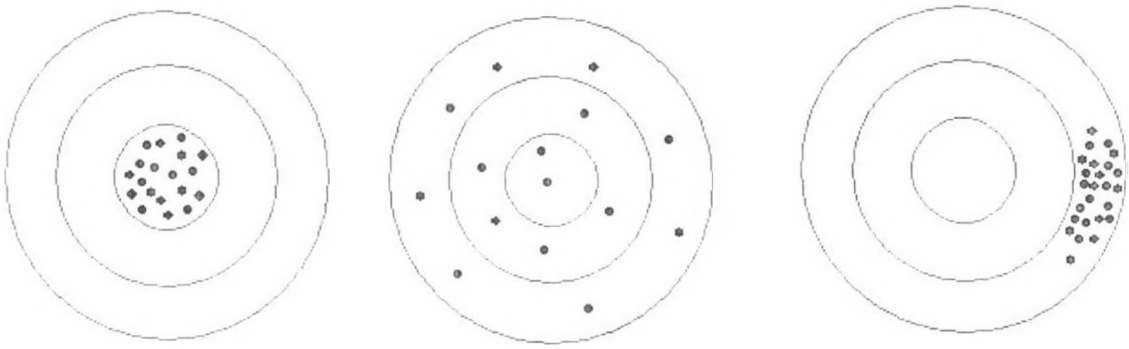
Equivalent forms of a test measure the same aspect of behaviour by using test items that are alike in content, level of difficulty, and other characteristics. Thus, one form of the test can substitute for the other, making it possible to test pupils twice in rather close succession without their answers on the first testing, influencing their performance on the second testing. The advantage of equivalent forms is readily seen in mastery testing in which we want to estimate the factors of memory while retesting pupils on the same domain of achievement. Equivalent forms of a test also may be used to verify a questionable test score. For example, a teacher may feel that a scholastic aptitude or achievement test score is spuriously low for a given pupil and

may easily check this by administering an equivalent form of the test. Many tests also provide comparable forms. These are especially useful in measuring development in the basic skills.

**(viii) Cost of Testing**

Testing is relatively inexperience, and cost should not be a major consideration.

Reliability is necessary but not sufficient for validity. Observe following figures regarding shooting incidents.



Reliable and Valid  Unreliable and Invalid  Reliable and Invalid

## References


1. Best, John, W. and James V. Kahn : *Research in Education* (7th Edition).
2. Koul, Lokesh. *Methodology of Educational Research* (2nd Edition).
3. Grounlund, Norman E. and Linn, Robert L. *Measurement and Evaluation in Teaching* (6th Edition).
4. Ronal d Czaja and Johny Blair : *Designing Surveys – A Guide to Decisions and Procedures.*
5. Saravanavel P : *Research Methodology.*

# 6. Educational Data Processing

**Objectives:**

After reading this module, the reader will be able to:

i) Understand the role that statistical data analysis plays in Education.
ii) Recognize the important characteristics, assumptions, terms and concepts of statistical data analysis.
iii) Calculate the important descriptive statistics.
iv) Perform the correlation analysis in comparing the similar type of data.
v) Understand the principles of statistical inference.
vi) Perform various statistical data analysis using calculators or some computer software.
vii) Interpret the statistical results.

**6.1 Introduction:** To carry out any research investigation, data must be gathered with which to test the hypothesis. Many different methods and procedures will be developed and used in the acquisition of data. These tools employ distinctive ways of describing and quantifying the data. Such data are only crude information and not knowledge by themselves. Data originally collected for an investigation are known as primary data. Generally the primary data will be collected through observation, interview and schedule/questionnaire methods.

Research is the systematic collection, analysis and interpretation of facts relating to a specific problem. Statistics, as a tool in research, deals with methods of collecting and interpreting numerical facts (data). Many of the educational problems, which invite research, relate to the measurement of differences between individuals, results of instructions, or organization and administration of the school or of institute. Statistics as a tool enable us to classify, organize and summarize numerical facts so that they can be more readily comprehended and interpreted. Statistics also enable us to draw conclusions with certain degree of exactness, based upon the less than complete evidence. Here an attempt has been made to explain the common statistical methods that are useful in Educational researches like, experimental, action researches.

**6.2 Stages of Statistical Analysis:**

Studying a problem through the use of statistical data analysis usually involves four basic steps:

1. Defining the problem.
2. Collecting the data.
3. Analysing the data.
4. Reporting the results.

For the purpose of statistical data analysis, distinguishing between cross-sectional and time-series data is important. Cross-sectional data are data collected at the same or approximately the same point of time. Time-series data are data collected over several time periods.

Data can be collected from existing sources or obtained through observation and experimental studies designed to obtain new data. In an experimental study, the variable of interest is identified. Then one or more factors in the study are controlled so that data can be obtained about how the factors influence the variables. In observational studies, no attempt is made to control or influence the variables of interest. A survey is perhaps the most common type of observational study.

Statistical data analysis, divide the methods for analysing data into two categories: exploratory methods and confirmatory methods. Exploratory methods are used to discover what the data seems to be saying by simple arithmetic and easy to draw pictures to summarize data. Confirmatory methods use ideas from probability theory in the attempt to answer specific questions. Probability is important in decision making because it provides a mechanism for measuring, expressing and analysing the uncertainties associated with future events.

The final results based on the analysis may be reported in the form of a table, a graph or set of percentages etc. The reported results must reflect the uncertainty through the use of probability statements and intervals of values.

### 6.3 Preparation of Data files for Analysis:

Data collected by means of questionnaires, interviews, diaries or any method mean very little until they are analysed and evaluated. We are constantly looking for similarities and differences, for groupings, patterns and items for particular significance. In questionnaires, it is helpful to identify question types and to workout ways in which responses can be analysed and presented. There will be seven types of questions - list, category, ranking, scale, quantity, grid and verbal.

**List Questions:** Suppose we want to find the qualification of the participants of a certain programme. We prepare a list of possible qualifications and the participants tick appropriate one from the supplied list. They may tick more than one box and so we will need to be ready to deal with multiple responses. Finally we prepare a frequency table indicating the number of participants having different qualifications. This information can be presented in a variety of ways. We may present this information in a simple table or a vertical bar chart.

**Category Questions:** Here the respondents will be asked to give one answer only. For example, age. In such cases we generally find the average— arithmetic mean, median, mode (central tendency). The selection of type of average depends on what you need to know and why. Normally we prepare the arithmetic mean. Even sometimes we may also be interested on dispersion or variation among the values of the responses— range, inter-quartile range, standard deviation and coefficient of variation. This type of data can be presented by a histogram or a pie chart.

**Grids:** Sometimes respondents have to answer two questions simultaneously. For example, age and experience of a teacher. Here there are two dimensions - age and

experience. The answers could be presented in a tabular form or it could be possible to produce a compound bar chart.

**Scales:** These are devices to discover strength of feeling or attitude. There are many different types of scales, some of which require complex construction and analysis. Thurstone and Guttman scales in particular require careful handling. The most straight-forward attitude scale is probably the Likert. Likert scales ask respondents to indicate strength of agreement or disagreement with a given statement or series of statements on a five or seven point range. Answers are then scored generally from 1(strongly disagree) to 5(strongly agree) and a measure of respondents feelings can be produced. A bar chart would also illustrate the range of response along with the tabular form.

**Verbal Questions:** Here the practice is to write all the responses on to separate sheets. This allows all items to be scanned in order to see whether there are any recurring themes. Some of the responses will probably provide useful questions to illustrate certain points in the report.

**Ranking:**Here respondents will be asked to indicate their order of preferences according the importance of the statements or items. This ranking of items makes a useful method of analysis. Items may be weighted in inverse order to the rankings. A composite judgement of the importance of the items could be determined by the weighted totals or averages for all the respondents. A bar diagram may be used to indicate the importances of the items.

### 6.4 Important Guidelines for Interpretation and Presentation:

1. Data must be recorded, analysed and interpreted.
2. Look for similarities, groupings and items of particular significance.
3. Try out different methods of presentation for responses to different question types, if any.
4. Prepare summary sheets.
5. Experiment with different ways of presenting findings – Table / Bar charts / Histogram / Other graphs.
6. If you need to discover the average of certain values, decide whether the mean, median or mode is the most suitable.
7. Study the dispersion, as the average may not give complete picture by calculating range or Inter-quartile range, standard deviation etc.
8. All the data require interpretation.
9. Do not attempt complex statistical techniques unless you have the expertise to cope.

Data are often recorded manually on data sheets. Unless the number of observations and variables are small the data must be analysed on a computer. The data will then go through three stages.

i) Editing or scrutiny of data is to be performed for the removal of inaccuracies, inadequacies of data before the coding.

ii) Coding: Coding involves the assigning of symbols (numbers) to each response of category. The purpose of symbols is to translate raw data into symbols, which may be counted and tabulated. The coding can be done by the observer or the interviewer at the time of data collection.

iii) Controlling Errors: Errors may occur at different levels. Recording, typing, transcription (copying), inversion (e.g.123.45 is typed as 123.54), repetition, deliberate errors are the common errors. Necessary precaution is required to control these errors.

A critical step in any study is the selection or naming of variables to be included. For example: an employee can be described among many variables, such as place of residence, education qualifications, experience and so on. The variables that are relevant to the problem under study must be chosen from the vast array of information available. All potentially relevant variables would be included in the study, since it is much easier to exclude unnecessary variables from analysis than to gather additional information.

**6.5 Recording the Data:** Once the variables have been selected, we must decide how they will be recorded. Recording is mainly depending on the nature of analysis. There are several methods of entering data into the computer. The entered data in this form is known as data file.

Computer plays very important role in statistical data analysis. The MS Exel , MS Access and SPSS (Statistical Package for Social Sciences) offer extensive data - handling capabilities and numerous statistical analysis routines that can analyse small to very large data statistics. The computer will assist in the summarization of the output to make inferences and predictions.

**6.6 Important Statistical Methods in Educational Research:**

**6.6.1 Classification:** After data collection, the fore most step is classification. Classification is the process of arranging the data into different groups or classes based on the similarities and dissimilarities. The principal objectives of classification are to:

i) facilitate comparison
ii) condense the mass of data into classes which can be readily apprehended
iii) indicate the most significant features of the data at a glance and
iv) make it possible the statistical application.

Broadly, classification is of four types:

i) qualitative classification
ii) quantitative classification
iii) geographical classification and
iv) chronological classification.

**Qualitative Classification:** In this, the basis for classification is quality or attribute such as sex, beauty, literacy, religion etc. Qualitative classification is studied under

i) simple classification — where only one attribute, say, presence or absence of attribute is studied.
ii) dichotomous classification — where two classes of attributes, say, sex into males and females are formed and again each category again into literates and illiterates.
iii) multiple classification — where two or more than two attributes are involved and divided into three or more than three categories are considered, sex, literacy and employment as indicated below:

| | **Literacy** | | | |
|---|---|---|---|---|
| **Sex** | Literates | | Illiterates | |
| | Employed | Unemployed | Employed | Unemployed |
| Male | | | | |
| Female | | | | |

***Tabulation.*** The main purposes of tabulation are to:

**Quantitative Classification:** This classification is based on the magnitude of the variable. For example, height, weight etc.

**Geographical Classification:** This classification is based on geographical area or location of the data concerned. For example, village, school, class etc.

**Chronological Classification:** Here the classification is based on the time period. For example, year, month, week, date etc.

**6.6.2 Presenting of Data:** The classified data can be presented normally in three ways: i) tabular form ii) diagrammatic form or iii) textual form.

**6.6.3 Tabulation:** The technique or process of presenting data, either qualitative or quantitative in the columns and rows is called

i) make the purpose of study clear
ii) express data in less space
iii) make comparison easy.

The main parts of a table are:

i) title of the table
ii) name of the rows (stubs) and columns (captions)
iii) body of the table
iv) footnote

When the data is classified on the basis of magnitude of the observations and presented in a tabular form, it is known as a frequency table.

**6.6.4 Frequency Table:** The statistical data arranged and classified into a number of groups in an orderly manner on the basis of magnitudes of the values, constitute a

frequency distribution, and a table presenting them is known as frequency table. For example, marks scored in mathematics by 20 students are presented in the form of the following frequency table.

| Marks | Number of students |
|---|---|
| 0 - 20 | 5 |
| 21 - 40 | 3 |
| 41 - 60 | 12 |
| 61 - 80 | 5 |

**6.6.5 Diagrammatic Representation:** While the numbers in the frequency table can be studied and compared, it is often useful to present results in a form that can be interpreted visually.

To simplify the complexity of statistical data and to draw the conclusions, data can be presented by means of various types of pictures or diagrams. The main types of diagrams are

a) Line diagrams
b) Bar charts- Absolute or Percentage and sub-divided
c) Pie diagrams
d) Histograms
e) Frequency Polygons
f) Frequency curves
g) Cumulative Frequency curves (Ogives)-less than, more than.



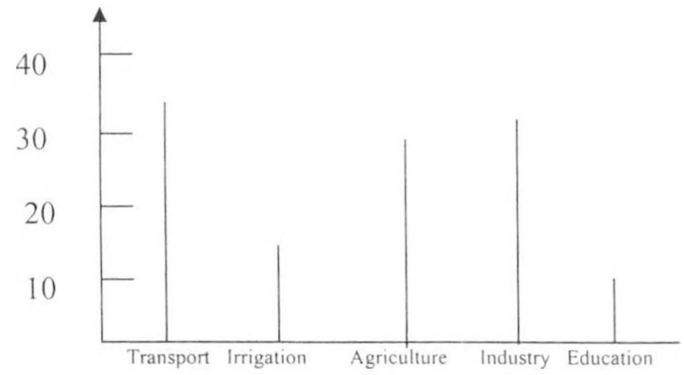
In the above-mentioned diagrams, the last four are related to the frequency distributions.

*Line Diagram:* Line diagram consists of a series of line segments to represent the magnitude of the item concerned as shown in the following example.

*Example:* Draw a line diagram to represent the following data obtained from plan outlay of the Central Government under the fourth five year plan.

| Heads of Expenditure | Rupees (crores) |
|---|---|
| **Transport and Communication** | 3237.26 |
| **Irrigation and Food Control** | 1086.57 |
| **Agriculture and Allied Sectors** | 2728.18 |
| **Industry and Minerals** | 3337.71 |
| **Education** | 822.66 |

The line diagram for the above data is drawn as below by considering the amount vertically and the Head of Expenditure items horizontally.



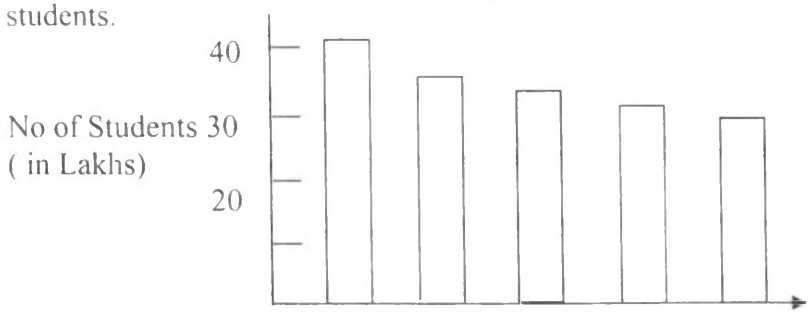

Head of Expenditure

*Bar Diagrams:* The term bar is used for a thick wide line. Bar diagram consists of a series of bars of equal width. The bars stand on a common base line with equal gap between one bar and another. The bars may be either horizontal or vertical. But vertical bars are preferred as they give a better look and facilitate comparison. The bars are constructed in such a way that their lengths are proportional to the magnitudes. Bar diagrams are called as one dimensional diagrams. Sometimes these are also known as column charts.

*Example:* The following data give the number of students in different classes. Represent the data by bar diagram.

| **Class** | **Number of Students** |
|---|---|
| I | 442839 |
| II | 354605 |
| III | 266718 |
| IV | 286045 |
| V | 263182 |

To represent the above data by means of bar diagram, five bars corresponding to those five classes will be drawn their lengths are proportional to the respective number of students.

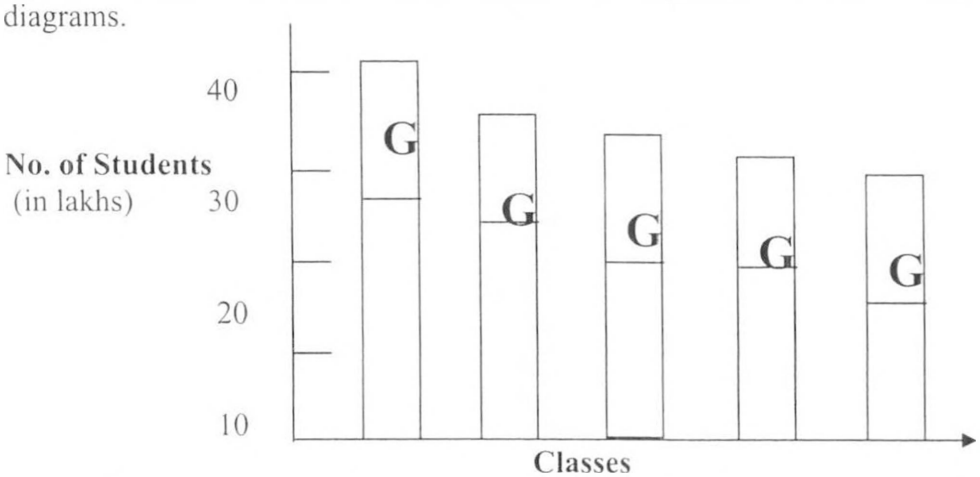


10

I II III IV V

Classes

The bar diagram used to represent only one variable (number of students in different classes) is known as simple bar diagram. In the above example suppose the number of students are given separately as boys and girls corresponding to each class. In this case, two bars one for boys and the other for girls of each class will be drawn to represent the data. Such bar diagrams are known as multiple bar diagrams. Sometimes, instead of drawing two bars separately for boys and girls, one bar will be drawn to represent the total students of each class and then that br will be divided into tow parts to indicate the boys and girls of that class as indicated below and such bar diagrams are known as subdivided bar diagrams. These will be used to represent two or more than two variables. In the bar diagram, if the lengths of bars are proportional to the percentage of figures instead of absolute values then those diagrams are known as percentage bar diagrams.




***Area Diagrams:*** Geometrical figures such as squares or circles are used to represent the data. They give a better comparison than bars and are convenient whenever bar charts are not feasible to draw as when one frequency is relatively large. The areas of these figures are proportional to the magnitude of the frequencies.

In the construction of squares, first of all square root of the various numbers of the data s circulated and then squares are drawn with the lengths of their sides in the same ratio as the square root of the original numbers given.

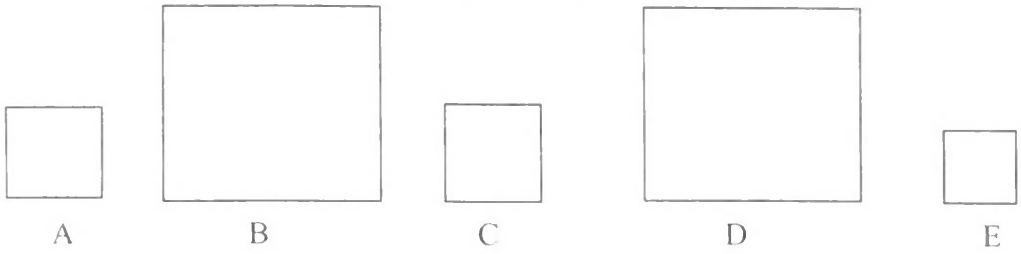
*Example:* Draw an area diagram for the following data.

The following table gives the number of operations performed by five doctors in a certain hospital during a period of one year.

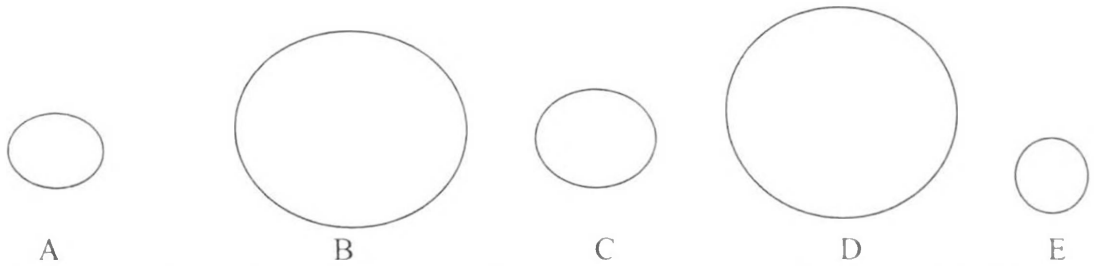
| Doctor | Number of Operations Performed |
|---|---|
| A | 16 |
| B | 64 |
| C | 36 |

| | |
|---|---|
| D | 81 |
| E | 16 |

Here five squares will be drawn corresponding to five doctors such that the area of each square is proportional to the respective number of operations as shown below:




Here, each side of the square is equal to the square root of the corresponding number of operations. To represent the above data, circles also can be drawn as follows:




*Pictograms:* Sometimes pictures are drawn to represent the relative values of items. The pictures drawn are of same size and the number of pictures are in proportion to the values of various items which have to be represented. For example, the number of thatched houses and pucca houses in a town can be represented by drawing the thatched house and pucca house figures. The number of thatched house figures should be proportional to the actual number of thatched houses in that town. This type of representation is more attractive and easy to understand.

*Pie Diagram:* Pie diagram is popularly used in many publications. It is used purely to impress the layman about the distribution of any quantity between different persons or trades, etc. The word pie is Latin origin indicating the shape of a cross section of an apple. Here, a circle is divided into different sectors so as to represent different quantities.

*Example:* The following table indicates the persons of different languages in R.I.E., Mysore. Represent these data by a Pie diagram.

| **Language** | **No. of Persons** |
|---|---|
| Tamil | 123 |
| Telugu | 146 |
| Kannada | 87 |
| Malayalam | 105 |
| Others | 24 |

| | |
|---|---|
| Total | 485 |

Consider the following steps to draw the required pie diagram.

Step 1: Draw a circle with a suitable radius (say 5 cm).

Step 2: Draw the circle into 5 different sectors in such a way that each sector represents one language group and the area of each sector is proportional to the number of persons of that group. This division of the circle into different sectors is possible after dividing the angle at the centre of the circle into five parts corresponding to the five language groups.

We know, the angle at the centre = $360^0$. Therefore, these $360^0$ will be divided into five parts.

The first part which is equal to $\left(\frac{123}{485} \times 360\right)$ = 91.3 because the ratio of the Tamil speakers to total speakers is $\frac{123}{485}$.

Similarly, corresponding to

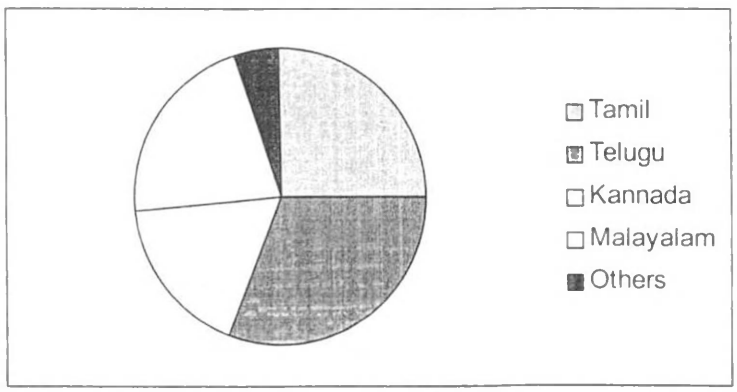
Telugu, the angle is $\left(\frac{146}{485} X 360\right)$ = 108.4

Kannada, the angle is $\left(\frac{87}{485} X 360\right)$ = 64.6

Malayalam, the angle is $\left(\frac{105}{485} X 360\right)$ = 77.9

Other languages, the angle is $\left(\frac{24}{485} X 360\right)$ = 17.8

Therefore, five different sectors making angles 91.3, 108.4, 64.6, 77.9, 17.8 at the centre will be used to represent the different language groups as shown below:




In addition to the above information, if the number of persons are given as males and females in each group, then that information also can be represented by a pie diagram by dividing each sector in the above diagram into two sectors so as to represent males and

females of that group. Similarly, data regarding more number of variables also can be represented with the help of a pie diagram in more legible way and in less space. If the number of variables are more, then the comparison is not so easy in a pie diagram s in the case of bar diagrams.

*Cartograms:* Sometimes, data will be represented by drawing suitable cartoons. For example, suppose we have the coal mines in India. These mines can be represented by indicating the locations on a map of India by a suitable symbol. For each mine one symbol will be used. By just looking at map, the location and the number of the coal mines will be spotted. Similarly, on the same map, other information like the number of major hydro-electric projects can also be represented. Places with the same rainfall can be indicated by drawing suitable lines passing through those places. Generally Cartograms are more informative.

As indicated earlier the other diagrammatic representations are related to frequency distributions.

*Histograms:* The data are represented as a series of rectangles. Class intervals are shown on the x-axis and the frequencies in terms of the lengths of rectangles are on the y-axis. There are as many rectangles as there are classes. The area of each rectangle represents the corresponding frequency of that class and the total area represents the total frequency. With a suitable scale, mark all the class intervals in the given frequency table on the x-axis, with the class intervals as bases, draw rectangles with areas proportional to the frequencies of the class intervals. For equal class intervals, the heights of the rectangles will be proportional to the frequencies.

*Example:* Draw histogram for the following data.

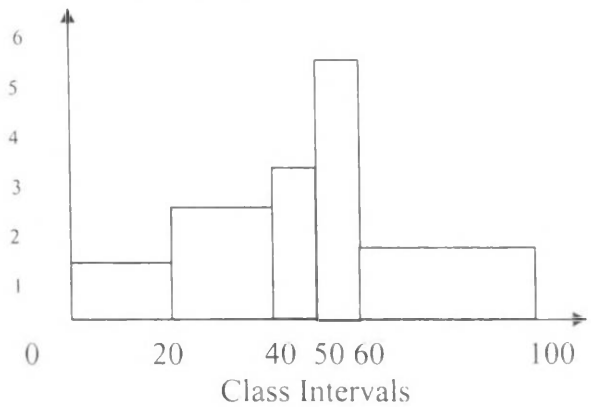
| Marks | Number of Students |
|---|---|
| 0-20 | 5 |
| 20-40 | 11 |
| 40-50 | 7 |
| 50-60 | 12 |
| 60-100 | 15 |

To draw the histogram corresponding to the give data, class intervals should be represented on x-axis with suitable scale (say 1" = 20 units). Then exact rectangles on these class intervals such that the areas of these rectangles are proportional to the corresponding frequencies of those classes. For these, the length of each rectangle should be calculated and to be represented on the y-axis. For example, the area of the first rectangle should be proportional to 5 units of frequency. But the base (class width) is one unit according to the scale considered. Therefore, the height of the rectangle can be considered as 1 unit on the y-axis so that the area of the rectangle as 1 square unit and this is proportional to the frequency of 5 units. The lengths of other rectangles are as follows:

Length of rectangle corresponding to the interval 20 – 40 = $\frac{11}{5} X \frac{1}{1}$ = 2.2 units because the ratio of the frequencies of first and this class is $\frac{11}{5}$ and the corresponding ratio of class width is $\frac{1}{1}$.

The lengths of other rectangles are $\frac{7}{5} \times \frac{1}{1/2} = 2.8$, $\frac{12}{5} \times \frac{1}{1/2} = 4.8$, $\frac{5}{8} \times \frac{1}{1/2} = 1.5$

The histogram corresponding to the given data is as below:




While drawing the histogram, class limits means the practical limits (or class boundaries) and therefore, generally continuous class intervals will be considered.

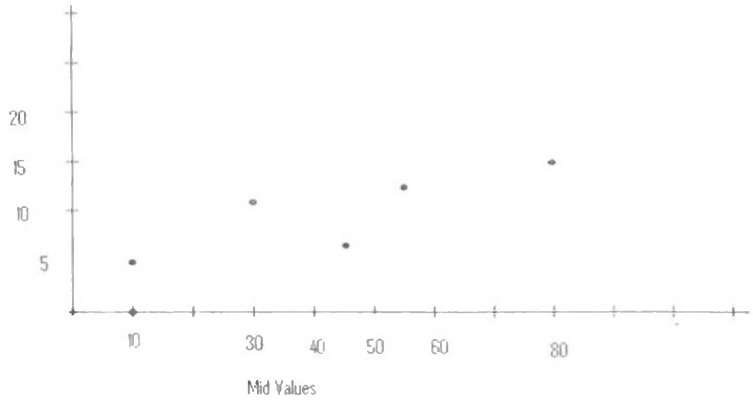
*Frequency Polygon and Frequency Curve:* For a grouped data, without class intervals, the frequency polygon is obtained by plotting points with abscissa as the variate values and the ordinates as the corresponding frequencies and joining the plotted points by means of straight lines. For grouped data with class intervals, the abscissa of the points are the mid-values of the class intervals. The frequency curve can be obtained by drawing a smooth free hand curve through the vertices of the frequency polygon.

*Example:* Draw the frequency polygon and the frequency curve of the data given in the previous example.

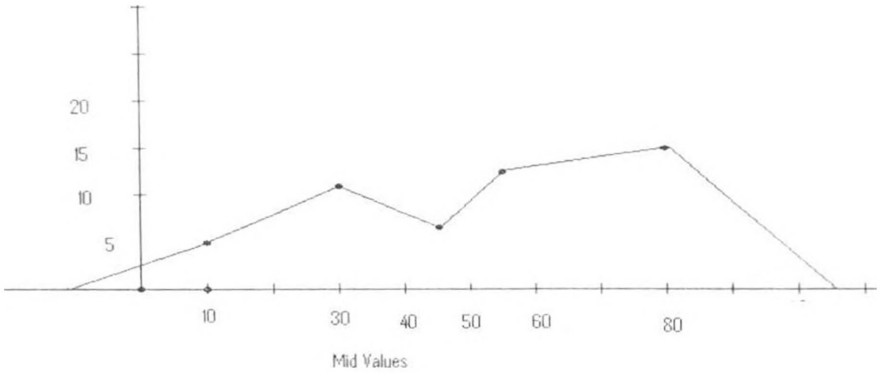
To draw the frequency polygon or frequency curve, it is necessary to calculate the mid-values of the class intervals. Therefore, the mid-values corresponding to the given data are as below:

| Mid Value | No. of Students frequency) |
|---|---|
| 10 | 5 |
| 30 | 11 |
| 45 | 7 |
| 55 | 12 |
| 80 | 15 |

The next step is to indicate these mid values on the x-axis with suitable scale and the corresponding frequencies to be considered as the ordinates. The five points are as indicated in the following figure.




Now, by joining those points by straight line segments, we get frequency polygon and by joining the points by free-hand curve, we get the frequency curve. But for all practical purposes, we join the end points with the points on the x-axis corresponding to the mid points of the preceding and succeeding class intervals to the first and last class intervals. Here, we assume the width of the preceding class is equal to that of the first class and the width of the succeeding class as equal to that of the last class. The frequencies of these data in the above example are indicated below:




***Cumulative Frequency Curves or Ogive:*** There are two types of cumulative frequencies- 'less than' cumulative frequencies and more than' cumulative frequencies. The less than cumulative frequency of any class is the number of observations having the values less than or equal to the upper limit of that class. Similarly, the more than cumulative frequency of any class is the number of observations having the values greater than the lower limit of that class. Firstly, form the cumulative frequency table and then plot the points with upper limits of classes as abscissa and the corresponding less than cumulative frequencies as ordinates. The points are to be joined by a free-hand smooth curve to get the less than cumulative frequency and therefore, the ogive is known as less than ogive. Similarly, corresponding to the lower limits and more than cumulative

frequencies, an ogive can be drawn and in that case, we call it as more than ogive or greater than ogive.

*Example:* Draw the less than and more than ogives for the following data

| Marks | Frequency |
|---|---|
| 0-10 | 2 |
| 10-20 | 18 |
| 20-30 | 19 |
| 30-40 | 35 |
| 40-50 | 40 |

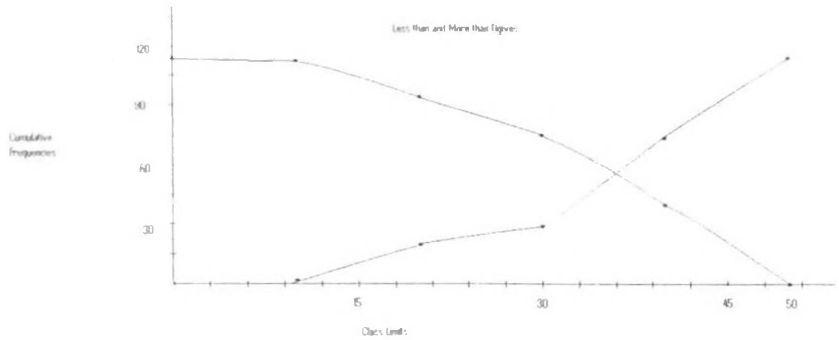
The first step before drawing n ogive is to calculate the cumulative frequencies. For the given data, the cumulative frequencies are as follows:

| Marks | Frequency | Less than Cumulative Frequency | More than Cumulative Frequency |
|---|---|---|---|
| 0-10 | 2 | 2 | 114 |
| 10-20 | 18 | 20 | 112 |
| 20-30 | 19 | 39 | 94 |
| 30-40 | 35 | 74 | 75 |
| 40-50 | 40 | 114 | 40 |

Now the different points to draw less than ogive are (10,2), (20,20), (30,29), (40,74), (50,114).

Similarly, the different points to draw more than ogive are (0,114), (10,112), (20,94), (30, 75), (40,40).

By plotting these points on a graph sheet and joining in order, we obtain the respective ogives as shown below:




From the ogives, it is easy to find out the number of observations having the value greater than any given value. Also, for any given number of observations, it is possible to obtain the value so that the number of observations greater than that value or less than that value. In some particular situations, ogives are very helpful to get the required information directly.

Frequency polygon, frequency curve and ogive can also be drawn by considering the percentage cumulative frequencies and percentage cumulative frequencies instead of the absolute frequencies. In those cases, we call them as percentage frequency polygon, percentage frequency curve and percentage cumulative curves (or ogives) respectively. After representing the data in the suitable forms, the next stage are comparing, analyzing and interpreting the data.

*Comparing, analyzing and interpreting data:* After the statistical data are collected and presented in a systematic manner by means of tables and charts, it must be summarized and condensed for the sake of comparison of the present data with the similar type of data. One of the ways of considering the data is finding a measure of central tendency or average.

**6.6.6. Measures of Central Tendency:** In order to reduce the complexity of data and to make them comparable, it is essential that the various phenomena which are being compared are reduced to one figure each. Such figures are called 'measures of central tendency' or averages. Average should give the gist and concise picture of the whole data. Therefore, an ideal (or good)average should represent the data for which it is considered.

There are many ways to consider the average for any given data.

*Arithmetic Mean:* Let the different values in the given raw (ungrouped) data be $x_1$, $x_2$, ..... $x_n$. Then the arithmetic mean (A.M) denoted by $\bar{x}$ is defined as the ratio of sum of the values and the number of values.

$$\bar{x} = \frac{\sum_{1}^{n} x_i}{n}.$$

In case of a frequency distribution with k class intervals, the A.M. is given by

$$\bar{x} = \frac{\sum_{1}^{k} f_i x_i}{\sum_{1}^{k} f_i}$$

where $x_i$ is the mid value of the i th class, $f_i$ is frequency of the i th class and k is number of classes.

Arithmetic mean is effected much by the end values. Also arithmetic mean cannot be calculated when the grouped data is having an open class interval and when the data are ordinal. Therefore, in these cases, another average 'Median' is used.

*Median:* Median is the value which exceeds and exceeded by the same number of observations or values.

If the n values in an ungrouped data are arranged in ascending (or descending) order of magnitude, the median is the middle value $\left(\frac{n+1}{2} th\ value\right)$ when n is odd; it is taken as

the arithmetic mean of the two middle values if n is even though in this case, any value between the two middle values will satisfy the definition.

Median in case of frequency data with class intervals is given by

$$\text{Median} = l + \frac{\frac{nx\,h}{2} - c}{f} \times h$$

where l = lower limit of the median class

f = frequency of the median class

h = width of the median class

N = total frequency

c = the largest cumulative frequency which is less than $\frac{N}{2}$.

Sometimes median is not changed even some observations are changed and also, it is not possible to calculate in case of frequency distribution where the median class coincide with the open class if any. In such cases, we consider another average known as 'mode'.

***Mode:*** Mode is the value which occurs more frequently than others. In ungrouped data, mode can be obtained by mere inspection. Even in case of grouped data without class-intervals, mode can be calculated easily.

If there are two or more values occurring with the same maximum number of items, then all those values are the modes of that data. The distribution having one mode is known as uni-modal distribution, having two modes is bi-modal distribution and having more than two modes is known as multi-modal distribution.

In case of grouped data with class intervals, mode is given by the following formula

$$\text{Mode} = l + \frac{f - f_1}{2f - f_1 - f_2} \times h$$

Where l= the lower limit of the median class ( model class means the class having the maximum frequency)

f = frequency of the modal class (or the maximum frequency)

$f_1$ = frequency of the class preceding the modal class

$f_2$ = frequency of the class succeeding the modal class

h = width of the modal class.

The above formula can be used only when the width of the modal class is not greater than the width of any other class. This is important, because as otherwise, the frequency of a class can be increased by increasing the width and any class may become a modal class.

Sometimes, the mode can be obtained in terms of the values of arithmetic mean and median by using the following relation;

Mode = 3 Median – 2 Arithmetic Mean.

It is not possible to obtain mode in the situations where the modal class coincides with an open end class or the preceding and succeeding classes and the modal classes are not having the same width to decide the maximum frequency correctly.

*Geometric Mean:* Whenever we want to measure the rate of change, another type of average other than arithmetic mean, median or mode is used. One such average is 'Geometric Mean' which is abbreviated as G.M. and indicated by G. This is defined only when all the observations are positive.

The geometric mean of n values is defined as the nth root of their product. That is the geometric mean of $x_1, x_2 \ldots, x_n$ is equal to $(x_1\ x_2 \ldots, x_n)^{1/n}$

In a frequency distribution, if $x_1, x_2 \ldots, x_k$ are the values and $f_1, f_2 \ldots, f_k$ are the corresponding frequencies, then the geometric mean is given by

$$G = \left(x_1^{f_1} x_2^{f_2} \ldots\ldots x_k^{f_k}\right)^{\frac{1}{N}} \quad \text{where } N = \sum_1^k f_i = \text{Total frequency.}$$

Similarly, in case of a frequency distribution with class intervals, G.M. is given by

$$G = \left(x_1^{f_1} x_2^{f_2} \ldots\ldots x_k^{f_k}\right)^{\frac{1}{\sum_1^k f_i}}$$

where $x_1\ x_2 \ldots, x_n$ are the mid values of the class intervals and $f_1, f_2 \ldots, f_k$ are the respective frequencies of those classes and G is the geometric mean. For all computational purposes, we use the following formula

$$\log G = \frac{\sum_1^k f_i \log x_i}{\sum_1^k f_i}$$

Geometric mean is very useful in the construction of index numbers and in calculation of population growth and the rate of interest.

***Harmonic Mean:*** In some of the situation like calculation of average speed when the distances and the corresponding times are given, we use another type of average known as 'Harmonic Mean' is used. Harmonic mean is denoted by H.

The harmonic mean of n values is defined as the reciprocal of the arithmetic mean of the reciprocals of these values.

Suppose the values are $x_1, x_2 \ldots, x_k$ and non-zero. Then the harmonic mean, H is given by $\dfrac{n}{\sum_1^n \frac{1}{x_i}}$

For a frequency distribution, $H = \frac{N}{\sum_{1}^{k} \frac{f_i}{x_i}}$ Where N is the total frequency, $x_i$ and $f_i$ are mid value and the frequency of the i th class respectively.

***Weighted Average:*** In the calculation of simple average, each item of the series is considered equally important. But there may be cases where all the items may not have equal importance, and some of them may be comparatively more important than others. An average which takes into account the relative importance of various items in its calculation is called 'weighted average'. The figures which indicate the relative importance of various items are called 'weights'.

There are three types of weighted averages which we use generally. They are

i)weighted arithmetic mean
ii) weighted geometric mean and
iii)weighted harmonic mean

The weighted arithmetic mean $\bar{x} = \frac{\sum_{1}^{n} w_i x_i}{\sum_{1}^{n} w_i}$ where $\bar{x}$ denotes the weighted arithmetic mean and $w_i$ is weight of the i th item $x_i$

Similarly, the weighted geometric mean and weighted harmonic mean are given by

$$\text{Log } G = \frac{\sum_{1}^{n} w_i \log x_i}{\sum_{1}^{n} w_i}$$

and $$H = \frac{\sum_{1}^{n} w_i}{\sum_{1}^{n} \frac{w_i}{x_i}}$$

***Selection of an average:*** No single average is suitable for all practical purposes. Each one of the averages has its own merits and demerits and thus its own particular field of importance and use. A judicious selection of the average depending on the nature of the data and the purpose of the enquiry is necessary for sound statistical analysis. Since the arithmetic mean possesses the good properties for an ideal average, it may be regarded as the best of all the averages.

**6.6.7 Measures of Dispersion:** When we compare two similar types of data the measures of central tendency may be the same in both cases, but still we find some differences in the observations of those. In one set, observations may be varying much from one another and in the second set the variation may be less. This type of difference

is known as the 'Dispersion'. Here, dispersion means 'Scatteredness'. These are different measures of dispersion which are given as below.

***Range:*** Range is the difference between the largest and the smallest observations or values. It is obvious that the range is depending upon the largest and smallest values only. Therefore, this will not give any idea regarding the variation of other values. Another measure know as 'Quartile Deviation' will be preferred.

***Quartile Deviation:*** Certain values can divide the given set of values into equal number of groups and these values which divide the given observations to form equal parts are known as 'quantiles'. Three values can divide the given set of observations into four equal parts. These three values are known as 'quartiles'. As already known, median will divide the observations into two equal parts. Also, we know the method of calculating median in case of any given data. Similarly, the three quartiles (denoted by $Q_1$, $Q_2$, $Q_3$) can be calculated.

The first quartile, $Q_1$ is the value which divides the given observed values into two parts such that one-fourth of the observations to have the values less than $Q_1$ and the remaining three-fourths observations to have values greater than $Q_1$. The third quartile, $Q_3$ is the value which divides the given observed values into two parts such that three-fourths of the observations to have the values less than $Q_3$ and the remaining one-fourth observations to have values greater than $Q_3$.

Quartile deviation is half of the difference between the third and first quartiles. Symbolically, this is equal to $\frac{Q_3 - Q_1}{2}$. It is also known as semi-interquartile range, $Q_3$ - $Q_1$ is the range of the quartiles. It is obvious that the quartile deviation may be same even some observations are changed, as the measure is based on the first and third quartiles. Another measure of dispersion which changes when some observations are changed is required and one such measure is the 'standard deviation'.

***Standard Deviation:*** The mean of the squares of the deviations of the observations about their arithmetic mean is known as 'Variance'. The positive square root of the variance is known as 'standard deviation'. Standard deviation is denoted by $\sigma$ and variances by $\sigma^2$.

If the observed values are $x_1, x_2, \ldots, x_n$ and their arithmetic mean is $\bar{x}$, then the standard deviation, $\sigma$ is given by $\sqrt{\frac{\sum_1^n (x_i - \bar{x})^2}{n}}$ or $\sqrt{\frac{\sum_1^n x_i^2}{n} - x^{-2}}$

The standard deviation in case of grouped data is given by the formula

$$\sigma = \sqrt{\frac{\sum_{1}^{k} f_i\left(x_i - \bar{x}\right)^2}{\sum_{1}^{k} f_i}} \text{ or } \sqrt{\frac{\sum_{1}^{k} f_i x_i^2}{\sum_{1}^{k} f_i}} - x^{-2}$$

where k is the number of class intervals, $x_i$ is the mid value of the i th class, $f_i$ is the frequency of the i th class and $\bar{x}$ is the arithmetic mean of the distribution.

The mean and hence the standard deviation will be difficult to calculate if the values are fractions or the frequencies are large. In such cases, one modified and short-cut method will be used as described below:

First change the given values into new values by using the following transformation. $u_i = \frac{x_i - A}{h}$ where $x_i$' s are the given values, $u_i$'s are the new values, A and h are some constants. Calculate the standard deviation of these new values. The standard deviation of the given values is equal to h times of the standard deviation of the new values. That is $\sigma_x = h\sigma_u$ where $\sigma_x$ is the standard deviation of given values, $\sigma_u$ is the standard deviation of new values. Particularly this method is very useful in grouped data with equal class intervals.

***Coefficient of Dispersion:*** Whenever, we have to compare the variations in two data, we first calculate the standard deviation of each data and then compare these standard deviations. If the standard deviation of the first data is more than the other, we say the variation is more in the first data when compared to the other. But this comparison is valid when the units of measurement in both the data are the same. Unless the units of measurement are the same. We cannot draw the conclusions regarding the variability on the basis of the magnitude of the measure of dispersion. Therefore for direct valid comparisons, it is necessary to make the measures of dispersion independent of the units of measurements. For this reason, we modify the different measures of dispersion so s to give pure numbers i.e., numbers free from the units used. These modified measures are known as 'coefficients of dispersion'. The different coefficients of dispersion based on different measures are listed below:

i) $\frac{Maximum\,value - Minimum\,value}{Maximum\,value + Minimum\,value}$. This is corresponding to the range.

ii) $\frac{Q_3 - Q_1}{Q_3 + Q_1}$. This is based on quartile deviation.

iii) $\frac{Standard\ Deviation}{Mean} = \frac{\sigma}{\bar{x}}$. This is based on standard deviation.

Standard deviation is generally considered as the best measure of dispersion. Therefore coefficient of dispersion based on standard deviation is considered as the best

value to use. For better comparison, we magnify this fractional value multiplying by 100. This final value is known as the coefficient of variation.

$$\text{The coefficient of variation} = \frac{S\tan dard\ Deviation}{Mean} \times 100 = \frac{\sigma}{\bar{x}} \times 100$$

This value can be used for the comparison of variation in any given data, even though the units of measurements are different.

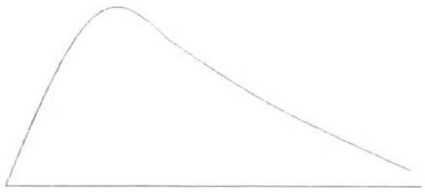
***Skewness:*** The distributions (or data) under comparison are not identical even when they have equal measures of central tendency and equal measures of dispersion. For example, in case of two distributions, the arithmetic means and the standard deviations may be the same but in one distribution the number of observations greater than the arithmetic mean may be smaller than the number of observations greater than the arithmetic mean of the other distribution. Such differences will be measured as the amount of asymmetry or skewness.

In a symmetrical distribution, the mean and median coincide with the mode which is the value corresponding to the central ordinate. The frequency curve of a symmetric distribution is generally a bell shaped curve as shown below.

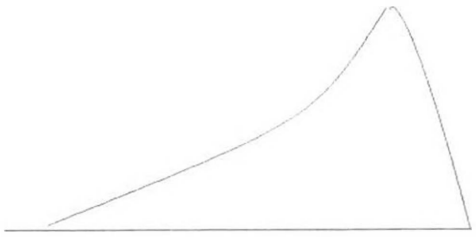
**Chart**



If we have more items to the right of highest ordinate of the frequency, the curve is said to be positively skewed and then the nature of the frequency curve is as given below.



Similarly, the curve is said to be negatively skewed if more items are found to the left of the highest ordinate. The type of curve is as shown below.



In positively skewed distribution, the arithmetic mean lies the right of mode and in case of negatively skewed distribution means lies towards left of mode. Always median lies between mean and median.

Literally skewness means lack of symmetry. We study skewness to have an idea about the symmetry of data – symmetric or skew. There are various measures for skewness. We consider (Mean-Mode) as one of the measures of skewness. But sometimes mode is not unique and therefore, we replace (Mean-Mode) by 3(Mean-Median). For comparison, we make this measure independent on units of measurement by dividing with standard deviation. This measure is known as the Karl Pearson's coefficient of skewness and is denoted by $S_k$. Therefore, $S_k = \frac{3(Mean - Median)}{Standard\ Deviation} = \frac{3(M - M_d)}{\sigma}$ where M is the mean, $M_d$ is the median and $\sigma$ is standard deviation.

Similarly, Bowley's coefficient of Skewness is given by $\frac{Q_3 + Q_1 - 2M_d}{Q_3 - Q_1}$ where $M_d$ is the median.

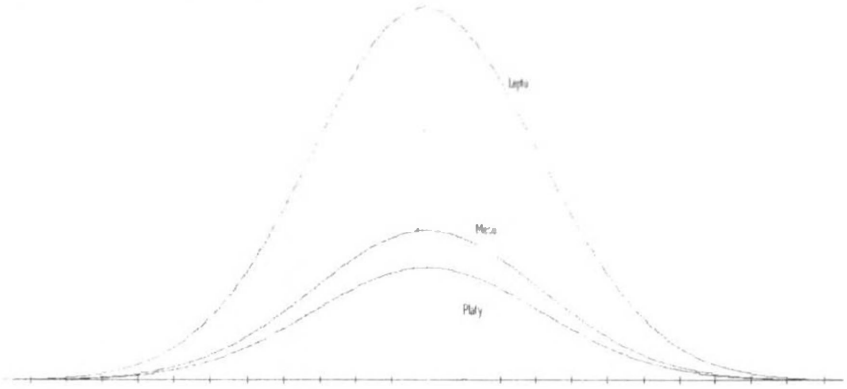
Another coefficient of skewness based on the moments is given by $\frac{\mu_3}{\mu_2^{\frac{3}{2}}}$ where

$\mu_3 = \frac{\sum_1^n (x_i - \bar{x})^3}{n}$, $\mu_2 = \frac{\sum_1^n (x_i - \bar{x})^2}{n}$ and are known as third and second.

The sign of any coefficient of skewness indicate the type of skewness. That is, if coefficient of skewness is zero then the distribution is symmetric, if the coefficient is positive then the distribution is positively skewed and if the coefficient is negative, the distribution is negatively skewed. The magnitude of the coefficient of skewness indicates the amount of skewness. For comparison, the coefficients of skewness of each distribution will be calculated and compared.

***Kurtosis:*** When two distributions are compared, suppose they have equal measures of central tendency, they have equal measures of dispersion and they are equal in skewness. Even then, in one data most of the items (or values) may be concentrated at some central

value than the other data. This type of more concentration gives high peak in the frequency curve s shown below.



On the other hand, if the concentration of items is less, then the peakedness of the frequency curve is also less and the tails on either side are longer. The degree of peakedness of the curve, that is, the concentration of values in the neighbourhood of mean is measured by kurtosis.

If the peak of the frequency curve of a given data is higher than the peak of the frequency curve of a normal (standard) distribution, then that distribution is known as 'leptokurtic' and if the peak is lower than that of normal it is known as 'platykurtic'. This normal distribution is said to be 'mesokurtic'.

One of the important measures of Kurtosis is given by $\frac{\mu_4}{\mu_2^2}$ -3 where $\mu_4 = \frac{\sum_1^n (x_i - \bar{x})^4}{n}$

**6.6.8 Correlation Analysis:**

So far we considered the data classified according to the values of a single variable (or on single character) and we saw how the observed values could be grouped into a frequency distribution, whose characteristics could be described by certain constants like the measure of central tendency, measure of dispersion, measure of skewness and measure of kurtosis. Now we have the case of two variables, in which each observation will exhibit two values, one for each of the variables under consideration. For example, the heights and weights of n students. Here two values – one for height and the other for weight, will be given corresponding to each student. Such values will be represented as

$(x_1,y_1), (x_2,y_2), \ldots\ldots, (x_n, y_n)$

This kind of data is called as 'bivariate data'. In such bivariate data situations, we have to measure the variation of one variable in relation to the variation in the other. Such measure of variation is known as the measure of covariation. This measure of covariation gives the idea of dependency of one data on the other. For example, the study

of relationship between the marks in mathematics and the marks in language of the same set of students requires knowledge of covariation.

A measure of covariation is known as covariance or correlation.

With a change in the values of one of the two variables the values of the other variable also change. Such a relationship between two variables is known as 'mutual dependence' and the variables are said to be 'mutually dependent'. The relationship between two variables need not always be the result of their mutual inter-dependence. Changes in the values of one variable may be the cause of changes in the values of the other variable; thus, there may be cause and effect relationship between the two variables. Sometimes, two variables may not have any type of apparent relationship between themselves. Yet we can find that when the value of one variable increases, the value of the other variable also changes. This may be due to a third factor which causes both variables to change simultaneously.

If with an increase in the value of one variable, the value of the other variable increases and if with a decrease in the value of one variable the value of the other variable also decreases, then we say the correlation between those two variables is positive. For example, the velocity of a projectile and its range. If with an increase in the value of one variable, the value of the other variable decreases and if with decrease in the value of one variable, the value of the other variable increase, then we say the correlation between those two variables is negative. For example, the length of the pendulum and its period of oscillation. Whenever, the changes in one variable bring the proportional changes in the other, then we say the correlation between those variables is linear correlation, otherwise, the correlation is non-linear or curve-linear correlation. If only two variables are considered and if correlation exists between these variables, the correlation is said to be 'simple'. If we consider more than two variables, then the association of one variable with the rest of the other variables is called 'multiple correlation'. For example, demand of a commodity depends upon the price, income of the consumer, taste of the consumer, etc. If we consider the relationship of two variables among the several variables assuming the values of other related variables to remain constant, then the correlation is said to be 'partial'.

First we consider our discussion to study of simple, linear correlation. Sometimes, there may not be any relationship between two variables and sometimes one variable may be completely dependent on the other. We say in the first case there is no correlation and in the second, there is perfect correlation.

Correlation can be studied by any one of the following methods.

1. Scatter diagram
2. Karl Pearson's coefficient of correlation and
3. Rank correlation coefficient.

***Scatter Diagram:*** The dependence or the relation between two sets of values can be seen by plotting the points corresponding to all the pairs of values, say (x,y) on a graph by taking the values of one variable x on the x-axis and the values of y on y-axis. The graph

of these points is known as the 'scatter diagram' as it reveals the scatteredness of the values. This diagram is also known as 'dot diagram'. The different possible scatter diagrams are given below:

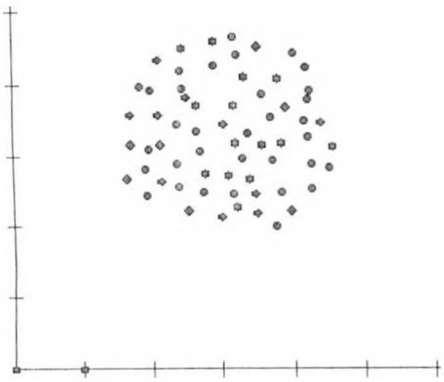


Fig1

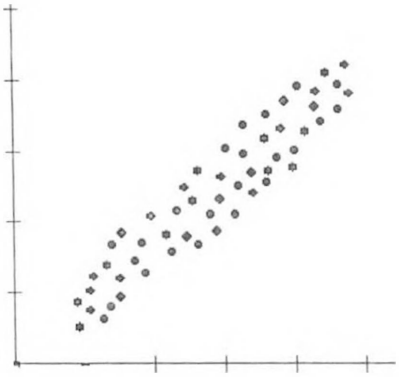


Fig2

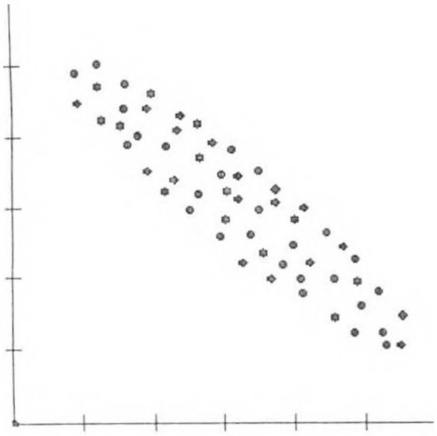


Fig3

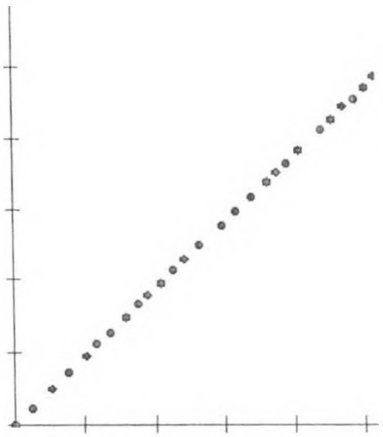


Fig4

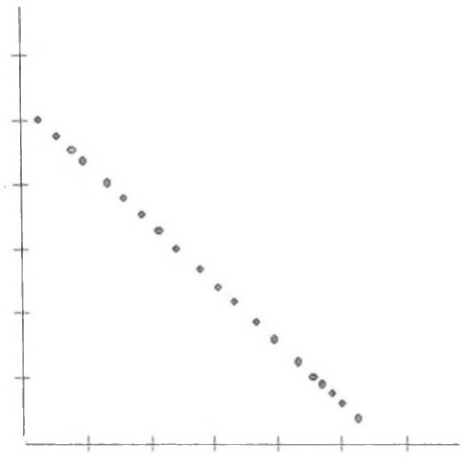


Fig 5

If the plotted points are scattered as in Fig.1, then we say there is no correlation. If the dots are lying around a straight line running from left to right in the upward direction as in Fig. 2 then the correlation between the two given variables is said to be positive. If the dots are lying around a straight line running from left to right in the downward direction as in Fig. 3 then the correlation is negative. If the dots in scatter diagram from a straight line running from left to right in the upward direction as in Fig. 4 then the correlation is 'perfect' ad 'positive'. If the dots form a straight line running from left to right in the downward direction as in Fig. 5, then the correlation is 'perfect' and 'negative'. Study of correlation from scatter diagram is the simplest and the quickest method. This diagram reveals the nature of correlation but not the amount of correlation. A better measure is the Karl Pearson's coefficient of correlation.

***Karl Pearson's Correlation Coefficient:***

If we represent the values of the bivariate data as $(x_1,y_1)$, $(x_2,y_2)$,....., $(x_n , y_n)$ then the Karl Pearson's coefficient of correlation, r is given by

$$r = \frac{\frac{1}{n}\sum_{1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sigma_x\sigma_y}$$

where $\sigma_x$ and $\sigma_y$ are the standard deviations of x and y values respectively.

The above formula can also be written as

$$r = \frac{\frac{1}{n}\sum_{1}^{n}x_iy_i-\bar{x}\bar{y}}{\sigma_x\sigma_y}$$

Here r can take all the values between –1 and 1, including these values. If r takes the value 0 we say that there is no correlation and the two variables are independent. If r =1, correlation is perfect positive correlation; if r =-1, the correlation is perfect negative correlation.

The Karl Pearson's Correlation coefficient is also known as Product moment correlation coefficient.

***Rank correlation Coefficient:***

Sometimes, the exact values of the observations may not be available and in such cases the product moment correlation coefficient cannot be calculated. But if we know the ranks (relative positions) of the observations, the correlation coefficient can be obtained in some other way. The correlation coefficient obtained on the basis of the ranks of observations is known as rank correlation coefficient.

Let $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ be the ranks of n individuals on two characteristics A and B respectively. Then the rank correlation coefficient, $\rho$ is given by

$$\rho = 1 - \frac{6\sum_{1}^{n} d_i^2}{n(n^2-1)}$$

Where $d = |x_i - y_i|$

This $\rho$ can take any value between $-1$ and $1$ including those values like the product moment correlation coefficient.

Sometimes, we use rank correlation coefficient instead of product moment correlation coefficient because of its easiness in calculation. In such cases, we first rank the x values and y values separately by giving rank '1' to the lowest value, rank '2' to the next lowest and so on. Sometimes two or more values of x may have the same value. In such cases, we give the average rank (average of the ranks that those observations are supposed to receive when they are different) to all such values. These ranks are known as common ranks. Because of these common ranks we make some adjustments in the rank correlation coefficient by adding the factor $\frac{m(m-1)}{12}$ to $d^2$ where m is the number of times a value is repeated. This correction factor is to be added for each repeated value.

## 6.6.9 Regression

Correlation analysis deals with the measure of amount of dependency among the variables. To study the amount of relationship between the variables one should use the regression analysis. In case there exists association or relationship between two variables x and y, the dots of the scatter diagram will be more or less concentrated round a curve which may be called the curve of regression and the relationship is said to be expressed by means of curvilinear regression. In the particular case when the curve is a straight line, it is called the line of regression and the regression is said to be linear. More precisely, the line of regression is the straight line which gives the best fit in the least square sense.

If the straight line is so chosen that the sum of squares of deviations parallel to the y axis is minimized, it is called the line of regression of y on x and it gives the best estimates y for any given value of x. If, on the other hand, the sum of squares of the deviations parallel to the x-axis is minimized, the resulting straight line is the line of regression of x on y for any given values of y. The equations of these two regression lines are

$$y - \bar{y} = r\frac{\sigma_x}{\sigma_y}(x - \bar{x})$$

And $$x - \bar{x} = r\frac{\sigma_x}{\sigma_y}(y - \bar{y})$$

where r is the correlation coefficient of x and y; $\bar{x}, \bar{y}$ are the arithmetic means of x and y and $\sigma_x, \sigma_y$ are the standard deviations of x and y respectively.

The coefficients $r\frac{\sigma_y}{\sigma_x}$ and $r\frac{\sigma_x}{\sigma_y}$ are called the regression coefficients of y on x and x on y respectively and give estimates of changes of y on x corresponding unit changes of x on y respectively. They are usually denoted by $b_{yx}$ and $b_{xy}$ respectively. If r=1 or –1, both the regression lines coincide. The departure of the value of $r^2$ from unity is a measure of departure of the relationship between two variates from linearity.

After deriving the regression equation, we want to know how powerful an explanation (or prediction) our regression equation provides. More technically, how well does the regression equation account for variations in the dependent variable? A preliminary judgment comes from visual inspection of the scatter-plot. The closer the regression line to the points, the better the equation "fits" the data. Which such "eyeballing" is an essential first step in determining the "goodness of fit" of an equation, we obviously need a more powerful measure, which the coefficient of determination ($R^2$) gives us. Here

$$R^2 = \frac{\sum_1^n (\hat{y}_i - \bar{y})^2}{\sum_1^n (y_i - \bar{y})^2}$$

Where $y_i$ is the observed value of the dependent variable corresponding to $x_i$ $\hat{y}_i$ is the estimated value of yi corresponding to $x_i$ and $\bar{y}$ is the average value of y.

The coefficient of determination, $R^2$, indicates the explanatory power of the regression equation. It records the proportion of variation in the dependent variable "explained" or "accounted for" by the independent variable. The possible values of this measure range from +1 to 0. At the one extreme, where R2 =1, the independent variable completely accounts for variation in the dependent variable. All observations fall on the regression line; so knowing x enables the prediction of y without error. At the other extreme, when R2 =0, the independent variable accounts for no variation in the dependent variable. The knowledge of x is no help in predicting y, for the two variables

are totally independent of each other. Generally, $R^2$ falls between these two extremes. Then, the closer $R^2$ is 1, the better the fit of the regression line to the points, and the more variation in y is explained by x. In regression analysis, we are virtually always pleased when the $R^2$ is high, because it indicates we are accounting for a large portion of the variation in the phenomenon under study. Further, a very high $R^2$ (say about 0.9) is essential if our predictions are to be accurate. In practice, it is difficult to attain an $R^2$ of this magnitude. However, a sizable $R^2$ does not necessarily mean we have a casual explanation for the dependent variable, instead, we may merely have provided a statistical explanation. The relationship between the coefficient of determination, $R^2$ and the estimate of the correlation coefficient, r, is $R^2 = r^2$. This equality suggests a possible problem with r, which is a commonly used measure of strength of association. That is, r can inflate the importance of the relationship between x and y. By relying on rather than $R^2$, the impact of x on y can be made to seem much greater than it is. Hence, to assess the strength of the relationship between the independent variable and dependent variable, the $R^2$ is the preferred measure.

Often it is necessary to find correlation between three or more variables. For example, the total marks are affected by marks in Physics, and the marks in Physics are affected by Marks in Mathematics. Whenever we are interested in the combined influence of a group of variates upon a variate not included the group, our study is that of multiple regression and multiple correlation. In such cases, to study the relationship between two variates, we have too methods. Firstly, we may consider only those members of the observed data in which the other members have specified values. Secondly, we may eliminate mathematically the effect of the other variates on the two variates under study.

The correlation between two variates when the linear effect of the other variates in them has been eliminated from both is called partial correlation. Suppose we are given n sets of corresponding values of three variates $x_1, x_2$ *and* $x_3$. Let these variates be measured from their respective means and the quantities obtained be $X_1, X_2$ *and* $X_3$. The regression equation of $X_1$ say, on $X_2$ *and* $X_3$ may be written in the form

$$X_1 = a + b_{12.3}X_2 + b_{13.2}X_3$$

Where the constants a and b are obtained based on least squared fit concept. Here

$$b_{12.3} = \begin{vmatrix} r_{12}\sigma 1 & r_{13}\sigma_3 \\ r_{12}\sigma_1 & \sigma_3 \end{vmatrix} \div \begin{vmatrix} \sigma_2 & r_{23}\sigma_2 \\ r_{23}\sigma_2 & \sigma_3 \end{vmatrix}$$

$$b_{13.2} = \begin{vmatrix} r_{13}\sigma 1 & r_{12}\sigma_2 \\ r_{13}\sigma_1 & \sigma_2 \end{vmatrix} \div \begin{vmatrix} \sigma_3 & r_{23}\sigma_2 \\ r_{23}\sigma_2 & \sigma_2 \end{vmatrix}$$

where $r_{12}$ is the correlation coefficient of $X_1$ *and* $X_2$. Similarly $r_{13}, r_{23}$ . These $b_{12.3}$ $b_{13.2}$ are called the Partial regression coefficients. The quantity $X_{1.23} = X_1 - b_{12.3}X_2 - b_{13.2}X_3$ is called the residual of second order.

***Multiple Correlations***:

Consider the regression equation of $X_1$ *on* $X_2$ *and* $X_3$ viz.,

$$X_1 = b_{12.3}X_2 + b_{13.2}X_3 .$$

Now the correlation between $X_1$ (observed value of the variate) and $X_{1.23}$ (expected value of the variate) is given by

$$R_{1.23} = \frac{\sum X_1 X_2}{\sqrt{(\sum X_1^2)}\sqrt{(\sum X_2^2)}}$$

$$\text{or} \quad 1\text{-}\ R_{1.23}^2 = \frac{\sigma_{1.23}^2}{\sigma_1^2}$$

where $\sigma_{1.23}^2$ is the variance of residual $X_{1.23}$.

If $R_{1.23} = 1$, then $X_1$ is a linear function of $X_2$ *and* $X_3$.

The partial correlation coefficients are as given below.

$$r_{12.3} = \frac{r_{12} - r_{12}r_{23}}{\sqrt{\{(1-r_{13}^2)(1-r_{23}^2)\}}}$$

$$r_{13.2} = \frac{r_{13} - r_{13}r_{23}}{\sqrt{\{(1-r_{12}^2)(1-r_{23}^2)\}}}$$

The concept of partial and multiple correlation and regression can be extended to more than three variables similarly.

## 6.7 Statistical Inference

Statistical inference is the problem of making conclusions about the population based on sample information. There are two aspects of statistical inference- one is Estimation and the other is Testing of Hypothesis.

*Estimation:* Here based on the sample information we have to find population value. A point estimate is a single sample value used as an estimate of the population value, parameter. Because of chance sampling variation, point estimates are in error – by an unknown amount. Interval estimates, on the other hand, incorporate sampling variation into the estimate and give a range within which the population value is estimated lie. Interval estimates are provided with a specified level of confidence, equal to (1- $\alpha$ )100 percent (usually 95% or 99%). A 95% confidence interval is constructed according to rule, $\bar{x} \pm 1.96\sigma_{\bar{x}}$, where a 99% confidence interval derives from the rule, $\bar{x} \pm 2.58\sigma_{\bar{x}}$. Once an interval has been constructed, it either will or will not include the population value, you do not know which condition holds.

***Testing O Hypothesis:***

The second aspect of statistical inference is testing of hypothesis. Much of our knowledge concerning the world we live in is the result of samples. We eat at a restaurant once and we form an opinion concerning the quality of the food and service at that restaurant. Quite often the opinions we form from the sample are not accurate. However,

in most cases, the opinions are more accurate than if no sample had been observed, and usually the larger the sample, the more accurate the opinion.

Our process of forming opinions may be placed within the framework of an investigation. We will refer to the collection of all elements under investigation as a population. A sample is a collection of some elements of a population. Scientific investigations are often concerned with obtaining information about some populations. Suppose a psychologist wishes to study the effect of constantly interrupted sleep on the emotional balance of a person. He might consider the population to be all human beings of contemporary times. To consider his experiment, he uses paid persons, obtained through an advertisement in a newspaper. He can broadly consider his subjects to be representative of the population and volunteer for a somewhat personal study. He is forced to use this type of sample for his experiment for practical reasons such as limited funds and limited time available for research. In order to obtain accurate information about a population, it would seem desirable to examine every element in that population. Usually this is impossible or impractical, so only a sample from that population is obtained. The sample may consist of those elements that are easily accessible or a haphazard selection of elements from the population. None of these methods of obtaining a sample permits the use of statistical techniques to aid in making inferences about the population, because they do not result in a random sample. Usually we assume that the sample is random even if it is not, but it is much better actually to have a random sample. We assume that the sample is a random sample for our discussion.

One of the primary purposes of sample information is to estimate unknown properties of the population. The unknown properties that may be estimated are necessarily numerical and include such as unknown proportions, means, probabilities and so on. Actually the estimate is based on a sample, a random sample if probability statements are to be made, and the estimate is an educated concerning some unknown property of the probability distribution, where the distribution represents some quantity of interest in the population. Hypothesis testing is the process of inferring from a sample whether or not to accept a certain statement about the population. This statement is called the hypothesis. Examples of hypotheses include such as:

i) Nursery school helps a child to achieve better marks in elementary school
ii) Women are not likely than men to have automobile accidents.

In such cases the hypothesis is tested on the basis of the evidence contained in the sample. The hypothesis is either rejected, meaning the evidence from the sample casts enough doubt in the hypothesis for us to say with some degree of confidence that the hypothesis is false, or accepted, meaning that it is not rejected. A test of a particular hypothesis may be very simple to perform and is a rule to accept or reject a hypothesis.

The different steps involved in such tests are the following:

1) The hypotheses are stated in terms of the population values

2) A test statistic is selected
3) A rule is made, in terms of possible values of the test statistic, for deciding whether to accept or reject the hypothesis
4) On the basis of a random sample from the population, the test statistic is evaluated, and a decision is made to accept or reject hypothesis.

To test any hypothesis on the basis of a sample of observations, we must divide this (all the possible sets of observations) sample space into two regions. If the observed sample point (value of statistic) falls into one of these regions, say C, we shall reject the hypothesis; if the sample point falls into the complementary region, $\overline{C}$, we shall accept the hypothesis. C is known as the critical region of the test, and $\overline{C}$ is called acceptance region. Sometimes the critical region is also called rejection region. The hypothesis being tested is called null hypothesis and is denoted by $H_0$.

There are two ways of making an incorrect decision in hypothesis testing. If the null hypothesis is true we might make the mistake of rejecting it, thus committing an error known as error of the first kind, or type I error. That is type I error occurs when $H_0$ is true and yet the outcome of our experiment is in the critical region. The second way of committing an error in hypothesis testing is by accepting the null hypothesis when it is false. This error is known as of the second kind, or a type II error.

These two types of errors have associated with certain probabilities, of the errors being made. The probability of rejecting a true null hypothesis is called as the level of significance or sometimes is called the size of the critical region. The level of significance may be found by first assuming $H_0$ is true and then ascertaining the probability of getting a point in the critical region. If $H_0$ is a hypothesis, the assumption that $H_0$ is true leads to only one probability function defined on the sample space, and $\alpha$, the significant level, may be found by addition of the probabilities of all points in the critical region. Usually, however, it is easier to find $\alpha$ by computing the probability that the test statistic will assume one of the values that results in rejection of $H_0$, under the assumption that $H_0$ is true.

It is desirable in hypothesis testing for probabilities, type I and type II errors to be close to zero. In practice the sample size helps to determine how small these errors may become. To detect a false hypothesis we consider (1 – probability of type II error) as it gives the probability of rejecting a false hypothesis.

If the test statistic has been chosen so as to result in a one or two-tailed test, the selection of a critical region depends on the experimenter's preference concerning the size of the critical region, the level of significance.

Usually a desirable decrease in the level of significance, $\alpha$ is accompanied by an undesirable increase in $\beta$. Our two objectives in hypothesis testing are to reject $H_0$ as

seldom as possible if $H_0$ is true and as often as possible if $H_0$ is false. As a result the critical region usually the set of points with the largest value of 1-$\beta$ from among those sets of points of some fixed significance level by convention more than any other reason. $\alpha$ is usually chosen near 0.05 or 0.01 and the critical region is then selected in terms of possible values of the test statistic. The results of a hypothesis test are much more meaningful if the value of the critical level is also stated.

A test, where the rejection region corresponds to the largest values ( or smallest values) of the test statistic, is called a one-tailed test. If the test statistic is selected so that the largest values of the test statistic and the smallest values of the statistic, combined, correspond to the rejection region, the test is called a two-tailed test.

Once the hypotheses are formulated, there are usually several hypothesis tests available for testing the null hypothesis. In order to select one of these tests, we consider carefully several properties of various tests. One of the most important question is, 'are the assumptions of this test valid assumptions in any experiment?' If the answer is, 'No' the test probably should be discarded. From among the tests that are appropriate, based on the above criterion, the best test may be selected on the basis of other properties like the following.

1. The test should be unbiased
2. The test should be consistent
3. The test should be more efficient in some sense than the other tests.

An unbiased test is a test in which the probability of rejecting $H_0$ when $H_0$ is false is always greater than or equal to the probability of rejecting $H_0$ when $H_0$ is true. A sequence of tests is consistent against all alternatives in the class $H_1$ if the power of the tests approaches 1.0 as the sample size approaches infinity, for each fixed alternative possible under $H_1$.

Let $T_1$ *and* $T_2$ represent two tests that test the same $H_0$ against the same $H_1$, with the critical regions of the same size $\alpha$ and with same levels of $\beta$. The relative efficiency of $T_1$ *to* $T_2$ (or efficiency of $T_1$ *relative to* $T_2$) is the ratio $\frac{n_2}{n_1}$, where $n_1$ *and* $n_2$ are the sample sizes of the tests $T_1$ *and* $T_2$ respectively. If the alternative hypothesis is composite, the relative efficiency may be computed for each probability function defined by the alternative hypothesis, resulting in a multitude of values for relative efficiency that may then be represented in a table or, occasionally, graphically.

A test is conservative if the actual level of significance is smaller than the stated level of significance.

A statistical method is nonparametric if it satisfies at least one of the following criteria.

1. The method may be used on data with a numerical scale of measurement
2. The method may be used on data with an ordinal scale of measurement

3. The method may be used on data with an interval or ratio scale of measurement where the distribution function of the random variable producing the data is either unspecified or specified except for an infinite number of unknown parameters.

Depending upon the research hypothesis and the data available, one should use the appropriate test to test that hypothesis.

### 6.9 Important Statistical Tests

**Test 1:** Rank correlation test for randomness of sample

***Purpose:*** To test the drawn sample have a random nature (i.e., to test for random sample)

***Data:*** $x_1, x_2, \ldots\ldots\ldots, x_n$ are drawn observations

***Assumptions:*** Observations are obtained by some procedure.

***Hypotheses:*** $H_0$ : The observations are drawn randomly or there is no presence of trend in the observations after arranging them in either in ascending or descending order.

$H_1$ : The sample is not a random sample.

***Test Statistic:***

$$Z = \frac{6D - n(n^2 - 1)}{n(n+1)\sqrt{n-1}}$$

Where $D = \sum_{i}^{n} (x_i - y_i)^2$ in which $x_i$ represents the rank of the i th observation as in the original drawn sample and $y_i$ is the rank obtained of the i th observation after arranging the observations in ascending order. N is number of observations in the sample.

***Decision*** : For n>30 the test statistic, Z is distributed as a standard normal variate. Hence if the value of |Z| is greater than the corresponding tabulated value (obtained from normal tables) we reject the null hypothesis by drawing the conclusion that the is not a random sample. Otherwise we accept the null hypothesis.

***Test 2:*** Z-test for a population mean

***Purpose***: To test the mean of a population.

***Data:*** $(x_1\ x_2 \ldots, x_n)$ is the random sample drawn from the given population.

***Assumptions:*** *The given population is a normal and its variance, $\sigma^2$ is known.*

***Hypotheses***: A- $H_0$ : $\mu = \mu_0$ B-$H_0$: $\mu = \mu_0$ C – $H_0$: $\mu = \mu_0$

$H_0 : \mu \neq \mu_0$ $\quad$ $H_0: \mu > \mu_0$ $\quad$ $H_0: \mu < \mu_0$

Where $\mu$ is the mean of the population from which the sample is drawn and $\mu_0$ is the specified value $\mu$.

***Test Statistic:*** $Z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

Where $\bar{x}$ is the sample mean, $\sigma$ is the population standard deviation and n is the sample size.

***Decision:*** Z is a standard normal variate.

A – If the value of |Z| is greater than the corresponding tabulated value we reject the null hypothesis stating that $\mu \neq \mu_0$. This is a two-tailed test.

B- If the value of Z is greater than the corresponding right-tail tabulated value (from normal tables) we reject the null hypothesis stating that mean is greater than $\mu_0$. This is known as a right-tail test.

C- If the value of Z is smaller than the corresponding left-tail tabulated value (from normal tables) we reject the null hypothesis stating that mean is smaller than $\mu_0$. This is known as a left-tail test otherwise we accept the null hypothesis.

***Test 3:*** t-test to test for a population mean

***Purpose:*** To test the mean of a population.
***Data:*** $(x_1\ x_2 \ldots, x_n)$ is the random sample drawn from the given population.
***Assumptions***: Given population is normal and its variance is not known.
Also n is small (say $\leq 30$)
***Hypotheses***: : A- $H_0 : \mu = \mu_0$ $\quad$ B-$H_0: \mu = \mu_0$ $\quad$ C – $H_0: \mu = \mu_0$

$H_1 : \mu \neq \mu_0$ $\quad$ $H_1: \mu > \mu_0$ $\quad$ $H_1: \mu < \mu_0$

***Test Statistic:*** $t = \dfrac{\bar{x} - \mu_0}{S / \sqrt{n}}$

Where $\bar{x}$ is the sample mean, $\mu_0$ is the assumed mean, n is the sample size and

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

***Decision***: t is a student t-variate with n-1 degrees of freedom when n < 30, otherwise t is a normal variate.

A – If the value of |t| is greater than the corresponding two-tail tabulated value (obtained form t-tables with n-1 degrees of freedom, when n<30 we reject the null hypothesis stating that $\mu \neq \mu_0$. This is a two-tailed test.

B- If the value of t is greater than the corresponding right-tail tabulated value (from normal tables) we reject the null hypothesis stating that mean is greater than $\mu_0$. This is known as a right-tail test.

C- If the value of t is smaller than the corresponding left-tail tabulated value (from normal tables) we reject the null hypothesis stating that mean is smaller than $\mu_0$. This is known as a left-tail test otherwise we accept the null hypothesis.

Similar conclusions as above can be drawn by referring the normal tables respectively when $n \geq 30$.

***Test 4:*** Binomial test for a proportion

***Purpose:*** To test the proportion of a category in a given population
***Data:*** A random sample of n elements in which the proportion of the specified category is p.
***Assumptions***: the sample is small (n<30)
***Hypotheses***: : A- $H_0 : p = p_0$ B-$H_0$: $p = p_0$ C – $H_0$: $p = p_0$
$\mu = \mu_0$

$H_1 : p \neq p_0$ $H_1$: $p > p_0$ $H_1$: $p < p_0$

Where p is the sample proportion and $p_0$ is the specified value of p i.e., populations proportion.

***Test Statistic:*** b = p

***Decision***: When n < 30, b is a binomial variate with n $p_0$ and variance $\frac{p_0(1-p_0)}{n}$

A – If the value of b is either smaller or larger than the tabulated two-tail value from Binomial tables we reject the null hypothesis stating that the proportion of specified category in the population is not equal to $p_o$. Otherwise we accept the null hypothesis.

B- If the value of b is greater than the right-tail tabulated value from Binomial tables we reject the null hypothesis stating that the population proportion of the specified category is larger than $p_0$ Otherwise we accept the null hypothesis.

C- If the value of b is smaller than the left-tail tabulated value from binomial tables we reject the null hypothesis stating that the population proportion of the specified category is smaller than $p_0$ otherwise we accept the null hypothesis.

***Test 5:*** Z-test for a proportion

***Purpose:*** To test the proportion of a category in a given population
***Data:*** A random sample of n elements in which the proportion of the specified category is p.
***Assumptions***: the sample is large ($n \geq 30$)
***Hypotheses***: : A- $H_0 : p = p_0$ B-$H_0$: $p = p_0$ C – $H_0$: $p = p_0$
$H_1 : p \neq p_0$ $H_1$: $p > p_0$ $H_1$: $p < p_0$

Where p is the sample proportion and $p_0$ is the specified value of p i.e., populations proportion.

***Test Statistic:*** $Z = \dfrac{p - p_0}{\{p_o(1 - p_0)/n\}^{\frac{1}{2}}}$

***Decision***: Here Z is a standard normal variate.

A – If the value of |Z| is larger than the two-tail tabulated value from normal tables we reject the null hypothesis stating that the proportion of specified category in the population is not equal to $p_0$. Otherwise we accept the null hypothesis.

B- If the value of Z is greater than the right-tail tabulated value from normal tables we reject the null hypothesis stating that the population proportion of the specified category is larger than $p_0$. Otherwise we accept the null hypothesis.

C- If the value of Z is smaller than the left-tail tabulated value from Binomial tables we reject the null hypothesis stating that the population proportion of the specified category is smaller than $p_0$. Otherwise we accept the null hypothesis.

***Test 6:*** Z-test for the equality of two proportions

***Purpose:*** To test the equality of the proportion of two populations of a specified category.

***Data:*** Proportion of the specified category in the random sample drawn from the first population, say $p_1$ and the proportion of specified category in the random sample drawn from the second population, say $p_2$ are available.

***Assumptions***: the sample is large ($n_1, n_2 \geq 30$)

***Hypotheses***: : A- $H_0 : p_1 = p_2$ B-$H_0$: $p_1 = p_2$ C – $H_0$: $p_1 = p_2$

$H_1 : p_1 \neq p_2$ $H_1$: $p_1 > p_2$ $H_1$: $p_1 < p_2$

Where p is the sample proportion and $p_0$ is the specified value of p i.e., populations proportion.

***Test Statistic:*** $$Z = \frac{p_1 - p_2}{\left\{ p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right\}^{\frac{1}{2}}}$$

Where $p_1$ and $p_2$ are sample proportions, $n_1$ and $n_2$ are the sample sizes and

$$P = \frac{np_1 + n_{21}p_2}{n_1 + n_2}$$

***Decision***: Here Z is a standard normal variate.

A – If the value of x is larger than the two-tail tabulated value from normal tables we reject the null hypothesis stating that the proportion of specified category in the population is not equal to $p_0$. Otherwise we accept the null hypothesis.

B- If the value of Z is greater than the right-tail tabulated value from normal tables we reject the null hypothesis stating that the proportion in the first population is larger than the corresponding proportion in the second population. Otherwise we accept the null hypothesis.

C- If the value of Z is smaller than the left-tail tabulated value from normal tables we reject the null hypothesis stating that the proportion in the first population is smaller than the corresponding proportion of the second population. Otherwise we accept the null hypothesis.

***Test 7:*** $x^2$ – for a population variance

***Purpose:*** To test the variance of a population.

***Data:*** A random sample ($x_1, x_2, \ldots\ldots, x_n$) drawn form the given population..

***Assumptions***: The population is a normal with unknown variance, $\sigma^2$

***Hypotheses***: : A- $H_0 : \sigma^2 = \sigma_0^2$ B-$H_0$: $\sigma^2 = \sigma_0^2$ C – $H_0$: $\sigma^2 = \sigma_0^2$

$H_1 : \sigma^2 \neq = \sigma_0^2$ $H_1$: $\sigma^2 > \sigma_0^2$ $H_1$: $\sigma^2 < \sigma_0^2$

Where $=\sigma_0^2$ is the assumed value of the population variance, $\sigma^2$ .

***Test Statistic:*** $x^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$

Where n is the sample size, $\sigma^2$ is assumed variance $S^2 = \dfrac{\sum_1^n (x_i - \bar{x})^2}{n-1}$

***Decision***: Here $x^2$ is a chi-square variate with (n-1) degrees of freedom.

A – If the value of $x^2$ is either smaller or larger than the two-tail tabulated values respectively (obtained from $x^2$ –tables with n-1 degrees of freedom as $\chi_1^2$ and $\chi_2^2$) we reject the null hypothesis stating that the variance is not equal to $\sigma^2$. Otherwise we accept the null hypothesis, $H_0$

B- If the value of $x^2$ is greater than the right-tail tabulated value (obtained from $x^2$ –tables with n-1 degrees) we reject the null hypothesis stating that the variance is greater than $\sigma^2$. Otherwise we accept the null hypothesis, $H_0$

C- If the value of $x^2$ is smaller than the left-tail tabulated value (obtained form $x^2$ –tables with n-1 degrees) we reject the null hypothesis stating that the variance is smaller than $\sigma^2$. Otherwise we accept the null hypothesis, $H_0$

***Test 8***: F-test for two population variances.

***Purpose***: To test the equality of the variances of two populations.
***Data:*** Two random samples ($x_1, x_2, \ldots\ldots, x_{n_1}$) and ($y_1, y_2, \ldots\ldots, y_{n_2}$) drawn from the two given populations respectively.
***Assumptions:*** Both the populations are normal with unknown variances $\sigma_1^2$ and $\sigma_2^2$ respectively. These populations are not related or the samples are drawn independently.

***Hypotheses***: A – $H_0 : \sigma_1^2 = \sigma_2^2$ B – $H_0 : \sigma_1^2 = \sigma_2^2$ C- $H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$ $H_1 : \sigma_1^2 > \sigma_2^2$ $H_1 : \sigma_1^2 < \sigma_2^2$

***Test Statistic:***

$F = \dfrac{s_1^2}{s_2^2}$ if $s_1^2 > s_2^2$

Where $s_1^2 = \dfrac{\sum_{i=1}^{n_1}(x_i - \bar{x})^2}{n_1 - 1}$ and $s_2^2 = \dfrac{\sum_{i=1}^{n_2}(y_i - \bar{y})^2}{n_2 - 1}$

***Decision:*** Here F is distributed as a Snedecors F with $[(n_1 - 1), (n_2 - 1)]$ degrees of freedom.

A. If the value of F is either smaller or larger than the two-tail tabulated values respectively (obtained from F-tables with $[(n_1-1), (n_2-1)]$ degrees of freedom as $F_1$ and $F_2$) we reject the null hypothesis stating that the variances are not equal. Otherwise we accept the null hypothesis, $H_0$.

B. If the value of F is greater than the right-tail tabulated value (obtained from F-tables with $[(n_1-1), (n_2-1)]$ degrees of freedom) we reject the null hypothesis stating that the variance of the first population is greater than the variance of the second population. Otherwise we accept the null hypothesis, $H_0$.

C. If the value of F is smaller than the left-tail tabulated value (obtained from F-tables with $[(n_1-1), (n_2-1)]$ degrees of freedom) we reject the null hypothesis stating that the variance of the first population is smaller than the variance of the second population. Otherwise we accept the null hypothesis, $H_0$.

***Remark:*** Always larger variance $s_1^2$ or $s_2^2$ normal should be in the numerator in computing the value of F. Accordingly the degrees of freedom should also be considered.

***Test 9***: Z-test for two population means.

***Purpose***: To test the equality of the means of two populations.
***Data:*** Two random samples $(x_1, x_2, \ldots\ldots, x_{n_1})$ and $(y_1, y_2, \ldots\ldots, y_{n_2})$ drawn from the two given populations respectively.
***Assumptions:*** Both the populations are normal with unknown variances $\sigma_1^2$ and $\sigma_2^2$ respectively. These populations are not related or the samples are drawn independently.

***Hypotheses***: A – $H_0 : \mu_1 = \mu_2$ $\quad$ B – $H_0 : \mu_1 = \mu_2$ $\quad$ C- $H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$ $\quad$ $H_1 : \mu_1 > \mu_2$ $\quad$ $H_1 : \mu_1 < \mu_2$

Where $\mu_1$ and $\mu_2$ are means of the first and second populations respectively.
***Test Statistic:***

$$Z = \frac{(\bar{x} - \bar{y})}{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)^{\frac{1}{2}}}$$

Where $\bar{x}$ and $\bar{y}$ are the means, $\sigma_1^2$ and $\sigma_2^2$ are the variances, $n_1$ and $n_2$ are the sample sizes.

***Decision:***

A. If the value of $|Z|$ is larger than the two-tail tabulated value from normal tables we reject the null hypothesis stating that the variances are not equal. Otherwise we accept the null hypothesis, $H_0$.

B. If the value of F is greater than the right-tail tabulated value from normal tables we reject the null hypothesis stating that the mean of the first population is larger than the mean of the second population. Otherwise we accept the null hypothesis.

C. If the value of Z is smaller than the left-tail tabulated value from normal tables we reject the null hypothesis stating that the mean of the first population is smaller than the mean of the second population. Otherwise we accept the null hypothesis.

***Test 10:*** t-test for two population means (variances unknown)

***Purpose:*** To test the equality of means of two populations

***Data:*** Two random samples ($x_1$, $x_2$, ....., $x_{n_1}$) and ($y_1$, $y_2$, ..... $y_{n_2}$) drawn from the two given populations respectively.

***Assumptions***: Both the populations are normal with unknown variances $\sigma_1^2$ and $\sigma_2^2$ respectively. Those populations are not related or the samples are drawn independently

***Hypotheses***: : A- $H_0$ : $\mu_1 = \mu_2$ B-$H_0$: $\mu_1 = \mu_2$ C – $H_0$: $\mu_1 = \mu_2$

$H_1$ : $\mu_1 \neq \mu_2$ $H_1$: $\mu_1 > \mu_2$ $H_1$: $\mu_1 < \mu_2$

Where $\mu_1$ and $\mu_2$ are means of the first and second populations respectively.

***Test Statistic:*** $$t = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{S\left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{\frac{1}{2}}}$$

Where $S^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$, $s_1^2 = \frac{\sum_1^{n_1}(x_i - \bar{x})^2}{n_1 - 1}$, $s_2^2 = \frac{\sum_1^{n_2}(y_i - \bar{y})^2}{n_2 - 1}$

***Decision***: Here t is a Student t with $(n_1 + n_2 - 2)$ degrees of freedom.

A. If the value of $|t|$ is larger than the two-tail tabulated values from t-tables we reject the null hypothesis stating that the means are not equal. Otherwise we accept the null hypothesis.

B. If the value of t is greater than the right-tail tabulated value from t-tables with $(n_1 + n_2 - 2)$ degrees of freedom we reject the null hypothesis stating that the mean of the first population is larger than the mean of the second population. Otherwise we accept the null hypothesis.

C. If the value of t is smaller than the left-tail tabulated value from t-tables with $(n_1 + n_2 - 2)$ degrees of freedom we reject the null hypothesis stating that the mean of the first population is smaller than the mean of the second population. Otherwise we accept the null hypothesis.

***Test 11:*** Paired t-test for two population means (unknown variances)
***Purpose:*** To test the equality of means of two populations
***Data:*** Two random samples $(x_1, x_2, \ldots\ldots, x_n)$ and $(y_1, y_2, \ldots\ldots y_n$ drawn from the two given populations respectively.
***Assumptions***: Both the populations are normal with unknown variances $\sigma_1^2$ and $\sigma_2^2$ respectively. These populations are not related or the samples are drawn independently

***Hypotheses***: : A- $H_0$ : $\mu_1 = \mu_2$ B-$H_0$: $\mu_1 = \mu_2$ C – $H_0$: $\mu_1 = \mu_2$

$H_1$ : $\mu_1 \neq \mu_2$ $H_1$: $\mu_1 > \mu_2$ $H_1$: $\mu_1 < \mu_2$

Where $\mu_1$ and $\mu_2$ are means of the first and second populations respectively.

***Test Statistic:*** $t = \dfrac{\bar{d}}{S/\sqrt{n}}$

Where $\bar{d} = \bar{x} - \bar{y}$, $S^2 = \dfrac{\sum_1^{n_1}(d_i - \bar{d}_1)^2}{n-1}$, $d_i = x_i - y_i$ and n is the sample size

***Decision***: Here t is a Student t with $(n - 1)$ degrees of freedom.

A. If the value of $|t|$ is larger than the two-tail tabulated values from t-tables we reject the null hypothesis stating that the means are not equal. Otherwise we accept the null hypothesis.

B. If the value of t is greater than the right-tail tabulated value from t-tables with (n-1) degrees of freedom we reject the null hypothesis stating that the

mean of the first population is larger than the mean of the second population. Otherwise we accept the null hypothesis.

C. If the value of t is smaller than the left-tail tabulated value from t-tables with (n-1) degrees of freedom we reject the null hypothesis stating that the mean of the first population is smaller than the mean of the second population. Otherwise we accept the null hypothesis.

***Remark:*** If n is large ($\geq 30$) we can refer normal tables for tabulated value and the decisions are same as above.

***Test 12:*** $x^2$ – test for goodness of fit.

***Purpose:*** To compare the observed frequencies with those obtained under specified assumption.

***Data:*** ($O_1, O_2, \ldots\ldots, O_k$) are observed frequencies and ($e_1, e_2, \ldots\ldots, e_k$) are the expected frequencies with the specified assumptions.

***Assumptions***: Expected and observed frequencies are to be computed if these are not given.

***Hypotheses***: : $H_0$ : Specified assumption (say the distribution is normal) is correct
$H_1$ : Specified assumption is wrong.

***Test Statistic:*** $x^2 = \sum_{1}^{k} \frac{(o_i - e_0)^2}{e_i}$

Where $o_i$ s and $e_i$ s are observed and expected frequencies.

***Decision***: Here $x^2$ is a chi-squared variate with (k-1) degrees of freedom. If the computed value is greater than the tabulated value obtained from $x^2$ –tables with (k-1) degrees of freedom we reject the null hypothesis. Otherwise we accept the null hypothesis.

***Test 13(a)*** : $x^2$ –test for independence.

***Purpose:*** To test the independence of two attributes or variables in a contingency table.

***Data***: Frequencies in a contingency table format. For example

| First Attribute / Variable | Second Attribute / Variable | | |
|---|---|---|---|
| | I | II | III |
| 1 | | | |
| 2 | | | |
| 3 | | | |

***Assumptions:*** The sample is a random sample. Each observation may be classified into exactly one of r categories according to first criterion and into exactly one of s different categories according to second criterion. Data is arranged in a contingency table format with r rows and s columns. Each cell value is greater than 5.

***Hypotheses:*** $H_0$ : The two attributes /variables are independent.
$H_1$ : The two attributes / variables are not independent.

***Test Statistic:*** $x^2 = \sum_1^r \sum_1^s \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$

where $o_{ij}$ and $e_{ij}$ are the observed and expected frequencies of the ij th cell.

***Decision:*** Here $x^2$ is distributed as a chi-squared variate with (r-1)(s-1) degrees of freedom. If the calculated value of $x^2$ is greater than the tabulated value (obtained form $x^2$ – tables with (r-1)(s-1) degrees of freedom) we reject the null hypothesis, $H_0$, stating that the attributes/variables involved are not independent. Otherwise we accept the null hypothesis.

***Remark:*** If any cell value is less than 5 we combine that column or row containing the cell with the adjacent column or row to make all the cell values are greater than 5. But finally we should have at least k x 2 or 2 x k table (k > 2). If not, we follow other $x^2$ –tests separately for 2 x 2 and k x 2 tables.

***Test 13(b):*** t – test of a regression coefficient.
***Purpose:*** To test the significance of the regression coefficient in simple regression context.
***Data:*** May be in the form of a set of independent values and the corresponding set of dependent values.
***Assumptions***: The dependent variable say y follows a normal distribution. The variance of the dependent variable is same for any given values of the independent variable. The regression equation is of the form $y = a + b(x - \bar{x})$
***Hypotheses:*** A – $H_0$ : $\beta = 0$
$H_1$: $\beta \neq 0$

Where $\beta$ is the population regression coefficient

***Test Statistic:*** $F = \frac{b^2 \sum_1^n (x_i - \bar{x})^2}{\sum_1^n (y_i - \bar{y})^2 - b^2 \sum_1^n (x_i - \bar{x})^2} \frac{n-2}{1}$

***Decision:*** Here t is distributed like a Student t variate with (n-2) degrees of freedom.

***Test 14:*** t – test of a correlation coefficient.
***Purpose:*** To verify the significance of the correlation coefficient (incase of a simple correlation)
***Data:*** ($y_1$, $y_2$, ....., $y_n$) and ($x_1$, $x_2$, ....., $x_n$) as the values of the dependent and independent variables respectively.
***Assumptions***: x and y values are drawn from a bivariate normal population. The relation is linear.
***Hypotheses:*** A – $H_0$ : There is no relation or the value of correlation coefficient, $\rho = 0$

$H_1$: There is a relation or the value of correlation coefficient, $\rho \neq 0$

B – $H_0$ : $\rho = 0$
$H_1$: $\rho > 0$
C - $H_0$ : $\rho = 0$
$H_1$ : $\rho < 0$

***Test Statistic:*** $t = \frac{r}{\sqrt{1-r^2}}\sqrt{n-2}$

Where r is the sample correlation coefficient and n is the sample size.

***Decision:*** Here t is distributed like a Student t variate with [n-2] degrees of freedom

A. If the value of |t| is greater than the two-tail tabulated value (obtained from t-tables with n-2 degrees of freedom) we reject the null hypothesis stating that the variables are related. Otherwise no relation.
B. If the value of t is greater than right –tail tabulated value (obtained from t-tables with n-2 degrees of freedom) we reject the null hypothesis stating that there is a positive relationship. Otherwise there is no relation.
C. If the value of t is smaller than left-tail tabulated value (obtained from t-tables with n –2 degrees of freedom) we reject the null hypothesis stating that there is a negative relationship. Otherwise there is no relation.

***Test 15***: Z-test of a correlation coefficient.
***Purpose:*** To test the specified value of the correlation coefficient.
***Data:*** ($y_1$, $y_2$, ..... , $y_n$) and ($x_1$, $x_2$,....., $x_n$) as the values of the dependent and independent variables respectively.
***Assumptions:*** x and y values are drawn form a bivariate normal population. The relation is linear.
***Hypotheses:*** A – $H_0$ : There is no relation or the value of correlation coefficient, $\rho = 0$
$H_1$: There is a relation or the value of correlation coefficient, $\rho \neq 0$

B – $H_0 : \rho = \rho_0$

$H_1 : \rho > \rho_0$

C- $H_0 : \rho = \rho_0$

$H_1 \ \rho < \rho_0$

Where $\rho$ is the actual value of the population correlation coefficient and $\rho_0$ is assumed value of $\rho$.

***Test Statistic:*** $Z = \dfrac{Z_1 - \mu_1}{\sigma_1}$ where $Z_1 = \dfrac{1}{2}\log_e\left(\dfrac{1+r}{1-r}\right)$, $\mu_1 = \dfrac{1}{2}\log_e\left(\dfrac{1+\rho_0}{1-\rho_0}\right)$, $\sigma_1 = \dfrac{1}{\sqrt{n-3}}$,

r is the sample correlation coefficient, and n is the number of pairs of observations of (x, y).

***Decision:*** Here Z is a standard normal variate.

A. If the value of |Z| is greater than the two-tail tabulated value (obtained from normal tables) we reject the null hypothesis stating that the value of the correlation coefficient is not equal to $\rho_0$. Otherwise $\rho = \rho_0$
B. If the value of Z is greater than right –tail tabulated value (obtained from normal tables) we reject the null hypothesis stating that the value of $\rho$ is greater than $\rho_0$ Otherwise $\rho = \rho_0$
C. If the value of Z is smaller than the left-tail tabulated value (obtained from normal tables) we reject the null hypothesis is stating that the value of $\rho$ is smaller than $\rho_0$. Otherwise $\rho = \rho_0$

***Test 16:*** Z-test for many correlation coefficients.

***Purpose:*** To compare the equality of several correlation coefficients.

***Data:*** $(x_{i1}, x_{i2}, \ldots\ldots, x_{in_i})$ $(y_{i1}, y_{i2}, \ldots\ldots, y_{in_i})$ are drawn from the i th bivariate population

(i =1,2, ……, k).

***Assumptions:*** x and y values are drawn from each of k-bivariate normal distribution. The vaiance of y is independent of x values in each of the k-distributions. The relationship is linear in all the cases.

***Hypotheses:*** $H_0 : \rho_1 = \rho_2 = \ldots\ldots = \rho_k$

Where $\rho_i$ is the correlation coefficient in the i th population.

***Test Statistic:*** $Z = \dfrac{\sum_1^k (Z_i - \mu_i)}{\left(\sum_1^k \dfrac{1}{n_i - 3}\right)^{\frac{1}{2}}}$ where $Z_i = \dfrac{1}{2}\log_e\left(\dfrac{1+\rho_i}{1-\rho_i}\right)$ and $n_i$ is the size of the i th sample drawn from i th population.

***Decision:*** Here Z is a standard normal variate. If the value of |Z| is greater than the two-tail tabulated value (obtained from normal tables) we reject the null hypothesis stating that all the correlations are not equal. Otherwise they are all equal.

***Test 17:*** F-test for k population means (ANOVA)
***Purpose:*** To test equality of the means of several independent populations.
***Data:*** ($x_{i1}$, $x_{i2}$,....., $x_{in_i}$) drawn from the i th population (i =1,2, ....., k).

***Assumptions:*** Populations are normal and have equal variances. Samples are drawn independently.

***Hypotheses:*** $H_0 : \mu_1 = \mu_2 = \ldots\ldots = \mu_k$

$H_1 : \mu_1 \neq \mu_2 \neq \ldots\ldots \neq \mu_k$

Where $\mu_i$ is the mean of the i th population.

***Test Statistic:*** $F = \dfrac{\sum_1^k n_{ii}\left(\overline{x_i} - \overline{x}\right)^2}{\sum_1^k n_i s_1^2} \dfrac{N-k}{k-1}$

where $n_i s_1^2 = \sum_1^{n_i}\left(x_{ij} - \overline{x_i}\right)^2$, $\overline{x_i} = \dfrac{\sum_1^{n_i} x_{ij}}{n_i}$, $\overline{x} = \dfrac{\sum_1^k \sum_1^{n_i} x_{ij}}{N}$, $N = \sum_1^k n_i$

***Decision:*** Here F is a Snedecor's F with [k-1, N-k] degrees of freedom. If the value of F is larger than the tabulated value (obtained from f-tables with [k-1, N-k] degrees of freedom) we reject the null hypothesis stating that all the means are not equal. Otherwise they are all equal.

***Remark:*** Always numerator sum of square should be large

***Test 18:*** Tukey test for k population means
***Purpose:*** To compare the means of k population means and to find which means differ if all are not equal.

***Data:*** ($x_{i1}$, $x_{i2}$, ....., $x_{in_i}$) drawn from the i th population (i = 1, 2,....., k)
***Assumptions:*** Populations are normal and have equal variances. Samples are drawn independently.
***Hypotheses:*** $H_0 : \mu_1 = \mu_2 = \ldots\ldots = \mu_k$

$H_1 : \mu_1 \neq \mu_2 \neq \ldots\ldots \neq \mu_k$

Where $\mu_i$ is the mean of the i th population.

(All observations in the sample which are larger than the median value are given a + sign and those below the median are – sign. A succession of values with the same sign is called a run and the number of runs is k.)

***Decision:*** Case1: For n < 30. If the value of k is either small or large when compared to the tabulated value (from run test tables) we reject the null hypothesis, otherwise we accept the null hypothesis.

Case2: For n $\geq 30$. Here k is distributed as a normal variate with mean (n+1) and variance $\frac{n(2n-2)}{2(2n-1)}$ and then

$$Z = \frac{k-(n+1)}{\left\{\frac{1}{2}n(2n-2)/(2n-1)\right\}^{\frac{1}{2}}}$$ is a standard normal variate. If the value of |Z| is greater than the two-tail tabulated value (from normal tables) we reject the null hypothesis. Otherwise we accept $H_0$.

***Test 27***: F-test for testing main effects and interaction effects in a two-way classification
***Purpose:*** To test the significance of the main effects and interaction effects.
***Data:*** $x_{ijk}$, i = 1,2,....., p (level A), j = 1, 2, ....., q (level B) and r = 1,2,.....k.
***Assumptions:*** The sample observations are normally distributed. Equal number of multiple observations in each cell.
***Hypotheses:*** A- $H_0$ : $\alpha_i = 0$ B- $H_0$ : $\beta_j = 0$ C- $H_0$ : $(\alpha\beta)_{ij} = 0$

$H_1$ : $\alpha_i \neq 0$ $H_1$ : $\beta_j \neq 0$ $H_1$ : $(\alpha\beta)_{ij} \neq 0$

Where $\alpha_i$ and $\beta_j$ are the main effects of i th level of the factor A and j th level of factor B respectively; $(\alpha\beta)_{ij}$ is the interaction between the i th level of A and j th level of B.

***Test Statistic:*** $F_1 = \frac{MSA}{MSE}$, $F_2 = \frac{MSB}{MSE}$, $F_3 = \frac{MSAB}{MSE}$

$$\text{Where MSA} = \frac{rq\sum_{i=1}^{p}(x_i - x_{....})^2}{p-1}$$

$$\text{MSB} = \frac{rp\sum_{j=1}^{q}(x_j - x_{....})^2}{q-1}$$

$$\text{MSAB} = \frac{r\sum_{i=1}^{p}\sum_{j=1}^{q}(x_{ij} - x_{....})^2}{pq-1}$$

$$\text{MSE} = \frac{\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r}\left(x_{ijk} - x_{\ldots\ldots}\right)^2}{rpq - pq - p - q - 2}$$

***Decision:***

A. Here $F_1$ is distributed like Snedecors F with [(q-1), pq(r-1)] degrees of freedom. If the value of $F_1$ is greater than the tabulated value (obtained from F-tables with [(p-1), pq(r-1)] degrees of freedom we reject the null hypothesis, stating that there are some effects due to factor A.

B. Here $F_2$ is distributed like Snedecors F with [(q-1), pq(r-1)] degrees of freedom. If the value of $F_2$ is greater than the tabulated value (obtained from F-tables with [(q-1), pq(r-1)] degrees of freedom we reject the null hypothesis, stating that there are some effects due to factor B.

C. Here $F_3$ is distributed as a Snedecors F with [(p-1)(q-1),pq(r-1)] degrees of freedom. If the value of $F_3$ is greater than the tabulated value obtained from F-tableswith [(p-1)(q-1), pq(r-1)] degrees of freedom we reject the null hypothesis, stating that there are no interaction effects.

***Remarks***: If there is a single observation in a cell then r =1 and hence

MSA =
MSB =
MSAB=
MSE =

Here $F_2$ is a Snedecors F with (p-1) degrees of freedom, $F_2$ is Snedecors F with (q-1) degrees of freedom, and $F_3$ is Snedecors F with (p-1)(q-1) degrees of freedom.

***Test 28***: F-test for linear regression.
***Purpose:*** To test the linear regression.
***Data:*** $\{(x_1,y_1), (x_2, y_2), \ldots\ldots,(x_n, y_n)$ is bivariate sample obtained form the given bivariate population.
***Assumptions:*** Observations are drawn independently. Y = a + bx
***Hypotheses:*** $H_0$ : The regression of Y on X is linear.
$H_1$ : The regression of Y on X is not linear.

***Test Statistic:*** $F = \dfrac{MSR}{MSE}$ where $\text{MSR} = b^2 \sum_{1}^{n} n_i\left(x_i - \bar{x}\right)^2$

$$\text{MSE} = \sum_{1}^{n}\left(y_i - \bar{y}\right)^2 - b^2\sum_{1}^{n} n_i\left(x_i - \bar{x}\right)^2 / (n-2)$$

***Decision:*** Here F is distributed as a Snedecors F with [1, (n-2)] degrees of freedom. If the value of F is greater than the tabulated value of F (obtained from F-tables with [1, (n-2)] degrees of freedom we reject the null hypothesis, stating that the regression is not linear. Otherwise the regression is linear.

***Test Statistic:*** W = $\frac{q.S}{\sqrt{n}}$ where $S^2 = \frac{\sum_{1}^{k}(n-1)s_i^2}{N-k}$, $N = \sum_{1}^{k} n_i$, $s_i^2 = b\frac{\sum_{1}^{n_i}(x_{ij}-\bar{x}_i)}{n_i - 1}$,

$n = \frac{k}{\frac{1}{n_1} + \frac{1}{n_2} + \ldots\ldots + \frac{1}{n_k}}$ and q is the studentized range available from tables.

***Decision:*** If W is greater than the absolute difference between any two sample means, we conclude that those means of populations corresponding to those samples differ. If all absolute differences are smaller than W, we accept null hypothesis.

***Test 19:*** Bartlett's test for of variances of k populations.
***Purpose:*** To test the equality of variances of k populations.
***Data:*** ($x_{i1}$, $x_{i2}$, ....., $x_{in_i}$) drawn from the i th population (i = 1, 2,....., k)
***Assumptions:*** Populations are normal. Samples are drawn independently.
***Hypotheses:*** $H_0 : \sigma_1^2 = \sigma_2^2 = \ldots\ldots = \sigma_k^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2 \neq \ldots\ldots \neq \sigma_k^2$

***Test Statistic:*** $x^2 = \frac{\left\{\sum_{1}^{k}(n_i-1)\log s^2 - \sum_{1}^{k}(n_i-1)\log s_i^2\right\}}{1 + \frac{1}{3(k+1)}\left\{\sum_{1}^{k}\frac{1}{(n_i-1)} - \frac{1}{\sum_{1}^{k}(n-1)_i}\right\}}$

where $S^2 = \frac{\sum_{1}^{k}(n-1)s_i^2}{\sum_{1}^{k}(n_i-1)}$, $s_i^2 = \frac{\sum_{1}^{n_i}(x_{ij}-\bar{x}_i)^2}{n_i - 1}$, and $n_1$ is the size of the sample from the i th population.

***Decision:*** If $x^2$ is chi-squared variate with k-1 degrees of freedom. If the value of $x^2$ greater than the tabulated value (obtained from $x^2$ –tables with k-1 tables) we reject the null hypothesis stating that all the variances are not equal. Otherwise the variances are equal.

***Test 20***: Kolmogorov – Smirnov test for goodness of fit.
***Purpose:*** To test the correctness of the assumed distribution of the population.
***Data:*** ($x_1$, $x_2$, ....., $x_n$) is the drawn sample.
***Assumptions:*** Sample is drawn format he given population.
***Hypotheses:*** $H_0$ : The assumed distribution is the correct one of the population

$H_1$ : Assumed distribution is not the right one for the population.

***Test Statistic:*** D = $|F-S_n|$ where F(x) is the assumed distribution function of the population and $S_n(x)$ is the distribution function of the sample. D is the maximum absolute difference, n is the sample size.

***Decision:*** If the value of D is greater than the tabulated value (obtained from Kolmogorov – Smirnov tables) for given we reject the null hypothesis stating that the assumed distribution is not the correct one. Otherwise we accept $H_0$.

***Test 21***: Kolmogorov – Smirnov test for two populations
***Purpose:*** To test the equality of the distributions two given populations.
***Data:*** $(x_1, x_2, \ldots\ldots, x_{n_1})$ and $(y_1, y_2, \ldots\ldots, y_{n_2})$ are the two samples drawn from the given populations respectively.
***Assumptions:*** $n_1$ and $n_2$ are sufficiently large.
***Hypotheses:*** $H_0$ : The two population distributions are identical
$H_1$ : The two population distributions are not identical.

***Test Statistic:*** D = $|S_1(x) - S_2(y)|$ where $S_1(x)$ and $S_2(y)$ are the distribution functions of the drawn samples and D is the maximum absolute difference of the sample distribution functions.

***Decision:*** If the value of D is greater than the tabulated value (obtained from Kolmogorov – Smirnov tables) we reject the null hypothesis. Otherwise we accept the null hypothesis.

***Test 22***: Median test for comparison of two population distributions
***Purpose:*** To test the equality of the distributions two given poplations.
***Data:*** $(x_1, x_2, \ldots\ldots, x_{n_1})$ and $(y_1, y_2, \ldots\ldots, y_{n_2})$ are the two samples drawn from the given populations respectively.
***Assumptions:*** $n_1$ and $n_2$ are sufficiently large.
***Hypotheses:*** $H_0$ : The two population distributions are identical
$H_1$ : The two population distributions are not identical.

***Test Statistic:*** $x^2 = \dfrac{\left\{|ad-bc| - \frac{1}{2}N\right\}^2 N}{(a+b)(a+c)(b+d)(c+d)}$ where a, b, c, d are given as

| | Sample 1 | Sample2 |
|---|---|---|
| Left of median | a | b |
| Right of median | c | d |

N=a + b + c + d

***Decision:*** Here $x^2$ is a chi-squared variate with 1 degree of freedom. If the value of $x^2$ is greater than the tabulated value (obtained from the $x^2$ – tables with 1 degree of freedom) we reject the null hypothesis stating that the distributions are not identical. Otherwise we accept $H_0$.

***Test 23***: Median test for k populations
***Purpose:*** To test the equality of the distributions of k given populations.
***Data:*** $(x_1, x_2, \ldots\ldots, x_{n_i})$ for i = 1, 2, ……, k are k samples drawn from k given populations.
***Assumptions:*** $n_1, n_2, \ldots\ldots, n_k$ are sufficiently large.
***Hypotheses:*** $H_0$ : The distributions of the given k populations are identical
$H_1$ : The distributions of the given k populations are not identical.

***Test Statistic:*** $x^2 = \sum_1^k \frac{(a_{1j} - e_{1j})^2}{e_{1j}} + \sum_1^k \frac{(a_{2j} - e_{2j})^2}{e_{2j}}$ where $a_{ij}$ and $e_{ij}$ are as given below

| | Sample 1 | Sample2 . . . . . Sample k | | Total |
|---|---|---|---|---|
| Left of median | $a_{11}$ | $a_{12}$ | $a_{1k}$ | A |
| Right of median | $a_{21}$ | $a_{22}$ | $a_{2k}$ | B |
| Total | $a_1$ | $a_2$ | $a_k$ | N |

$e_{1j} = \frac{A \times a_j}{N}$ and $e_{2j} = \frac{B \times a_j}{N}$ for j = 1,2, . .. . . . . ., k

***Decision:*** Here $x^2$ is distributed as chi-squared variate with k-1 degrees of freedom. If the value of $x^2$ is greater than the tabulated value (obtained from $x^2$ –tables with k-1 degrees of freedom) we reject the null hypothesis otherwise we accept the null hypothesis.

***Test 24***: Wilcoxon – Mann-Whitney rank test for two populations
***Purpose:*** To test the equality of the means of two populations.

***Data:*** ($x_1, x_2, \ldots\ldots, x_{n_1}$) and ($y_1, y_2, \ldots\ldots, y_{n_1}$) are the two samples drawn from the given populations respectively.
***Assumptions:*** Populations are having continuous distributions.
***Hypotheses:*** $H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$
Where $\mu_1$ and $\mu_2$ are the means of the given populations.
***Test Statistic:*** $R_1 = n(N+1) - R$
Where R is the sum of the ranks of smaller sample, n is the size of the smaller sample and N is the size of the combined sample.

***Decision:*** If the value of $R_1$ is smaller than the tabulated value (obtained from Wilcoxon-Mann-Whitney tables) with given $n_1$ and $n_2$ we reject the null hypothesis, stating that the means are not equal. Otherwise we accept $H_0$.

***Test 25***: Kruskall –Wallis test for k populations
***Purpose:*** To test the equality of the means of k given populations.
***Data:*** ($x_{i1}, x_{i2}, \ldots\ldots, x_{in_1}$) for i = 1, 2, . . .. ....., k are k samples drawn from k given populations.
***Assumptions:*** Each sample size should be at least 5. The distributions of the populations are continuous.
***Hypotheses:*** $H_0$ : All the means of the given k populations are equal.
$H_1$ : All the means of the given k populations are not equal.

***Test Statistic:*** $x^2 = \left\{ \frac{12}{N(N+1)} \sum_1^k \frac{R_i^2}{n_i} \right\} - 3(N+1)$ where $R_i$ is the sum of the ranks of i th sample observations, $n_i$ is the size of the i th sample and N is the size of the combined sample, which is equal to $\sum_1^k n_i$

***Decision:*** Here $x^2$ follows a chi-square distribution with k-1 degrees of freedom. If the value of $x^2$ is greater than the tabulated value (obtained from $x^2$-tables with k-1 degrees of freedom) we reject the null hypothesis, stating that all the means are not equal. Otherwise we accept the null hypothesis.

***Test 26***: Run test for randomness in a sample.
***Purpose:*** To test the random nature of the sample.
***Data:*** ($x_1, x_2, \ldots\ldots, x_n$) are the drawn observations from the given population.
***Assumptions:*** All observations obtained in the same conditions.
***Hypotheses:*** $H_0$ : The observations in the sample are drawn randomly.
$H_1$ : The observations in the sample are not drawn randomly.
***Test Statistic:*** k = number of runs

***Test 29***: F-test for coefficient of multiple regression

***Purpose:*** To test the significance of multiple regression coefficient.

***Data:*** $x_{ij}$ I = 1, 2, ....., k and j = 1, 2, ....., n is multivariate sample drawn from a multivariate population.

***Assumptions:*** Each $x_{ij}$ is normally distributed. Sample observations are drawn independently.

***Hypotheses:*** $H_0$ : R=0 or $\beta = \beta_2 = \ldots\ldots\ldots = \beta_k$ for k = 1,2, ....., p-1

$H_1$ : $R \neq 0$ or Not all $\beta_k = 0$

Where R is the multiple correlation coefficient and $\beta_i$ is the regression coefficient corresponding to i th variate.

***Test Statistic:*** $F = \dfrac{MSH}{MSE}$ where MSH = Mean sum of square under $H_0 = \sum_{11}^{k} b_i \dfrac{\sum_{i}^{n} y_i x_{ji}}{k}$

MSE = Mean sum of square due to error = $\sum_{1}^{k}(y_i - \bar{y})^2 - \sum_{1}^{k} b_i \sum_{1}^{k} \dfrac{y_i x_{ji}}{(n-k-1)}$

$b_i$ is the least square estimate of $\beta_i$.

$$MSE = \sum_{1}^{n}(y_i - \bar{y})^2 - b^2 \sum_{1}^{n} n_i (x_i - \bar{x})^2 / (n-2)$$

***Decision:*** Here F is distributed as a Snedecors F with [k, n-k-1] degrees of freedom. If the value of F is greater than the tabulated value of f tables with [k, n-k-1] degrees of freedom we reject the null hypothesis. Otherwise we accept the null hypothesis.

***Test 30*** : Mc Nemar test for significance of changes.

***Purpose:*** To test the change in the population from one situation to another is significant or not.

***Data:*** $(x_i, y_i)$, i = 1, 2, ....., n is a bivariate random sample drawn form the given bivariate population.

***Assumptions:*** The pairs $(x_i, y_i)$ are mutually independent. The measurement scale is nominal with two categories, for all $x_i$ and $y_i$. The difference scale is nominal with two categories, for all $x_i$ and $y_i$. The difference $P(x_i = 0, y_i = 1) - P(x_i = 1, y_i = 0)$ is negative for all I, or zero for all I, or positive for all i.

***Hypotheses:*** $H_0$ : $P(x_i = 0, y_i = 1) = P(x_i = 1, y_i = 0)$ for all i

$H_1$ : $P(x_i = 0, y_i = 1) \neq P(x_i = 1, y_i = 0)$ for all i

OR

$H_0$ : $P(x_i = 0, y_i = 1) = P(x_i = 1, y_i = 0)$ for all i

$H_1$ : $P(x_i = 0, y_i = 1) \neq P(x_i = 0, y_i = 0)$ for all I

OR

$H_0$ : $P(x_i = 1) = P(y_i = 1)$ for all i

$H_1 : (x_i = 1) \neq P(y_i = 1)$ for all i

***Test Statistic:*** $T_1 = \dfrac{(b-c)^2}{b+c}$ or $T_2 = b$ where

| | $y_i = 0$ | $y_i = 1$ |
|---|---|---|
| $x_i = 0$ | a | b |
| $x_i = 0$ | c | d |

***Decision:*** Case 1: When $b+c \leq 20$, $n = b + c$. If the value of $T_2$ is smaller or larger than the two-tail tabulated value (obtained from Binomial tables) we reject the null hypothesis, stating that the change is significant. Otherwise the change is not significant.

Case 2: $b + c > 20$. Here $T_1$ is a chi-squared variate with one degree of freedom. If the value of $T_1$ is greater than the right tail tabulated value (obtained form $x^2$ –tables with 1 degree of freedom we reject the null hypothesis, stating that the change is insignificant.

***Test 31***: Cox and Stuart test for trend.
***Purpose:*** To test for the existence of trend in the population.
***Data:*** $x_1, x_2, \ldots\ldots\ldots, x_n$ are the observations arranged in either i ascending order or descending order.
***Assumptions:*** The random variables $x_1, x_2, \ldots\ldots\ldots x_n$ are mutually independent. The measurement scale is at least ordinal. $x_{1+c}$ is the first value in the second group where $c = \dfrac{n}{2}$ when n is even and $\dfrac{n+1}{2}$ when n is odd and reject the central value.
***Hypotheses:*** A - $H_0$ : No trend
$H_1$ : There is a trend.

B - $H_0$ : There is no upward trend
$H_1$ : There is an upward trend

C - $H_0$ : There is no downward trend
$H_1$ : There is a downward trend.

OR

A - $H_0$ : $P(x_i < x_{i+c}) = P(x_i > 1\ x_{i+c})$ for all i
$H_1$ : $P(x_i < x_{i+c}) \neq P(x_i > 1\ x_{i+c})$ for all i

B - $H_0$ : $P(x_i < 1\ x_{i+c}) \leq P(x_i > 1\ x_{i+c})$ for all i
$H_1$ : $P(x_i < x_{i+c}) > P(x_i < x_{i+c})$ for all i

C - $H_0$ : $P(x_i < 1\ x_{i+c}) \geq P(x_i > 1\ x_{i+c})$ for all i
$H_1$ : $P(x_i < x_{i+c}) < P(x_i > x_{i+c})$ for all i

***Test Statistic:*** T = number of pluses. (If $x_i > x_{i+c}$ then + is assigned and if $x_i < x_{i+c}$ then – is assigned )

$$Z = \frac{T - \frac{n}{2}}{\sqrt{\frac{n}{4}}}$$

***Decision:*** Case 1: for ≤ 20.

A. If the value of T is either smaller or larger than the two-tail tabulated value (obtained from Binomial tables we reject the null hypothesis. Otherwise we accept the null hypothesis.
B. If the value of Z is larger than the right-tail tabulated value obtained form normal tables we reject the null hypothesis. Otherwise we accept the null hypothesis.
C. If the value of Z is smaller than the left-tail tabulated value obtained from normal tables we reject the null hypothesis. Otherwise we accept the null hypothesis.

Case 2: For > 20. Z is a standard normal variate.

A. If the value of |Z| larger than the two-tail tabulated value obtained from normal tables we reject the null hypothesis. Otherwise we accept the null hypothesis.
B. If the value of Z is larger than the right-tail tabulated value obtained from normal tables we reject the null hypothesis. Otherwise we accept the null hypothesis.
C. If the value of Z is smaller than the left-tail tabulated value obtained form normal tables we reject the null hypothesis. Otherwise we accept the null hypothesis.

***Test 32***: Cochran's test for means of related populations.
***Purpose:*** To test the equality of the mean effects of several related populations.
***Data:*** Each of s treatments is applied independently to each of r blocks and the result of each treatment as "success" or "failure". The results are then given in the form of a table with r rows representing the blocks and s columns representing the s treatments.
***Assumptions:*** The blocks were randomly selected from the population of all possible blocks. The outcomes of the treatments may be dichotomized in a manner common to all treatments within each block, so the outcomes are listed as either "O" or "1".

***Hypotheses:*** $H_0$ : The treatments are equally effective
$H_1$ : There is difference in effectiveness

***Test Statistic:*** $$T = \frac{(s-1)\sum_{1}^{s} S_j^2 - (s-1)N^2}{sN - \sum_{1}^{r} R_i^2}$$

Where $S_j = \sum_{1}^{r} x_{ij}$, $R_i = \sum_{1}^{s} x_{ij}$, $N = \sum_{1}^{r}\sum_{1}^{s} x_{ij}$, s = number of columns and r = number of rows.

***Decision:*** Here T is distributed like a chi-squared variate with s-1 degrees of freedom when number of rows are sufficiently large. If the value of T is greater than the right-tail tabulated value (obtained from $x^2$ –tables with s-1 degrees of freedom) we reject the null hypothesis, stating that there is a difference in effectiveness among treatments. Otherwise we accept the null hypothesis.

***Test 33***: Cochran and Stuart test for two population means.
***Purpose:*** To test the equality of the mean of two populations (when the variance is not equal and unknown)
***Data:*** $(x_1, x_2, \ldots\ldots, x_n)$ and $(y_1, y_2, \ldots\ldots, y_n)$ drawn from the two populations respectively.
***Assumptions:*** The populations are normal with unknown and unequal variances. Two samples are independently.

***Hypotheses:*** : A- $H_0 : \mu = \mu_0$, $H_1 : \mu = \mu_1$ B- $H_0 : \mu = \mu_0$, $H_1 : \mu > \mu_1$ C- $H_0 : \mu = \mu_0$, $H_1 : \mu < \mu_1$

***Test Statistic:*** $$t = \frac{\bar{x} - \bar{y}}{\left(\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}\right)^{\frac{1}{2}}} \quad \text{where } s_1^2 = \frac{\sum_{1}^{n_1}(x_i - \bar{x})^2}{n_1}, \quad s_2^2 = \frac{\sum_{1}^{n_1}(y_j - \bar{y})^2}{n_1}$$

***Decision:*** Here t s a student t with $(n_1 + n_2 - 2)$ degrees of freedom.

A. If the value of $|t|$ is larger than the two-tail tabulated value from t-tables with
$(n_1 + n_2 - 2)$ degrees of freedom we reject the null hypothesis stating that the means are not equal. Otherwise we accept the null hypothesis.
B. If the value of t is greater than the right-tail tabulated value from t-tables with $(n_1 + n_2 - 2)$ degrees of freedom we reject the null hypothesis stating that the mean of the first population is larger than the mean of the second population. Otherwise we accept the null hypothesis.

C. If the value of t is smaller than the left-tail tabulated value from t-tables with
$(n_1 + n_2 - 2)$ degrees of freedom we reject the null hypothesis stating that the mean of the first population is smaller than the mean of the second population. Otherwise we accept the null hypothesis.

Now a days for almost all computations of the values of test statistics, computer softwares available. The prominent among them MS Excel, SPSSS , SYSTAT etc. These softwares are helpful even in interpretation of test results. For advanced Statistical techniques one may refer either SPSS or SYSTAT softwares. With little effort one can make use of these softwares for all common data analysis.

(1) ***An area:*** is used to demarcate a geographic boundary of a DIET, which covers every thing there that relates to school education in general, and elementary education in particular.

(2) ***Jurisdiction***: is used to indicate the geographic boundary in which a DIET operates.

(3) ***Experiment***: is used as a field experiment. By field experiment, what is meant is a systematic activity of studying changes in social setting, which one can see objectively, as a result of the manipulation of certain variables. In field experiment, an idea, a process is tested and its results are verified. It could include all tryouts and innovations too.

[* A definition generated by a group of Educationists in a workshop held at RIE, Mysore]

For example: (a) Experimenting on efficient ways of introducing grades in lieu of marks in schools through tryout of different methods.

(b) Trying an alternate model of internship, or practice teaching in a DIET.

(4) ***Study***: is used as an activity of trying to understand meaning out of existing situation/ information/ data. It is different from experiment in the sense that here nothing is manipulated. The status is studied. After anything is implemented one may like to see how that is working.

For example: (a) Impact of mid day meal scheme on enrollment, quality of pupil participation and their achievement.

(b) Preparedness of D.Ed students to a new internship model.

(c) Evaluation of the worthwhileness of new textbooks introduced.

(5) ***Research:*** is an activity of generating new knowledge. Any activity undertaken systematically which can generate new knowledge in educational context is an educational research. A research may have an experimental possibility built into it. Experiment by itself does not become research. In an experiment, it is attempted to study a cause and effect relationship. Building a new theory /model based on tested conjectures, there by generating a new knowledge is the task of research. There are very thin demarcating lines between studies and researches. Studies use secondary source of information mostly, while researches use primary sources of information mostly.

For example: (a) Comparative study of different methods of teaching in a particular context.

(b) Assessing the systemic preparedness in the introduction of English from class III.

(6) ***Professional***: Teaching is a profession. A teacher is a professional and he has the moral responsibilities of working towards all round development of the personality of his students. Therefore, a teacher provides plenty of learning opportunities by design, and thereby enables students to grow to their full potential. A teacher does not take teaching as a job, but as a profession. Therefore, their responsibilities are professional in nature.

(7) ***Improvement of quality elementary education***: refers to all those processes covering activities and initiatives undertaken by which provision of access, better enrollment, higher retention of students in schools, arrest of wastage and stagnation, improving better standards of transaction and functioning, higher rate of learning as well as higher levels of attainments are achieved at elementary level in a district.

With the above explanation of the definition, it becomes clear that in the DIET context, a lab area is a chosen part of a district, which a DIET can recognize and adopt in order to work intensely investing its human and material resources and undertake activities by which new things are tested; innovations are tried out; new relationships are established; and new knowledge is generated. Putting together all these activities, it can empower a DIET in becoming more professional as well as become capable of informing and influencing educational planners and decision makers leading to qualitative improvement of school education. Therefore, understanding different issues and concerns of Lab area becomes necessary for all the faculty of DIET.

Every DIET has many different functions to perform. There are many functions, which DIET faculty will have to implement through out the district. These routine activities do not become the activities of a lab area, as they are to be implemented uniformly throughout. These are activities, which are non-negotiable in nature.

A lab area is an area, which is not necessarily confined only to DIETs, or its context but also to any professional body, which may need to try out its own activities, programmes and innovations in order to test their efficacy, suggest certain changes based on tested strategies in order to bring qualitative changes in its functioning. In this sense, it is important to know the concept of lab area. This could perhaps be, understood better if it is seen from its attributes and characteristics.

**(ii) Attributes and Characteristics of a Lab area**

Some of the attributes and characteristics of a lab area in the DIET context could include the following:

(1) ***A chosen piece of a district:*** A lab area is a chosen piece of a district, which is chosen by a DIET itself, by design, based on certain criteria.

(2) ***Used for field experimentation:*** This chosen piece of an area is used for field experimentation. Experiments are done in a reality situation, which may lead to professional development of DIET faculty as well as enable a district in working better towards improving the quality of elementary education.

(3) ***Lab institutions are a part of a lab area:*** A lab area is an area, which could have a lab school, a lab NFE center or a lab adult education center, etc. It only means that a lab area covers lab institutions. For example, if a DIET chooses a pocket as its lab area, all the formal schools, NFE centers, Adult education centers, the community, all will become a part of Lab area. All institutions, which have a

bearing on elementary level education and adult education, will become the part of that lab area.

(4) ***Different lab areas can also co-exist:*** There can be different lab areas which may focus on different concerns, if need be. Certain pockets can focus on formal elementary education, certain pockets on adult education, & non-formal education, etc. The point that is attempted to be made here is that it is not always necessary that there must be only one lab area at a given point of time. Yet, logistically speaking, it would be all right if a particular geographic piece is chosen as a lab area, as managing different lab areas could pose problems.

(5) ***Can be a part of annual plan:*** A lab area activities can emerge out of activities of annual planning covering different activities. It implies that activities of a lab area can be carved out of annual plans.

(6) ***Lab area activities are non routine activities:*** Activities that are undertaken in a lab area are those, which try to test some thing new and the outcome of which can enhance professional insights of DIET faculty, or provide suggestions to the educational planners, managers, administrators and organizers, or can also help a district education implementation authorities in implementing new ways of doing things based on tried out modes. Therefore, they are not routine activities, which any DIET does.

(7) ***Activities are planned by DIET faculty themselves***: Lab area activities are those activities, which are planned by the faculty of DIET themselves.

(8) ***Members of DIET have a role in it:*** Indeed, faculty members of DIET have their roles in planning and execution of lab area of a district. From this viewpoint collective wisdom and work needs to be ensured in order to have good lab area plans. This point needs to be appreciated by all DIET faculty members.

(9) ***Resources of DIET converge on lab area:*** The DIET will have to converge its resources, lend support in working on lab area activities. Though there are different branches, their respective contributions will have to converge. Therefore, there is a need to develop a healthy coordination within and between different branches of DIET and they all will have to work as one team.

(10) ***Cannot go beyond its legitimate DIET jurisdiction:*** While planning a lab area, the DIET cannot go beyond its DIET boundary. With in the district boundary, it can select any pocket, on its own pre-determined criteria and priority and select as its lab area.

**(iii) Possible Misconceptions about Lab Area**

One can visualize certain possible misconceptions, which may obscure the clarity of the concept of lab area. In this light, it is perhaps necessary to discuss a few possible misconceptions about lab area. They can be explained in DIET context as follows.

(1) ***Lab area planning is prepared by higher ups / state department:*** Lab area planning that will have to be done locally, by DIET faculty, based on the identified priorities. Therefore it is a misconception to think that Lab area planning is prepared by higher ups or state department.

(2) ***Lab area should be physically attached to DIET:*** The lab area is not an area, which is physically attached to the wall of a DIET. As the DIET has a responsibility for the entire district, lab area can be chosen belonging to any part of the district. Therefore, it is a misconception to think that a lab area should be physically attached to DIET.

(3) ***There cannot be more than one lab area:*** There can be more than one lab area at a time. Looking at the responsibilities of the DIET, on priority basis different localities or pockets of a district can be considered as a lab area for different concerns. For instance, a DIET might be studying the effectiveness of certain incentive schemes on enrolment and participation in schooling process of tribal children in one pocket, while they might be studying a different model of practice teaching/internship in another pocket, or they might be interested in undertaking a research study to see the systemic preparedness for the introduction of English at class III level in rural areas. A DIET can afford to have 3 lab areas at a time too. Therefore, there can be different lab areas at a time. But, it should be manageable by a DIET.

(4) ***It is difficult to understand the concept and activities of lab area:*** This misconception can be questioned by all of us if we intend to be professionals in our endeavor to serve the cause of school education. Therefore, it is left to us whether we want to retain this apprehension or throw it.

(5) ***Only DIET faculty has to work for Lab area:*** The concept of a lab area is not confined only to DIET but all professionals and professional institutions can have their lab areas. For example, Mysore Medical College has K.R. Hospital as its lab hospital. Theory is taught in the college and real cases are seen and discussed in hospital.

(6) ***It has nothing to do with DIET's annual plan and programmes and it is an additional burden to DIETs:*** It is erroneous to say that the lab area concept has nothing to do with DIETs annual plans. Activities related to Lab area can emerge from annual plans. Lab area activities enable a DIET to work systematically and meaningfully. This can facilitate them to be more methodical and professional. Therefore it is not an additional burden at all. On the contrary, it is a platform to become more professional.

(7) ***It's planning and implementation is the responsibility of only P&M branch and not others in the DIET:*** Though the guidelines for DIETs clearly mentioned lab area under P&M functions, it needs to be understood as a responsibility of the

entire DIET, but, P&M can perhaps coordinate some of these activities. DIET as a unit has to generate basic data about the district seeking the cooperation of sub district functionaries. Based on the data, all the branches will have to provide necessary inputs in formulation of lab area planning of the district for identifying different roles of different branches and different faculty members. From this perspective, lab area planning and implementation cannot be the sole responsibility of P&M.

(8) ***Identification of lab area is time consuming***: It is indeed a misconception to think that identification of a lab area is time consuming. It is in fact an intelligent activity to identify a lab area. All DIETs have demographic details of the district. Educational indicators are also available. The nature of the proposed activity and its demand will enable a DIET to identify its lab area. Therefore, it is not to be understood as time-consuming activity.

(9) ***Lab area activity is expensive***: It is a misconception to think that lab area activity is expensive. It is a miniature activity undertaken systematically. In fact these activities need to grow more and more. These activities are not at all expensive looking at the power they have in informing and influencing those who matter.

(10) ***Specialized training is necessary to understand the concept and implement it:*** There is no need for any specialized training in order to understand and implement lab area concept. A simple one time training would be all right. Once, one gets clarity of the concept, one can continue to plan and implement.

**(iv) Objectives of Lab Area**

Any lab area in the DIET context will have the following objectives. Acceptance of the Lab area concept enables a DIET faculty:

(a) To conduct experiments, carry out researches and studies thereby enhancing its professional competence, and

(b) To work towards qualitative improvement of school education in the district based on its own experiments, studies or researches.

From this viewpoint, in a lab area, if any activity of DIET needs to be understood whether it could be considered a fit activity under lab area, one can ask herself / himself whether the activity satisfies any one of the objectives. If the answer is yes, then it can be an activity, which can be accepted under lab area.

**(v) Functions of a lab Area**

Any lab area can serve two distinct functions, enabling functions and facilitating functions. Enabling functions are those functions, which enable a DIET to become more professional in their perspectives and practice. Facilitating functions are those, which facilitate a district to achieve quality elementary education in the district.

Therefore, acceptance and adoption of lab area concept not only facilitate a DIET to grow professionally, but it also enables a district to provide quality elementary

education. Thus, it serves twin purposes, i.e., professional development of the institute, and educational development of the district, both based on experiments, studies and researches.

**(a) Enabling functions:** As explained above are those functions, which enable a DIET to become more professional, as the lab area planning and implementation requires systematic and methodical approaches and professional perspectives. Some of the possible suggestive functions can be listed as follows. Adoption of lab area by a DIET can:

(1) Encourage try out of experiments and innovative ideas in a practical situation in lab schools, lab NFE centers or Lab Adult Education centers or community.
(2) Enable DIET faculty to demonstrate how certain ideas theorized can work in actual situations.
(3) Educate DIET faculty by orienting them to real problems that exist in the field and provide a reality orientation. This in turn can tune a DIET faculty to see problems of education realistically and work.
(4) Enable DIET faculty to be methodical, systematic and reality driven in planning programmes based on priorities. Thus, it sharpens the professional perspectives.

**(b) Facilitating functions**: As explained above, these functions facilitate a district to provide quality elementary education based on tryouts, experiments, studies and researches, thereby working towards achieving the objectives of UEE. These functions though not exhaustive, can be listed as follows. The outcomes of Lab area activities, which are based on experiments/ studies/ researches, will be capable of ;

(1) Suggesting alternative ways to facilitate schools to provide access to all children, who are in the age group of 6 to 14.
(2) Suggesting ways to provide alternative education to all those who have missed formal education.
(3) Proposing different models to facilitate enrollment of all children in the district irrespective of caste, sex, disability, language and religion based on its own try out.
(4) Proposing different strategies to retain students in schools and complete elementary level.
(5) Providing tips to school education system, in providing quality education.
(6) Influencing the school education system to provide elementary education, which is rooted in their own culture and thereby making schooling enjoyable and relevant through some suggested methods.
(7) Motivating schools to ensure that minimum levels are achieved by all learners at different levels of schooling, by suggesting certain tried and tested teaching-learning techniques.

**(vi) Identification of a Lab area**

Lab area identification is important as all other activities are to be undertaken there. There are different considerations for identification of a lab area. A lab area is a pocket of a district;

(1) Close by to DIET as well as potential area where research and experiments can be conducted conveniently by DIET faculty.
(2) Where certain variations are desirably needed to be studied.
(3) Which is the requirement of the research / experimental work. (Ex. If certain institutions like multi-grade schools, tribal schools, minority schools, NFE Centers, EGS centers etc are to be studied, the area must have them.) So the research activity concern has to decide.
(4) Where trying innovative ideas require an institution, which can match the context of innovation where it needs to be tried out. For example, if a DIET faculty wants to try out an innovation in an NFE center, the lab area must have an NFE center there.
(5) Where certain indicators of access, enrollment, retention, drop out, wastage, stagnation, quality and achievement are not satisfactory. **Here in some cases, distance may not be a barrier**.
(6) Where activities such as evaluation of textbooks, school effectiveness, testing innovations and such can be undertaken, which can truly represent all variations of the district. **Here again in some cases, distance may not be a barrier.**

With regard to the pre-service activities, a DIET has to reflect on improving its internship styles/practice teaching styles and work towards betterment and making it relevant. So objective evaluation and critical reflection becomes a necessity here.

From the above, it can be inferred that, while selecting a lab area, it is desirable that it satisfies the following conditions.

(1) ***Potentiality of the pocket***: The pocket identified as a lab area need to be potential enough to cover the components meant for study /experiment /innovate /evaluate /study.
(2) ***Manageable in size and activities***: Make sure that the selection of a lab area is such which a team of DIET can manage easily without making it a burden from the viewpoint of management of functions of a lab area. Lab area activities are those activities, which are over and above routine activities. Therefore, let the activities not overshadow their regular functions.
(3) ***Accessibility***: Easy accessibility is almost a requirement for identification of a lab area for testing reflective concerns. Almost the reverse is true for educational development concerns of the district.

**(vii) How to plan a Lab Area activity?**

Planning is an activity, which requires a systematic understanding of objective to be achieved, optimum utilization of the resources that are available including human and

material through well thought out strategies, and considering the time targets that are to be honored. Interplay of these can enable one to plan well. Therefore, what are the objectives that are to be achieved for which we want to plan? What area the available resources with which we can achieve the goals? What are the strategies that can be used? And within what time frame that needs to be achieved? Are the questions, which are to be asked and addressed?

At this point, for the sake of illustration, it is worthwhile to list out a few activities, which can be considered under lab area. If these are found useful in the lab area, they could be recommended for wider implementation. Some of the possible activities that could be listed are as follow:

(a) Trying the alternate models of pre-service training programme. Ex. NCERT is trying how far two year B.Ed programme is better than the traditional one year B.Ed programme, through its lab institutions, the Regional Institutes of Education, located at Ajmer, Bhopal, Bhubaneswar and Mysore.
(b) Study of Systemic preparedness in the introduction of English at earlier levels.
(c) All innovative ideas proposed by DIET faculty could be tried out to see how far they might be effective.
(d) Acceptability and preparedness of trimester system in schools by stake holders.
(e) Tryout of different training models related to switching over to grading.
(f) Try out of strategies, which can persuade migratory population to enroll and continue education.
(g) Trying out of different methods by which community could be made to demand quality education from schools with their active collaboration.

While, planning activities under lab area, different activities require different styles, depending upon the nature of the activity. Let us try to plan a couple of activities as follow. These are only suggestive but not prescriptive.

**Activity 1**

Systemic Preparedness for the introduction of English at class III level

Lab Area: A representative area of a district. [This can be explained with more details]
*Objective:* To study the preparedness of the system for the introduction of English at class III level.
To achieve the above objective, the following strategies can be planned as follows.

| Sl.No. | Sub activity | Resources required | | | | Remarks |
|---|---|---|---|---|---|---|
| | | Human | Material | Money | Expected Time | |
| 1 | Listing attributes / components | DIET faculty | - | - | 7 days | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | of the system and planning activities | | | | | |
| 2 | Identification and searching of tools | DIET faculty and some technical help | Literature | - | 7 days | |
| 3 | Workshop on development and finalization of tools | DIET faculty and some technical people | Stationary items, | Rs.1,000 | 3 days | |
| 4 | Field Work: Assessment of (a) Teachers' competence to use English, (b) TLMs' suitability to introduce English at class III level (c) Views of SDMC members | DIET faculty | Tools, Tape recorders, | Rs. 5,000/- | 8 days | |
| 5 | Scoring and analysis | DIET team | Table work | - | 10 days | |
| 6 | Report Writing and presentation | Activity Coordinator | Typing, Xeroxing Binding Multiple copies making | Rs 1000/- | 10 days | |
| | TOTAL | DIET team | - | Rs 7000/- | 45 working days | |

The above activity requires 45 working days or nearly 2 months and requires around Rs.7,000/-. The output of this quick study could inform and influence the district authorities to feed to policy making. **On these concerns, it would still be a grand idea if all DIETs undertake the same study based on the same design, which will have a much better use in taking macro-level decisions.**

**Activity 2**

*Training schoolteachers and educational administrators to switch over to grading in schools.*

Lab Area: A representative area of a district. [This can be explained with more details]

*(a)* *Objective:* To train schoolteachers and educational administrators to switch over to grading in schools.

The above objective can be achieved by following the suggested steps.

| Sl.No. | Sub activity | Resources required | | | | Remarks |
|---|---|---|---|---|---|---|
| | | Human | Material | Money | Expected Time | |
| 1 | Collect materials on need and importance of grading | DIET | Relevant materials | - | 5 days | |
| 2 | Initial discussion with teachers across, on issues about evaluation in general and problems and inadequacies in specific | DIET team, school teachers and HMs | - | Rs.1,000/- | 12 days | |
| 3 | Assessment of training needs on feedback – Workshop | DIET team | - | Rs.250/- | 3 days | |
| 4 | Planning for a training programme | DIET team | - | - | 2 days | |
| 5 | Training programme | DIET team | Training materials | Rs. 3,000/- | 3 days | |
| 6 | Assessment of | DIET team and | - | - | 1 day | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | effectiveness of training programme | participants | | | | |
| 7 | Report Writing | Activity coordinator | Typing, Xeroxing, Binding, Multiple copies making | Rs. 3000/- | 10 days | |
| | TOTAL | DIET team | - | Rs 7,250/- | 36 working days | |

The above activity requires 36 working days and requires around Rs.7, 250/-. The output of this training programme could inform and influence the schoolteachers to develop their preparedness to shift to grading.

Keeping in view, the above lab area planning needs to be done. The above guidelines are only suggestive of lab area planning and it is not prescriptive. After the planning and its approval, it can be executed suiting time of the faculty responsible for this activity. The execution of lab area activities can go on as per the plan. However, depending upon the nature of the activity, certain flexibility can also be built in. The main objective of using the lab area must not be defeated in the name of other things.

**(viii) Documentation and dissemination of Lab area activities:**

As professionals, DIET faculty need to be familiar with documentation of different activities. Different activities that are undertaken by DIET faculty in lab area need different kinds and styles of documentation. The purpose of documenting any activity of lab area is to chiefly disseminate to those who matter.

(1) **Research:** Documentation of research reports in lab area cannot be different from researches of non-lab area. Therefore, all research activities need to be documented alike following standard steps. But, the dissemination of these findings has different meanings and implications for different personnel in education system. They can be as follows.

(a) Those researches undertaken in lab area on teacher concerns can through new light on the DIET faculty themselves and other teachers at school level. It needs wider dissemination through reports, presentations, and discussions. It thus feeds back to DIET faculty themselves and teachers at large in the district.

(b) Those researches, which have been undertaken to improve the quality of education in the district, or improve teacher conditions in schools, which

may have some policy implications, should reach educational planners, administrators, policy makers and managers of school education. This is a professional responsibility of DIET faculty.

(2) **Experimentation:** DIET faculty can undertake a number of field experimentations in lab area. These tried out experiments need to be reported in simple style which practitioners as well as policy makers can understand and implement at their own levels. What is desirable in these documents is a systematic write up which can convince a reader about the seriousness and sincerity of the experimenter and the experiment itself. Different experiments may need different styles but the documentation has to keep the end users in mind.

(3) **Studies:** Include all those attempts, which have been made /conducted systematically, are generally based on, secondary source of information. These studies are conducted based on available data in the field. With this a DIET faculty can attempt to relate many other factors, which might be responsible for the situation. Thus, these studies are potential enough to understand and analyze the situation. Status studies and qualitative studies based on facts and data can be good examples of 'studies'.

In terms of rigor and its usefulness, they are no less important than research studies. The documentation is a must and depending upon the concern of the study, they need to be shared among professionals as well as policy makers to facilitate policy making which are rooted in empirical realities.

In the above backdrop, it is necessary that all our lab area activities need to be documented and disseminated.

**Sum up:** The concept of lab area needs to be clear, only then all other things can be undertaken.

XXX

## Reference

MHRD.(1989). *District Institutes of Education and Training: Guidelines.* New Delhi: GOI.

# 7 (B) Multi-Centric Studies (MCS)

**Objectives:** At the end of reading this module, a reader will be able to:

(a) Understand the concept of MCS
(b) Enumerate the characteristics of MCS
(c) Understand the need and importance of MCS in education
(d) Identify problems and issues for MCS
(e) List out potential areas for MCS
(f) List out issues related to methodology, and designing and execution of MCS
(g) List out issues related to the execution of MCS
(h) Understand issues related to the reporting of MCS

**(i) Background:** Education is a sub set of social reality. Understanding education and facilitating its dynamic movement is the responsibility of all concerned. It includes educational planners, educational administrators, educational managers, educational researchers, teacher educators, teachers, parents and community at large. If students are to be trained efficiently and nurtured in to future good citizens, all must contribute. Only then the educational sector can progress. This amounts to brining about changes in the educational system.

Brining about changes in any system has to be done based on solid foundations. In the context of education, educational changes need to be brought about, based on needs and requirements of the system. This could be conveniently done based on research results. Quite often, it is felt that many educational changes that are brought about by educational planners are not based on sound empirical evidences, but they are based on gut feelings and political considerations. While, policy formulations done based on the research evidences can really bring about changes.

Researches in education are normally conducted by individual researchers, which are sporadic in nature. Such assorted attempts will not be able to inform and influence any educational policy of a state. In this context, what is desirable is undertaking mega-researches at the state level, which will be capable of informing and influencing policy formulations or capable of studying the effectiveness of the implementation of the policies of the state. One such strategy is being in use in medicine, which we are using conveniently these days in education is called 'multi-centric studies'.

## (ii) Multi Centric Study- Concept

Multi-Centric Study (MCS) is an approach or a strategy of research to understand issues and problems of different local contexts, at a time, using a common research design. It involves micro or case studies, where one wants to uncover deeper issues in

different contexts. In our educational context, it involves multiple researchers engaged in studying a particular problem or issue independently as it exists in their own locations, such as at district and sub district levels. At the end all those are consolidated to get a macro picture. This will be more scientific as the larger picture captured is based on the same methodology used by different collaborators who have worked on a common problem. The scenario that exists at different districts or sub districts could be different also. The strength MCS has its ability to see local as well as global picture that has been studied by a network of researchers at a given point of time.

MCS is seen as an attempt to provide solutions to the operational and managerial problems with in a system, which can clearly spell out policy recommendations and implementation strategies. It is also seen as an important strategy to promote and sustain systemic reforms to enhance quality of school education. Perhaps it is time we think and act on this issue seriously.

**(iii) Definition:** In our context, Multi-centric Study could be defined as ***a study carried out by multiple partners or collaborators located in different parts of a state, on a particular identified theme or aspect, with an intention to identify local context specific issues, capable of informing and influencing policy formulation or testing the efficacy of a policy implementation, leading to systemic reforms.***

**Explanation of the terms used in the definition:** The terms used in the definition are to be understood in our context as follows.

*(a)* ***Study carried out by multiple partners***: In multi-centric studies, a number of partners or researchers take part. In Karnataka SSA context, all the DIETs are the partners. Thus, the study carried out under multi-centric involves multiple partners who are located at different DIETs.

*(b)* ***Particular identified theme:*** In multi-centric studies, all the collaborators or partners undertake a study on the same identified theme. The core team conceptualizing the study would take the lead and involve all the partners or collaborators in this activity.

*(c)* ***Local specific issues*** are those issues, which are specific to the local realities of a district or sub-district level. This indicates that the study does not ignore the ground realities that prevail in a particular local context in the name of aggregating the evidences of different districts. The local issues indeed have their own importance. The district has to develop the capacity to reform itself based on the findings too.

*(d)* ***Capable of informing and influencing policy formulations:*** The study undertaken under multi-centric context is capable of informing policy makers as well as influencing them on what would be an appropriate policy direction. The outcomes of the multi centric studies and the subsequent recommendations will be capable of facilitating the policy makers to take decisions in either reformulating the existing policies or arrive at new policies rooted in empirical reality. This can also be called a 'policy study' as it does the function of facilitating policy formulation.

(e) ***Testing the efficacy of a policy implementation:*** The study undertaken under multi-centric context is also capable of testing the efficacy of policy implementation. It means it can study the extent of success of policy implementation too. From this viewpoint also, a multi-centric study can be called a policy study too.

(f) ***Systemic reforms***: Refers to reformation of the system. Or to put it simply, it means bringing about qualitative changes in the system. In educational context, it means bringing about qualitative changes in all aspects of education and among all functionaries of the education system thereby effecting qualitative improvements in the education of learners.

## (iv) Features and characteristics of Multi-Centric Study

**(1) It involves participation of multiple researchers located at different parts of a larger unit:** Multi-centric study, the name itself suggests that it is a study carried out in different centers. For this to happen multiple researchers are necessary. Therefore, it involves the participation of multiple researchers located at different parts of a larger unit of which they are parts. In Karnataka context, the larger unit is the state and the parts are districts.

**(2) MCS is conducted by different researchers under the overall supervision and direction of a lead team or a lead researcher:** It is important to understand that in a multi centric context; there is a need for a nucleus. This nucleus could be a core team or a lead researcher himself or herself. This nucleus conceptualizes studies in collaboration with its units, provides all guidance, monitors the smooth conduct of study, and facilitates possibility of local variation and global aggregation of out puts. In Karnataka SSA context, the nucleus is the team, which is responsible for Lab area and multi-centric studies.

**(3) MCS is simultaneous activity, which is also collaborative in nature**. The study carried out under MCS is simultaneously carried out by all the partners at the same time in their own local contexts. Though many researchers are engaged in the same study in their own context, the basic design of the study is finalized in collaboration with all the partners. Hence, it is an activity undertaken simultaneously as well as it is collaborative in nature.

**(4) MCS can cover a large area.** Because of its collaborative nature, multi-centric studies can cover larger area or sites at the same time. However, it is to be noted that the same study design is replicated in different geographical areas. The issues of the study are analyzed intensively.

**(5) MCS creates scope for obtaining a macro picture about the problems and issues concerned:** MCS has the capacity to make its data amenable for aggregation from different partners, thereby enabling the core group at the macro level to understand, analyze and explain the obtained results at macro level. This in turn is capable of informing and influencing policy formulation or change or reformulation.

**(v) Purposes of MCS**: MCS serves certain purposes. Therefore, they are very useful. They can be explained as follows:

1. **In understanding the ground realities which can inform and influence a particular policy or its successful implementation by the state**: MCS helps in understanding ground realities quickly with the help of a large number of research teams which in turn can inform and influence a particular policy formulation and it can also assess the effectiveness of the implementation of any policy.

2. **In obtaining a macro picture of the state as a whole by putting together the study outcomes of all the micro studies conducted on a particular theme or area at some point of time:** It would be difficult for any state to have a macro picture based on some small observations or some small scale studies. On the contrary, MCS facilitates obtaining a macro picture of the state as a whole as it not only talks of studying the local realities, but also has a built-in potentiality of making its data amenable for aggregation leading to macro level picture.

3. **In assessing the successful implementation of a particular educational policy:** Many a times, policies are formulated and communicated and executed. They are not assessed for its success rate or failure rate. It goes on till it is revised. While, a well-conceived MCS can assist any system, in our context the state, in getting the assessment of the implementation of that policy. This has a great advantage for policy makers.

4. **For obtaining feed back on an existing policy to make midway corrections and reformulation:** A well conceived MCS would also be capable of enabling policy makers to make certain revisions in the policy implementation or re-formulations of the policies too.

5. **The MCS approach is very useful where problems and issues are deeply entrenched in larger systems:** There are different sectors where MCS are useful like Education, Health and Nutrition etc, which are huge in structure where with in a sector there are different strata. It becomes unmanageable in those sectors to truly understand the changing system and reflect the desirable change in policy formulations.

6. **Elucidate the causal factors underlying specific educational problems in the State as they are encountered under real-life conditions in the field:** MCS can capture causal factors underlying specific educational problems in a state and sub state systems as it captures local as well as global scenarios. With this reforms can be implemented at local as well as global levels. This can include the following.

   (i) ***Gain insights into the different manifestations of the educational problems in backward regions:*** MCS can be an efficient way of

attempting to gain insights into different manifestations of educational problems in backward regions. A specific policy formulation can be made based on this. An honest attempt through MCS can be a sure way of making a realistic policy leading to change and reforms.

(ii) ***Identify factors inhibiting the application of available knowledge for the control of specific educational problems:*** With the undertaking of MCS, it is possible to identify those factors, which are inhibiting the successful implementation of programmes, which are rooted in available knowledge.

*(iii)* ***Identify appropriate strategies to combat these problems, which are capable of application in the current context:*** It is not sufficient if one gains insight but it is necessary to identify appropriate strategies to combat these problems. These strategies must be such which can be implemented in current context.

*(iv)* ***Arrive at overall recommendations leading to the generation of policies at the macro level or assessing the level of success of implementation of policies leading to policy reformulations:*** The ultimate objective of MCS is to arrive at overall recommendations leading to the generation of macro policies or assessing the level of success of the implementation of policies. If the assessment indicates that an implemented policy is not sound, the reformulation of the same should also be possible. The MCS has so much of potentiality. But, ultimately, it depends upon the administrative will of the people who matter to use it to its best.

## (vi) Why Multi-Centric Study in Education?

The multi-centric study/research approach is particularly appealing to educational researchers for addressing issues relating to systemic reforms. Education being a sub-system in a larger social system, the problems and issues are not only complex but also are systemic in nature. Besides, in the Indian context, with a society, which is culturally diverse and pluralistic in nature, the educational outcomes are influenced by the socio-political factors that prevail in a particular given context. The school system, which imparts education apparently, gets aligned within the existing socio-political context thereby influencing the broader educational goals and the immediate learning outcomes as well. Therefore, under the circumstance, the educational problem may manifest itself in different magnitude and forms within a Nation or State or even a District. In such a situation, the educational issues need to be analyzed at disaggregate level so that the proximate causes and explanations are identified with more clarity and specificity so as to arrive at macro policies for bringing about improvement in the system as a whole. It is in this context the multi-centric study/research come in handy for educational researchers and practitioners. Thus, multi-centric research helps us in educational reformations.

## (vii) Multi Centric Research in UEE Context

It is a known fact that even after 58 years of independence, realizing the goal of UEE in India has remained elusive. Various policy interventions in the past both by the center and the individual states to realize the UEE goal have encountered several challenges at the implementation and operational levels owing to deeply entrenched problems in the Indian education/school system. It is in this context systematic research attempts would help in deepening our understanding and in generating scientific knowledge systems with respect to critical problems and issues that prevail in different contexts. Such knowledge systems are of great help to planners, policymakers and practitioners to undertake appropriate corrective steps to realize the intended goals in the direction of effecting systemic reforms.

The need for undertaking multi-centric research assumes relevance in the context of the current national flagship programme of SSA launched in 2000. The SSA has placed the UEE goals in a clear time frame perspective, with a major emphasis on addressing concerns relating to equity, quality and community partnership. Provision of huge resources to individual districts not only for planning and prioritizing educational development, but also for enhancing capacities of various stakeholders in the system have been a distinct feature of the SSA. Separate resources have been allocated under SSA for carrying out research and evaluation by the districts and the DIETs have been identified as potential institutions, which could actively engage in such meaningful activities. The DIET faculty has already been oriented to undertake small research studies, undertake studies and experimentation under Lab area context and therefore engaging in multi-centric research studies should be seen as a next logical step and as a complementary activity. This can enhance their professional capacities and in improving the quality of school education in the district as well.

Apart from the above, there are a variety of innovative interventions launched by SSA both at the national and at the state level. However, successful implementation of such interventions is influenced by the district specific factors, which operate at the systemic levels. In this context, it becomes important to identify such factors, which thwart the progress of school education reform initiatives. Hence multi centric studies approach seems to help the State to address some of these crucial issues.

## (viii) Identification of problems and issues for MCS:

For identifying issues and problems for MCS, it is necessary that we keep the following in mind:

***(a) The problem chosen should be of major concern affecting large number of people:*** For multi-centric study, the problem chosen must be of a major concern affecting a large number of people belonging to different sub systems. Such concerns become more relevant for MCS. The output of MCS must be capable of informing and influencing a policy having implications at different levels. Therefore, the ability to cover different

layers of a system can determine the efficiency of MCS designers. So, the issues covered must have larger representation of subsystems.

(b) ***There are problems, which have already been studied, but some of its aspects require further elucidation through carefully designed field studies*:** There might be some attempts by some researchers but either it is not comprehensive, or it is sporadic not representing the larger macro level and such problems naturally become subjects for MCS. Therefore, the genesis of a concern for a MCS could even be other assorted studies too, which have found some truth but is not capable of generalizing it to the entire state.

***(c)That there are problems, which can provide concrete and feasible recommendations leading to its practical implementation by the government*:** The problem chosen for MCS should be such that the recommendations flowing out of it are realistic in nature and feasible to implement by the government both in terms of cost and operation.

**(ix) Problem identification**: The following may serve as potential source for problem identification under MCS:

***(i) From the available data*:** Many a times, there are facts and figures presented by officials and researchers in different contexts for different purposes which may provoke one to probe further. This may not necessarily to check or verify the accuracy of the data but to understand why it could be true. Such situations can be plenty, and in education, the available statistical database may be one such source. They may serve as one potential source for undertaking studies, which could fit MCSs.

***(ii) Through extensive informal consultations with researchers, academicians and practitioners who have intimate first-hand knowledge of the problem areas*:** The other set of sources of problem identification could be extensive informal consultations with well meaning researchers, practitioners or academicians who have first hand knowledge of different problem areas. Many forums where such concerns are expressed could be converted in to a potential area of study. [For instance, the famous public speech made by Sri R.K.Narayan, the famous author of Malgudi Days, who raised a very important issue in the Rajya sabha, on load of school bag among young children lead up to setting up Yashpal Committee to study and come out with the famous 'Learning Without Burden'.]

***(iii) Through actual field visits to potential project sites*:** There are a number of supervision activities undertaken by different functionaries at different levels. These visits can also through up different areas potential for systemic change. This could be done based on MCS.

***(iv) Issues raised by general public:*** It is not always necessary that problems to be identified for MCS have to be done by specialists or special people. It could even be based on any public debates and public concerns raised by common people in any platform, even on media covering newspapers, TV, or in any associations, etc.

## (x) Potential Areas for Multi Centric Studies under SSA

There are a vast number of areas that have the potential for being considered for multi centric studies in the overall UEE context. However, within the SSA framework, the following problems appear to be significant. To name a few:

(1) School-community relationship and school accountability
(2) Leadership and school management issues
(3) NPEGEL and girls' participation
(4) First generation learners and learning attainments
(5) Teacher perceptions, motivations and tasks
(6) Teacher training and utility value
(7) Curricular innovations and pedagogical management concerns for teachers
(8) Trimester System
(9) Mid day meal scheme implementation and its consequences
(10) Girls education
(11) Effectiveness of in-service training programme
(12) Status study of inclusive education
(13) Implementation of Out of school strategies
(14) Job satisfaction of teachers and pupils achievement
(15) Teachers absenteeism and learners' opportunity time
(16) Linkages-among CRC-BRC-DIET-DSERT, any many more.

**(xi) Methodology and Design:** Multi-centric study approach consists of 'sub-projects' being carried out in different parts of the state using a uniform study design. These studies, therefore, involve identification of competent, dedicated and experienced partners in the different regions, each with a proper institutional base. Because of this multi-centric nature, it requires coordination, monitoring, to and fro feedback, and periodic comparison of notes and experiences. Therefore, it is desirable that researchers engaged in a multi-centric study to work together in perfect harmony on a common project based on a mutually agreed study design and would be willing to come together frequently to discuss and sort out issues.

Generally multi-centric study will have assorted research teams spread out in different regions. However, they will be accountable for one single lead team, which sponsors study activity. The idea behind such an approach is that the individual centers by virtue of being located in different geographical areas will have the first hand knowledge and experience with respect to demographic characteristics and the nature of the institutional set up and the characteristic features of the functionaries working therein. Therefore, it is easier to grasp the ground realities for such researchers rather than some one who comes from outside to examine the problems. Hence, there is an advantage in having such an arrangement. In the present context the research teams will emerge from each of the DIETs located in the districts of Karnataka.

***(a) Planning and designing a MCS:*** Generally multi-centric study is planned by an agency, institution or organization, which is interested in obtaining a macro picture either at the state or some global level with respect to a particular theme or area but also is interested in capturing local context specific issues which thwart the progress of an intervention or a programme. The State MCS committee will develop a common research design for undertaking MCS. However, the state team through participation of the DIETs, the district and state SSA functionaries would evolve the design.

The agency, which plans for a multi centric study, will also generally arrange to provide funding support. It is the responsibility of the sponsoring agency or the lead team to identify a single research theme or topic for one multi-centric study. After identifying the research theme the lead team will come out with a concept or theme paper to situate the research problem. In addition, the lead team will also provide a common research design, common instruments for data collection; common analysis design formats and common sampling design. The team will prescribe research norms and guidelines (what should be done and how should be done, duration of the study etc,) to the participating collaborators, the DIETs in the present context. The lead team will also facilitate the collaborators through periodic academic inputs, technical and logistic support.

In the present context, the SSA lab area and multi-centric team is going to lead the multi-centric study and initiate the DIETs to be the collaborators in it. During the course of research the lead team would play facilitative role in helping the DIETs to carry forward the research activity as per schedule.

***(b) The role of collaborators in multi-centric study:*** The role of collaborators, in this case, the DIETs will be to engage themselves in carrying out a study meant for multi-centric purpose, which has the potential of aggregating the results at the state level eventually. Their specific role and collaboration will depend upon the nature of the study chosen for the MCS.

***(c)Nature of collaboration:*** The research partners enjoy autonomy with respect to conduct study and gather data professionally. The collaborators in a multi-centric study will of course have all the freedom to collect data based on the accepted design. In case if any district wishes to collect some more data they are free to do so using additional instruments. However DIET will have to adhere to the broad research guidelines given by the lead team so far as quantity and quality of the data to be collected, the use of analyses, and valid instruments for data gathering are concerned.

It is however the responsibility of the lead team to define the nature of collaboration that is proposed in a multi-centric study. The lead team would also come out with a schedule of activities under the MCS, which would clearly indicate at what stage the researchers would come together and for what purpose and the periodicity and duration of such meetings etc.

**(xii) Execution of Multi-Centric Studies:** The MCS lead team would provide schedule of research activities, which will have been prepared in consultation with the DIETs only. It also provides necessary research resources and guidelines to the DIETs. However, the State SSA Mission office would provide the required logistic support apart from the administrative and financial resources to the individual DIETs. The DIETs will have the responsibility of conducting the study as accepted. The study is also expected to be over more or less on the same duration and each DIET will have the responsibility of analyzing the data and preparing reports. In the data analysis stage, they can take the help of professionals.

**(xiii) Reporting of the Study:** There is a need to have a common format for writing the study report by the individual DIETs after its completion. The DIETs may refer the guidelines furnished by the MCS lead team in this regard.

**(xiv) Consolidation of Study Findings at Macro Level:** The lead MCS team would ultimately collate and consolidate all the study reports of the individual DIETs to evolve a State report capturing a macro picture about the problem studied. This is also an important activity of the lead team. This consolidation is not only a physical aggregation, but also doing a kind of 'analysis of analysis', which is technically called a 'meta analysis'. From this consolidation process, it is also expected that an executive summary be written. This executive summary will contain recommendations too. From this executive summary, implications for policymaking or revision etc emerge.

**(xv) Policy Implications:** The report consolidated at the state level includes the highlights of district reports also as part B of the report. Apart from this all the district reports are also going to be presented to the state with equal seriousness. The implications for policy making or assessment of policy implementation would emerge naturally from this process. Hopefully, the state policy makers and the multi-centric study team (covering both the lead team members and the representatives of collaborators, i.e., DIETs) meet and present the report. After the presentation, discuss the policy implications academically. Also understand the problems of implementation of any recommendation given by MCS teams. This is important. Many a time researchers become blind to realistic problems in implementing certain recommendations as much as policy makers also become incapable of benefiting from research findings. Therefore, a platform has to be created where; both the researchers and policy makers meet and share their concerns. This is perhaps the ultimate goal of conducting multi-centric studies. Then on, it becomes the moral responsibility of the policy makers to formulate policies as realistically as possible. The state's success on brining about systemic reforms depends upon the intrinsic and genuine concern of the policy makers in being realistic in their approach and to bring about changes based on empirical reality. Any multi-centric study can utmost facilitates policy formulation to this extent.

**(xvi) Sum up:** It is attempted to discuss the concept, purpose, approach, and likely consequences of multi-centric studies. It is hoped that in Karnataka we take pride in saying that we are fore runners in working for systemic reforms, which is the core

concern of SSA. We hope that all those who matter understand its significance and contribute for the success of such a wonderful opportunity we all have with us.

XXX

## References

Miles MB & Huberman AM. (1984). Qualitative Data Analysis: A Sourcebook of New Methods. California: Sage.
http://www.nutritionfoundation of India.res.in
http://www.ialir.net.in
www.vvgnli.net
http://dbtindia.nic.in

# 8. Quality assurance of classroom practitioners through Action Research

**Objectives:** At the end of reading this module, you will be able to:
(a) Understand the concept of Action Research (AR)
(b) List out the steps and process of AR
(c) List out the characteristics of AR
(d) Enumerate possible misconceptions about AR

**8.1. Background:** Action research has been a mid-20th-century concept, which has spread over different fields including education. It is still an evolving concept. It is still getting refined periodically. The development of the idea of action research is generally attributed to Kurt Levin, a Psychologist a phenomenologist who in the immediate post-world war period used it as a methodology for intervening in and researching the major social problems. Lewin maintained that through action research advances in theory and needed social change might simultaneously be achieved. Action Research (AR) according to Lewin consisted of analysis, fact-finding, conceptualization, planning execution, more fact-findings or evaluation and then a repetition of this whole circle of activities, indeed a spiral of such circles. (In, Kemmis, 1982: 13)

Lewin's ideas on action research were almost immediately applied to education as well as social science more generally. It was the work of Stephen Corey at Teachers' College, Columbia University, however and in particular this book Action Research to improve school practice (1953) that spreads the word about Action Research into the mainstream of American education. (In Hopkins, 1998: 46)

**8.2. Definitions:** Action Research has been defined as an approach to research that is based on a collaborative problem-solving relationship between researcher and the client. Which aims at both solving the problem and the generating new knowledge. It developed largely from the work of Kurt Levin and his associates and it involves a cyclical process of diagnosing a change situation or a problem, planning, gathering, taking action and then fact-finding about the results of that action in order to plan and take further action. The central feature of action research is that it uses a scientific approach to study the resolution of important issues together with those who experiences these issues directly. (Coghan and Brannick, 2001).

According to Rapport, R (1970), Action Research aims to contribute to both to the practical concerns of people in an immediate problematic situation and to the goals of social science by joint collaboration with in mutuality.

According to Kemmis, S. (1983), Action Research is a form of self reflective enquiry and undertaken by participants in social (including educational) situations in order to improve the rationality and justice of (a) their own social or educational

practices, (b) their understanding of these practices, and, (c) the situations in which the practices are carried out. It is most rationally empowering when undertaken by participants' collaborative, though it is often undertaken by individuals and sometimes in cooperation with outsiders. In education, action research has been employed in school-based curriculum development, professional development, school improvement programmes and systems planning and policy development.

According to Ebbutt, D. (1985), Action Research is the way in groups of people can organize the conditions under which they can learn from their own experiences. John Eliot, (1991: 61) defines Action Research as the study of the social situation with a view to improving the quality of action within it. He further laments that "Action Research aims at feeding practical judgments in concrete situations and the validity of the theories or hypothesis it generates depends not so much as scientific tests of truth, as on their usefulness in helping people to act more intelligently and skillfully. In action research, theories are not validated independently and then applied to practice. They are validated through practice.

While commenting on reflective teachers, Ross Bondy and Kyle, (1993), opined that "Reflective teachers are never satisfied that they have all the answers. By continually seeking new information, they constantly challenged their own practices and assumptions. In the process, new dilemmas surface and teachers initiate a new cycle of planning, acting, observing and reflecting." Reflective practice is the centrality of action research. Ideally all teachers need to be reflective (to) practitioners. Teachers also need to become autonomous practitioners. Only then, action research will become relevant to them. In this context, the Stenhouse (1984: 69) described the ideal role of the teacher as follows: "Good teachers are necessarily autonomous in professional judgment. They do not need to be told what to do. They are not professionally the dependants of researchers or superintendents or innovators or supervisors. This does not mean that they do not welcome access to ideas created by other people at other places or in other times. Nor do they reject the advice, consultancy support. But, they don't know that ideas and people or not of much a real use until they are digested to the point where they are subject to the teachers' own judgments. In short, it is the tasks of all educationists outside the classroom to serve the teachers, for only teachers are in the position to create good teaching."

Action research has a distinct identity. However whilst it has particular characteristics, action research is also a super ordinate term, one in which a variety of types and models are subsumed. Similarly as action research has been distinguished from other methodologies, authors have also differentiated between different types and models of action research. This could be explained by the inclusion and interplay between the dual dimensions of action and research. As Tripp (1995) notes the way varying, the importance placed upon these two concepts can produce recognizably different kinds of practice.

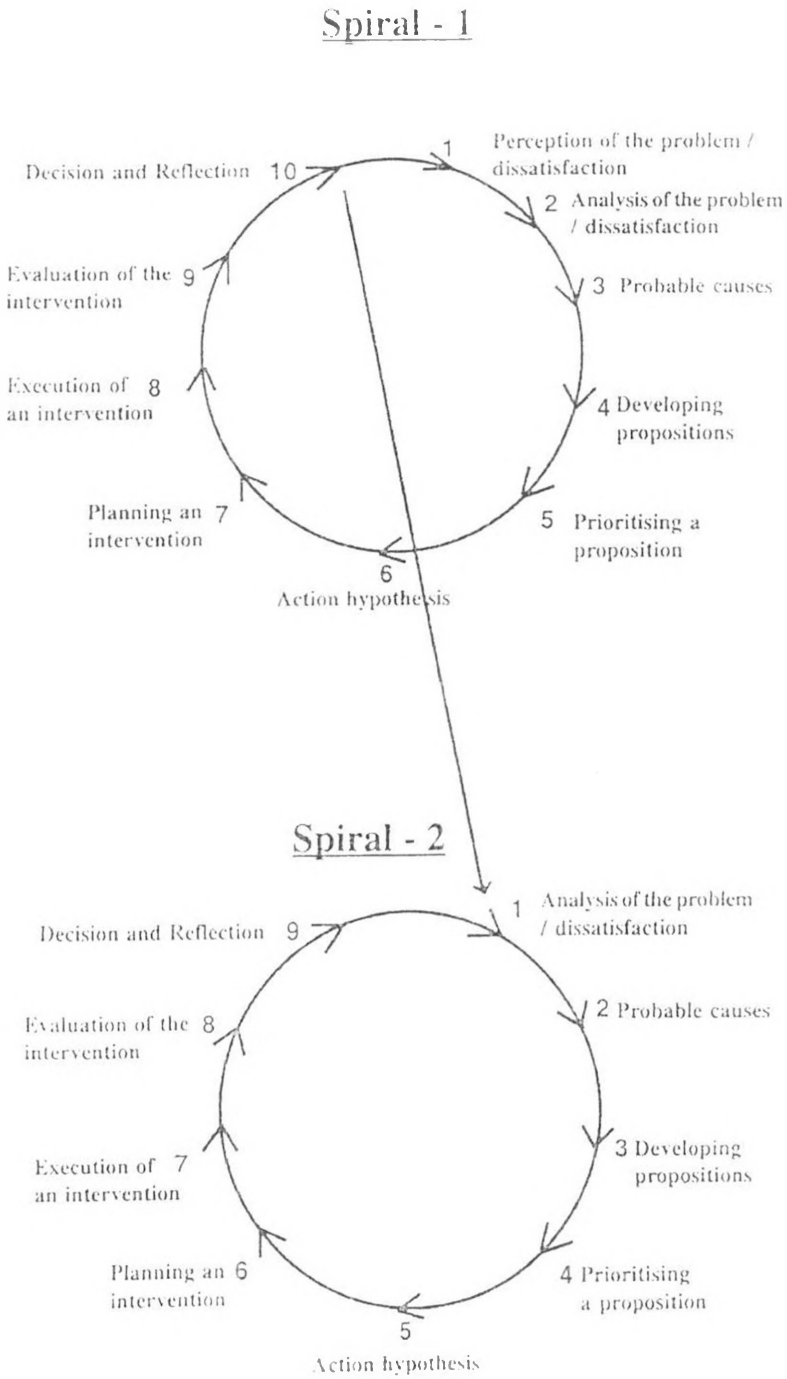
## 8.3. Action research model depicting different steps:

There are different models proposed by different specialists while explaining action research process. The model proposed by the author is as follows.

(1) Perception of the problem / dissatisfaction
(2) Analysis of the problem/ dissatisfaction
(3) Understanding the probable causes
(4) Developing propositions
(5) Prioritizing the proposition
(6) Developing an action hypothesis
(7) Planning an intervention
(8) Execution of an intervention
(9) Evaluation of the intervention
(10)Decision-making, (reflection, explanation and under understanding action.

Education is practice, and it is not just theory. So every teacher is a practitioner. Any practitioner can undertake action research and hence, any teacher can undertake action research. For undertaking action research, a practitioner has to be sensitive to his or her profession. If one is complacent about everything going on around, and then perhaps, one does not have any scope for action research. If one is dissatisfied with any aspect of one's practice there is a score for action research. Therefore no one can compel anyone to undertake action research.

Model




One undertakes Action Research if one feels to bring in some change in his/her practice. Under these premises, the following steps are explained and elaborated.

**(1) Perception of the problem/ dissatisfaction:** If a practitioner is dissatisfied with certain aspects of one's practice, she/he should be sensitive to them. Or there may be

a problem and the practitioner need to be sensitive to the problem, and then a beginning can be made.

Let us take an example to explain the same. A teacher is teaching English to class V students. He is uncomfortable to note that in spite of his best efforts, his students are not able to pronounce words in English satisfactorily. He feels that he needs to do something about it because he is convinced that it is his responsibility to develop good pronunciation among his students. This is an indication that he feels there is a problem and he is dissatisfied with the situation. On the contrary, if he were to be complacent about whatever is happening around him, perhaps, there would not have been scope for Action Research. So this is the first step in action research.

**(2) Analysis of the problem/ Dissatisfaction**: Having felt that he is dissatisfied, he needs to analyze the dissatisfied state or the problem from all possible angles. For this he must objectively be able to explain the dissatisfaction.

Taking the same example, here, at this stage, the teacher must be capable of understanding by way of asking himself;

(a) How many students are unable to pronounce English words?
(b) Are there any specific kinds of problems associated with this?
(c) Are there any specific patterns of errors they have in expression; and based on the answers he gets, he has to move to the next step.

**(3)Understanding the probable causes**: For the kind of dissatisfaction or problem he has to start listing all possible causes from different perspectives. These causes are only possible and plausible causes. These causes are based on his intelligent guesswork too.

In our example, some of the causes a teacher could think of could include the following:

(a) Lack of training
(b) Lack of support in the family
(c) First-generation Learner
(d) Friends in school do not speak in English
(e) Teachers do not encourage students to speak in English
(f) Children speak only in Kannada/ Telugu/Malayalam in their community
(g) Inadequate training in the past classes
(h) Lack of opportunities in school
(i) Second-grade given to English in school
(j) Lack of English listening opportunities
(k) Not interested in English
(l) Students ridicule if someone attempts to speak in English
(m)A feeling that without English also one can survive well
(n) Diffidence in pronouncing English.

The above are some of the hypothetical causes, attempted to be listed. All these causes are to be noted down as assorted points.

**(8) Execution of the intervention:** Here, the actual implementation of the intervention takes place. A good planning could take away all the vagueness of the implementation is when all minor details are planned in advance. This will facilitate focused direction and will save time and energy.

While executing, the practitioner should have a schedule covering activities, men and materials. As per the plan they need to be executed, so that systematically certain evidences or data can be generated. All the tools and techniques that are planned to be used with care and proficiency.

The collected data need to be scored, evaluated and treated qualitatively or quantitatively. This activity will be followed by the next activity, i.e., evaluation of the intervention.

**(9) Evaluation of the effectiveness of intervention:** Evaluation is an activity of assigning a value judgment to a measured attribute. In the action research context, evaluation may cover the following.

(1) Assess the worth or usefulness of the intervention in altering the dissatisfaction state/ or in minimizing the intensity of the problem.
(2) Explain how comprehensive, dependable and relevant, was the intervention?
(3) Assess whether the intervention on the whole succeeded completely, partially or failed totally.

To do this, the kind of data or evidences one collects must be dependable and must be drawn from multiple sources.

**(10)Decision and reflection (based on explanation and understanding the action):** Evaluation is made because certain decisions will have to be taken based on them. Based on the evaluation, the practitioner first has to 'reflect'. 'Reflection' is the centrality of action research. Reflection is the process of stepping back from experience to process what the experience means, with a view to planning further action (Daudelin, 1996; Kolb, 1984; Raelin, 2000: Rigano and Edwards, 1998).

Reflection is the critical link between the concrete experience, the interpretation and taking new action. According to Raelin, (2000) reflection is the key to learning as it enables one to develop an ability to uncover and make explicit to yourself what you have planned, discovered and achieved in practice. He further advocates that reflection must be brought into the open so that it goes beyond your privately held, taken for granted assumptions and helps one to see how knowledge is constructed. In action research, reflection is the activity, which integrates action and research.

Some of the following suggested questions may help in reflection.

(1) What did I think was the problem? Was that all right?
(2) Did the intervention bring about improvement to a satisfactory level?

(3) Is there any scope for further improvement?
(4) What went right at what went wrong?

Based on such questions, certain decisions will have to be taken. These decisions are based on the reflections made by the practitioner. The decisions may include the following suggested ones.

(1) Should I terminate the intervention?
(2) Should I move to the next spiral because the results yielded are not fully satisfied?
(3) What kinds of planning needs to be made keeping in view the present planning which did not help me fully?
(4) How can I make my effective intervention a part of my regular practice?
(5) What efforts and preparations are necessary to naturalize the tested strategy/strategies?
(6) What kind of changes I had to bring in myself in order to naturalize the tested interventions?

Keeping in view the above guidelines, one can take a decision whether one can terminate or move to the next spiral. If one wishes to move to the next spiral, all the steps had to be followed afresh. If terminated, one has to see how tried out strategy/strategies can become a part of one's natural practice.

**8.4. Characteristics of Action Research**: Some of the salient features of action research are as follow:

1. **Small-scale intervention:** Action Research, as a whole is a small scale intervention provided by a practitioner as a part of one's own legitimate activity or practice.
2. **Context specific in nature:** The whole action research is specifically studied, planned and implemented in a specific context and hence the strategies tested here are not generalisable. So in tact groups of beneficiaries are studied.
3. **Reflective practice:** The entire action research process enables the practitioner to reflect on his or her own efficiency and enables him or her in to improve their practice. So, 'reflective practice' is the central feature of action research.
4. **Enhancement of efficiency:** In the process of reflecting on the efficiency, it prompts the practitioner to enhance one's own efficiency. So this is one of the important features of action research.
5. **Practitioner's privilege:** The entire perspective of becoming a 'reflective practitioner' is possible only if one wants to become in reflective practitioner. If one is happy with whatever that is around him/her in whatever quality or manner, one need not undertake action research. So, one can undertake action research if only one wants. Nobody can compel anyone to do action research. One undertakes action research in order to improve one's own practice and efficiency. So it is indeed a practitioner's privilege. An intelligent practitioner considers action research as a boon and hence it becomes his or her privilege.

6. **It is a management mantra:** In an action research, the practitioner does not do different things but he does them differently. Therefore, it is only a management issue for a professional practitioner.
7. **Eliminates monotony in practice:** A reflective practitioner does not suffer from monotony in his or her practice as he/she always looks forward for qualitative betterment. This requires new ways of doing things and freshness in one's practice.
8. **Any practitioner can undertake:** An action research is one, which can be undertaken by any professional practitioner, who wants to enhance one's own professional skills.
9. **Action research is an activity, which has to be undertaken on one's, own initiative and cannot be enforced on the practitioner:** Action research is not an activity which would be enforced upon by yet upon a practitioner. If one is complacent about one's practice, he or she has no scope to undertake action research.
10. **The steps move in spirals:** All the steps on Action Research move in spiral. On reflection, if the practitioner finds that he/she has not been able to turn the situation around, he or she can move to another spiral and try another intervention strategy. So different steps move in a spiral.

## 8.5. Professional concerns to be effective reflective practitioners:

Every functionary in education, whether he or she is a teacher, teacher educator, an administrator, manager or planner can use action research as a strategy to make their profession rewarding. Some of the points every teacher / teacher educator reciprocated needs to keep in mind are as follows.

(1) Teaching is a profession and not a job.
(2) To become effective as the teacher or teacher-educator, it is in the hands of a practitioner himself or herself.
(3) Action Research / Reflective practice need to be understood conceptually and tried out by every practitioner whenever he or she feels that they need to improve a situation around where they have their legitimate role.
(4) Action research is not just problem-solving approach it is beyond that.
(5) Every practitioner must strive to do better and better.
(6) A teacher/ teacher educator must be sensitive to different issues surrounding himself / herself. Anything, which needs improvement, which is within one's own legitimate boundary, needs to be attempted to by a practitioner.
(7) Every practitioner needs to develop professional concerns, professional competence and conviction about doing things and doing things professionally.
(8) Teaching is to be understood as a well-planned and designed intervention and not as a formality of visiting classroom and engaging children irrespective of relevance.
(9) A good professional does not compromise on the quality.
(10) A good professional also attempts to influence the environment around in himself/ herself for qualitative transformation.

If we all seriously think of the issues discussed above, to become a part of our professional life, there is little scope to be pessimistic about becoming effective reflective practitioners.

**8.6. Misconceptions about action research:** There are certain misconceptions about action research and they can be discussed as follows.

(1) **Action research can improve others**: It is to be internalized amply, clearly, that a practitioner conducts Action Research in order to become a better professional. Therefore, the first beneficiary is the practitioner himself/ herself as a practitioner cannot function in isolation, the clients also benefit. Therefore it is not appropriate to say that action research is conducted to improve others is as the primary beneficiary is the professional himself/ herself.

(2) **Only school teachers can do Action Research:** It is a gross misconception that only a school teacher can conduct Action Research. All professionals can conduct action research. Since, a teacher is also a professional, who practices teaching, a teacher can also conduct action research.

(3) **Action research involves money in all cases:** It is not true that all action research requires money. There can be many action researches, which may not require any money at all. This is potentially so with teachers.

(4) **Action research is conducted based on the advice orders of higher up:** It is a gross misconception that action research is conducted based on the advice or orders of higher ups. It is the practitioner's choice and privilege. No higher ups can insist that his or her subordinates to do action research. He or she cannot utmost insist on quality and assurance and suggests that action research would be used as a tool but it cannot be insisted upon by the higher ups.

(5) **Action research can fetch a degree:** Action research is a reflective practice. It does not fetch any degree. It is only a professional management strategy of becoming more effective as a professional. It is only a strategy and a tool at the best. Therefore, it is not to be mistaken for any activity, which could fetch any degree.

(6) **Action research is a one-time affair. At best a couple of action researchers are sufficient in one's career:** It is totally erroneous to think that action research is a one-time affair. A practitioner can be a continuous innovator who believes in moving to higher levels constantly. As a practitioner does not undertake action research either to get a degree or any other benefit but for professional satisfaction he or she remains active and whenever, there is dissatisfaction one tries to use action research in converting the situation of dissatisfaction to a situation of satisfaction. Therefore, it is not the number of action researches that is important to a reflective practitioner.

(7) **Sophisticated tools and techniques are necessary in action research:** Unlike other researches, Action Research does not require sophisticated tools or techniques. An Action Research is a practitioner's research and a reflective practice. Depending upon the nature of action research one may need different tools. Generally they are not those sophisticated tools and techniques as a practitioner in education is not trained(add) professionally to

handle them. That does not mean that an action research is an activity that is to be understood apologetically. No, it is not so. It has a dignity of its own and it has a personality of its own. However, it does not require any sophisticated tools or techniques for conducting action research.

(8) **Rigorous training is a must in action research:** It is totally erroneous to say that action research requires rigorous training. To be able to conduct action research one has to understand the concept and the purpose of action research and for this is a simple orientation is sufficient and does not call for any rigorous training as a must to undertake action research.

To sum up, it is important to remember that action research facilitates a practitioner to become more professional. It is an activity that any sensitive practitioner undertakes to bring about qualitative improvement in one's own practice. This is also called a reflective practice and this is to be undertaken at one's own will.

## References


Coghlan, D. and Brannick, T. (2001). *Doing Action Research.* New Delhi : Sage Publications.

Daudelin, M.W. (1996). Learning from Experience through Reflection. *Organizational Dynamics,* 24 (3), 36-48.

Ebbutt, D. (1985). Educational Action Research : Some general concerns and specific quibbles : In Burgess, R. (Ed). *Issues in Educational Research.* Lewis: Falmer Press.

Elliot, J. (1991). *Action Research for Educational Change.* Milton Keynes : Open University Press.

Hopkins, D. (1998). *A Teacher's Guide to Classroom Research.* (Second Edition). Philadelphia : Open University Press.

Kolb, D. (1984). *Experiential Learning.* EaglewoodCliffs, NJ :Prentice Hall.

Kemmis, S. (1982). Action research in retrospect and prospect. *The Action Research Reader.* Deakin University : Victoria Deakin University Press.

Kemmis, S. (1983). Action Research. In, Husen, T. and Postlethwaite, T. (Eds). *International Encyclopedia of Education: Research and Studies.* Oxford : Pergamom.

Raelin, J.A. (2000. *Work-based Learning : The new frontier of management development.* Upper Saddle : NJ : Prentice Hall.

Rapaport, R. (1970). Three Dilemmas in Action Research. *Human Relations,* 23, 1-11.

Rigano, D. and Edwards, J. (1998). Incorporating Reflections into Work Practice: A case study. *Management Learning,* 29 (4), 431-446.

Ross, D., Bondy, E., and Kyle, D. (1993). *Reflective Teaching for Student Empowerment : Elementary Curriculum and Methods.* New York: MacMillan.

Stenhouse, L. (1984). Artistry and Teaching : The teacher as Focus of Research and Development. In, David Hopkins and Wideen, M. (Eds). *Alternative Perspectives on School Improvement.* Lewis : Falmer Press.

Stringer, E. (1996). *Action Research: A Handbook for Practitioners.* Thousand Oaks: Sage Publications.

Tripp, D. (1995). *Action Inquiry.* Paper posted on PARtoolbox@cornell.edu.

# Part B

# Educational Evaluation

# 9. Programme Evaluation

*Objectives*: At the end of reading this module the reader will be able to;

(a) understand the meaning of programme evaluation

(b) elaborate different steps involved in programme evaluation.

**9.1 Meaning of Programme Evaluation**: The programme is a systematic activity carried out by the institute or organisation which will be having certain specific goals. These goals are the major focus in the implimentation. It can be a small one or a big one. The programme evaluation can be process evaluation or product evaluation. Product evaluation is rather easy, as something that has been worked out after the implementation. But process evaluation is difficult as it is different from situation to situation.

A programme evaluation covers a specific programme as a whole, or selected aspects thereof. Most often, the aim is to provide an overall assessment of the programme, and programme evaluations typically encompass all components that influence programme quality. This includes components like the study environment and the organisational framework in which the programme operates. We can take up an example of an in-service programme organizing by an institute. The need for the programme is decided by the head of the institution in consultation with his/her colleagues or other members of the institute. The head of the institute can also arrive at the title of the programme based on the feed back given in the previous programmes or based on the need analysis done by the institute.

## 9.2 Steps in the Programme Evaluation:

1. Need Identification
2. Planning
3. Implementation
4. Reporting
5. Feed back

**1. Need Identification:** The need identification is very important step in the Programme implementation. This planning and organizing a programme can be done by the institute. This can be done based on field survey or discussion with the field functionaries or conducting a seminar of the focus area. This will enable the planners to know the field requirements. The success of any programme is depending on the field requirements.

**2. Planning:** The in-service programme has to be planned properly. This process will include the availability of physical resources and human resources. This includes the preparation of programme proposal based on the various resources of the institute, fixing

**B. Association with the ABL Programme**

1. Did you attend the ABL (or) any other activity based training programme in the previous years ?

Yes / No

If yes, give details.

| Sl. No. | Training Programme | Duration | Venue |
|---|---|---|---|
| | | | |

**C. Preparation for training programme**

1. Were you given sufficient time to enable you to participate in the key persons training programme?

Yes / No

If no, how many days of time do you require?

2. Were you aware of the nature of work that would be transacted at the training programme?

Yes / No

If yes, what preparations did you make before attending the programme? Please specify.
If no, what alternatives do you suggest?

**D. Transactions in the Programme**

Tick against your response.

1. How cooperative were the Resource Persons with you in the training programme?

- Extremely cooperative ( )
- Very cooperative ( )
- Fairly cooperative ( )
- Somewhat cooperative ( )
- Not at all ( )

2. To what extent were you given freedom to express your views in the programme?

- All the time ( )
- Some times ( )
- Often ( )
- Rarely ( )
- Not at all ( )

3. How appropriate was the approach used by the Resource Persons in dealing various activities under different school subjects?

- Highly ( )
- Very much ( )
- Fairly ( )
- Somewhat ( )
- Not at all ( )

4. How do you rate the competency of the Resource Persons in demonstrating ABL activities?

- Extremely competent ( )
- Very much competent ( )
- Fairly competent ( )
- Somewhat competent ( )
- Not at all competent ( )

5. How far were you benefited by the training programme to discharge your duties as a Resource Person at the field level programme?

- Highly ( )
- Very much ( )
- Fairly ( )
- Somewhat ( )
- Not at all ( )

6. Mention any three activities in different curricular areas which were handled most efficiently by the Resource Persons.

| Mathematics | Language | EVS |
|---|---|---|
| 1. | 1. | 1. |
| 2. | 2. | 2. |
| 3. | 3. | 3. |

7. Mention three activities which are not demonstrated efficiently by the Resource Persons in different curricular areas.

| Mathematics | Language | EVS |
|---|---|---|

| | | |
|---|---|---|
| 1.<br>2.<br>3. | 1.<br>2.<br>3. | 1.<br>2.<br>3. |

8. To what extent did you get guidance and assistance from the Resource Persons in the planning and demonstration of the activities?
   - Very great extent ( )
   - Great extent ( )
   - Some extent ( )
   - Very little ( )
   - Nil ( )

9. What other guidance and help do you require from the Resource Persons (BRTE) to enable you to conduct the district level programme successfully and purposefully.

   1.
   2.
   3.
   4.
   5.

**E. Print Material**

1. Did you receive the print materials sufficiently before the commencement of the field level training programme?

   Yes / No

2. Did you receive the print material sufficiently to distribute the teachers during the training programme?

   Yes / No

3. How far the activities in the package are self-explanatory?

   - Completely
   - Mostly
   - Quite a bit
   - Somewhat
   - Very little

4. List the activities demonstrated by you in the field level programme.

a.

b.

c.

5. Were the participants able to comprehend the contents of the modules?

Yes / Not sure / No

6. Were the suggested activities in the training package practicable?

Yes / Not sure / No

7. How many suggested activities were actually tried out in the field?

- All
- Most
- Some
- Very few
- None

8. Mention the reasons for the activities which were not tried out.

a.

b.

c.

9. List out the activities (on the order of priority) which generated greater participation in the training centre.

10. Name the activities which generated least participation.

11. Mention then activities which you have found to be difficult to organize.

12. Was there any facility to cyclostyle / Xerox and supply the material that was produced in the training programme?

Yes / No

**F. Audio-Visual Material**

1. Did you used AV material in the training programme?

Yes / No

2. Mention the situations where AV materials have been used.
   1.
   2.
   3.
3. List the AV materials which you used in the training programme. Also, indicate the sources from where you obtained.

| Name of the material | Source |
|---|---|
| 1. | |
| 2. | |
| 3. | |

4. How actively the other peer members participated in the proceedings of the programme?

   - Very much
   - Moderately
   - Neutral
   - Indifferent
   - Hostile

**G. Follow-up Programme**

1. Were you involved in the conduct of field level training programmes?

Yes / No

2. Were you consulted by SSA authorities for guidance in the various phases of the implementation?

Yes / No

If yes, mention the nature of help / guidance provided by you.

3. Is there a need for follow-up programmes?

Yes / No

If yes, what should be the nature of such programmes?

4. What support / facilities one needs to take up the follow-up work? Give details.

5. What measures do you suggest for the effective implementation of ABL programme at school level?

**H. Problems and Suggestions**

1. Mention the problems, if any, encountered in the organization of the training camps.

| Nature | Problems | Measures to be taken to tackle the problems |
|---|---|---|
| a. Academic<br>b. Administrative<br>c. Financial<br>d. Any other (specify) | | |

**Sum up:**

# 10. Project Evaluation

## Objectives

At the end of this session, you will be able to:

- Define the terms project evaluation, types of project evaluation and its importance.
- Describe the ways in which project evaluation is relevant to project revision.

## 10.1 Meaning of Evaluation and Project

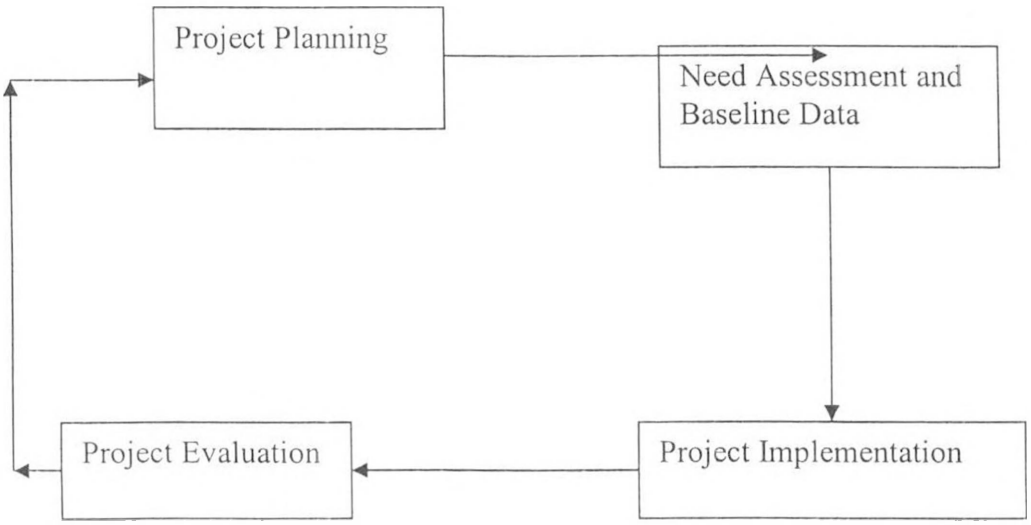
Evaluation has a wider meaning. Evaluation refers to assigning a value judgment to a measured attribute. A project is a social activity carried out in a group setting. The activity can be connected with all the school subjects at the various levels of school education.

***A project follows steps like:***

1. Identification of the project idea.
2. Planning
3. Execution of the project.
4. Reporting

## 10.2 Process of Project Evaluation:

Figure1: The project evaluation cycle



## 10.3 Purpose of Project Evaluation:

1. Evaluation provides information to help to improve a project

2. Evaluation provides information for communicating to a variety of stakeholders of the project.

3. Evaluation helps to know the impact and influence of the project on the target group of the project.

## 10.4 Types of Project Evaluation:

### *1. Formative Evaluation:*

Formative evaluation begins during project development and continues throughout the span of the project. Its purpose is to assess ongoing project activities and provide information to monitor and improve the project. It is done at several points in the process of a project and its implementation. According to Bob Stake, "When the cook tastes the soup, that's formative; When the guests taste the soup, that's summative." Formative evaluation has two components: **(i) implementation evaluation (ii) progress evaluation**.

**(i) Implementation Evaluation**: The purpose of implementation evaluation is to assess whether the project is being conducted as planned. This type of evaluation, sometimes called "process evaluation," which may occur once or several times during the span of the project. The underlying principle is that before you can evaluate the outcomes or impact of a program, you must make sure the program and its components are really operating and, if they are operating according to the proposed plan or description. A series of implementation questions guides an implementation evaluation. For example, questions that might be posed for that are as follows:

- Was appropriate target group selected?

- Was target group is identified in advance to provide the transitional support they needed?

- Do the interventional strategies and activities are as per the plan?

- Was target group given both academic and personal supports?

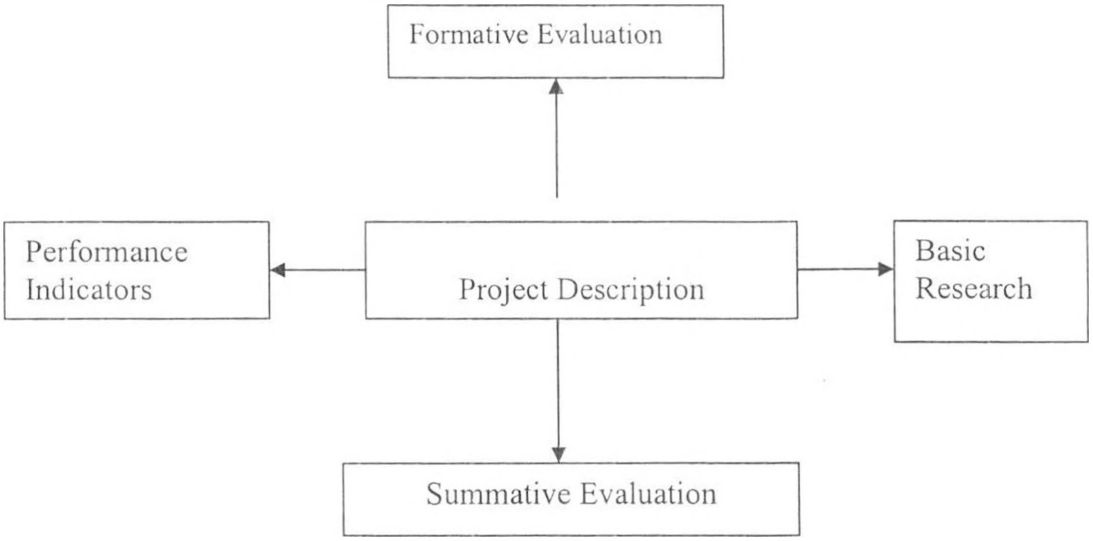
- To what extent the interventional strategies and activities are implemented to provide meaningful opportunities to conduct research to know the impact of the interventions provided?

- Was a solid project management plan developed and followed?

Sometimes the terms "implementation evaluation" and "monitoring evaluation" are confused.. They are not the same. An implementation evaluation is an early check by the project staff, or the evaluator, to see if all essential elements are in place and operating.

for the project evaluation is to be collected across the project in a consistent and systematic manner. In addition, some programmes have to be added to program-specific modules aimed at collecting tailored data elements.

**Figure2:—Types of data gathering activities**




Formative and summative evaluations are intended to gather information to answer a limited number of questions. Evaluations include descriptive information, but go well beyond that. Generally, formative and summative evaluations include more in depth data collection activities, are intended to support decision making, and are more costly. Performance indicators fall between general program statistics and formative/summative evaluation.

A performance indicator system is a collection of statistics that can be used to monitor the ongoing status of a program against a set of targets and metrics. Going beyond descriptive statistics, performance indicators begin to provide information that can be measured against a set of goals and objectives. Indicator systems are typically used to focus policymakers, educators, and the public on (1) key aspects of how an educational project is operating, (2) whether project is being made, and (3) where there are problems (Blank, 1993). Because performance indicators focus on tangible results, they often go beyond traditional reviews of project expenditures and activity levels. In fact, the term "performance" underscores the underlying purpose of indicator systems, i.e., to examine a project's accomplishments and measure progress toward specific Project Description Basic Research goals.

Hence, **Performance indicators** provide a snapshot of accomplishments in selected areas; however, in contrast to evaluations, the information is limited and is unlikely to provide an explanation of why a project may have succeeded or failed. Research studies include descriptive information and provide targeted in depth exploration of issues, but differ along other dimensions. Instead of being intended for

decision-making, research efforts typically are designed to explore conceptual models and alternative explanations for observed relationships.

## References


1. Annie W. Ward and Mildred Murray Ward. (1999). *Assessment in the Classroom,* London: Warsworth Publishing Company.

2. R.L.Thorndike (Ed.) (1971). *Educational Measurement.* Washington DC: American Council on Education.

3. Daniel D. Roman. *Managing Projects – A systems Approach*

4. http://www.nio.org/uploads/bec.pdf

5. http://www.rocw.raifoundation.org/computing/BCA/softwareprojectmanagement/lecture-notes/lecture-24.pdf

6. Janice M. Morse. (1994). Critical Issues in Qualitative Research Methods. SAGE.

7. Oya Lcmeli Tukel et.al, (2001). *An Empirical Evaluation of Project – Criteria.* London: MCB UB Limited.

# PART C

# MONITORING OF EDUCATIONAL ACTIVITIES

# 11. Monitoring Of Educational Activities

**Objectives:** At the end of reading this module, a reader will be able to understand:

(a) the meaning of the term monitoring
(b) why monitoring is needed
(c) when monitoring is to be done
(d) how to monitor an educational programme
(e) documentation related issues of a monitoring activity
(f) the role and importance of follow up of monitoring activity

## A: Monitoring Activities:

### 11.1 What is Monitoring?

Monitoring means checking to see if we are reaching our goals, and if not, making some changes to come up with a better plan. Monitoring is generally defined as ***the act of checking, overseeing or tracking the progress of an event to see whether the intended goals are achieved and what are the corrective steps necessary to make midcourse corrections if the goals are not achieved as per the plan.*** In case of development projects, monitoring represents an on-going activity to track project progress against planned tasks. It aims at providing regular overview of the implementation of an activity in terms of input delivery, work schedules, targeted outputs, etc.

Often the term monitoring is invariably used with another term, evaluation and is generally clubbed together as two interrelated activities of monitoring & evaluation (M&E). M&E are increasingly recognized as indispensable management functions, which are closely related, interactive and mutually supportive. Through routine tracking of project progress, monitoring can provide quantitative and qualitative data useful for designing and implementing project evaluation exercises. On the other hand, evaluations support project monitoring. UNICEF provides a comparative perspective of M&E wherein one can see the commonalities as well as distinctions between the same. The following table indicates the same.

| Item/Attribute | Monitoring | Evaluation |
|---|---|---|
| Frequency | Periodic, regular | Episodic |
| Main action | Keeping track/oversight | Assessment |
| Basic purpose | Improving efficiency<br>Adjusting work plan | Improve effectiveness,<br>impact, future |

| | | programming |
|---|---|---|
| Focus | Inputs/outputs, process outcomes, work plans | Effectiveness, relevance, impact, cost-effectiveness |
| Information sources | Routine systems, field observations, progress reports, rapid assessments | Same plus surveys/studies |
| Undertaken by | Project managers, Community workers, Community beneficiaries, Supervisors, Funders | Program managers, Supervisors, Funders External evaluators Community (beneficiaries) |

Source: UNICEF, A UNICEF Guide for Monitoring and Evaluation: Making a Difference? New York, 1991

In the field of education, monitoring of educational activities; be it learning, teaching or managing is considered as a very important function of educational management. Even here the monitoring is done to see whether a particular activity is progressing as per plan and what the gaps are between expected and observed and what necessary midcourse corrections are necessary for effective progress and outcomes.

Let us take a simple example of *monitoring children's attendance in schools*. Here the parameter identified is *attendance* and there are two reference points. One is the *children, whose* attendance needs to be monitored and the other one is the *school*. In this case, monitoring can be done in several ways. It can be done on daily basis or weekly basis or monthly basis or quarterly basis or half yearly basis or even on annual basis. Therefore there is a *periodicity* dimension attached to the monitoring activity. Further, attendance can be monitored for a particular grade or class, or for a particular subject, or for a particular group of children or even for particular category of schools. Hence there is the dimension of *situation or condition* against which monitoring can be undertaken. Both these dimensions are determined by the purpose or goal of monitoring. Hence monitoring is purpose or goal driven. That is monitoring always results in some kind of outcomes, which can help us to assess the progress of a given activity at a given point of time.

It may also be noted that the monitoring can be done with reference to both quantitative parameters and qualitative parameters. In the former, the parameters would reflect certain numerical indicators such as enrolment, attendance, examination results, teachers, PTR, institutions, expenditure etc. In the latter, the parameters might reflect certain process indicators such as the quality of implementation in terms of adequacy, fullness, timeliness, adherence to the given norms, physical, material and human resource inputs either facilitating or impeding the progress etc,.

## 11.2 Why Monitoring?

Project monitoring is normally carried out by project management, staff and other stakeholders. Effective monitoring needs adequate planning, baseline data, indicators of performance, and results and practical implementation mechanisms that include actions such as field visits, stakeholder meetings, documentation of project activities, regular reporting, etc.

In case of development projects, monitoring aims at:

(i) Providing project management, staff and other stakeholders with information on, whether progress is being made towards achieving project objectives. In this regard, monitoring represents a continuous assessment of project implementation in relation to project plans, resources, infrastructure, and use of services by project beneficiaries.

(ii) Providing regular feedback to enhance the ongoing learning experience and to improve the planning process and effectiveness of interventions.

(iii) Increasing project accountability with donors and other stakeholders;

(iv) Enabling managers and staff to identify and reinforce initial positive project results, strengths and successes. It also alerts managers to actual and potential project weaknesses, problems and shortcomings before it is too late. This would provide managers with the opportunity to make timely adjustments and corrective actions to improve the program/project design, work plan and implementation strategies.

(v) Checking on conditions or situations of a target group, and changes brought about by project activities. In this regard, monitoring assists project management to check whether the project continues to be relevant to the target group and/or geographical area, and whether project assumptions are still valid.

## 11.3 Why monitor educational activities?

There are different reasons offered for carrying out monitoring work. These reasons assume importance in different contexts depending on the stage of educational development. Primarily, the most important reason anywhere for monitoring is to find out how the education system is performing so that remedial action can be taken where performance is below the expected level. Particularly in less developed countries, where achievement in education is not satisfactory in terms of outcomes, quality of teaching and learning, poor participation of children in schools, the need to establish a sound system of monitoring is assuming crucial significance (Livingstone 1985).

It is noticed that several countries, which are on the path of educational reforms are adopting national developmental plans. For such countries, it is becoming increasingly important to produce baseline information on levels of achievement, against which the effectiveness of various national development plans can be measured at sometime in the future. Hence regular monitoring of such plans is becoming increasingly important for tracking progress on different parameters at frequent intervals of time.

An important feature of all national developmental plans is to improve the quality and relevance of education. In this context, reforms in school curriculum have remained important concern. Constant up gradation and updating of curriculum require objective data on achievement, so that curriculum developers, advisors, and teachers can plan for improvements in the quality and relevance of the education offered to students. Hence monitoring helps in gathering empirical data with respect to outcomes at different points of time so that appropriate actions can be initiated wherever there is shortfall in expected outcome.

With rapid expansion of education systems, the countries are faced with the demand to move from elitist form of education to universal systems wherein the need to equalize educational opportunities is becoming extremely apparent. Under the circumstance, the national educational authorities wish to know the achievement levels which exist in the different regions of the country, between urban-rural, male-female and different social groups in performance and the way these are related to the allocations of resources, both human and financial. Therefore periodic and regular monitoring across these dimension would help the planners and policymakers to assess the benefits and impact of the investments made for several interventions in education, be it for certain geographic regions or for particular target clientele.

As education continues to claim proportionately a larger share of national budgets in many developed countries along with health and social welfare, some form of national monitoring of standards preferably by an independent body is seen to be legitimate and necessary. Besides, governance reforms such as voucher systems, school based management, school based evaluation in the context of educational decentralization also demands a well endowed monitoring system in place to inform the education authorities from time to time the outcomes and accomplishment of intended goals and objectives. Particularly, where administration of the education system is decentralized and supported through resources raised by local taxpayer's money, there is all the more pressing need to be more accountable to those who support the system through the taxes. The demand for accountability becomes more apparent in times of national economic constraints and increasingly education system comes under closer scrutiny. As such, it becomes obligatory for the education system to allay public anxieties and convince the tax payers that students are receiving the best education worth the money spent. Therefore a good monitoring system is essential for the same.

Within the framework of educational reform for raising standards and improvements in quality of education service delivered, constant monitoring of national assessment is becoming a regular feature across all countries. Besides, complaints from

employers and the general public about alleged declines in standards in the quality of school education have also raised the need to put a constant vigil and continuous monitoring system in place. Reforms in education governance through rigorous accountability mechanisms, public-private partnership, increased transparency, right to information acts, stringent regulatory mechanisms for ensuring satisfactory standards have all triggered public conscience. In addition, curricular reforms in terms of pedagogy, evaluation and teaching learning materials and resources, incorporation of ICT into teaching, learning and management of education delivery, demand greater professionalism among teachers. With raising societal aspirations for good quality education and positive affirmative action policies, there is an increased call from the public for more objective ways of measuring the output of school systems. It is in this context that monitoring can inform the public from time to time about the quality of education delivered and the ultimate outcomes.

Even in India, we notice that over the years, the elementary education system has grown substantially. Along with this, the physical, material and other academic facilities have also seen substantial enhancement in schools. However, the outcomes in elementary education with respect to enrolment, retention and learning attainments are not commensurate with quantitative growth. Low retention, poor completion rates and poor quality of education in large number of schools have remained the major concerns and challenges for achieving the goal of Universalisation of Elementary Education. The global monitoring agencies such as the UNDP & UNICEF have been painting gloomy picture about primary education performance in India in terms of attendance and achievement despite huge resource support to primary education. This is worrisome.

In addition, the survey of educational outcomes conducted periodically by the national education agencies such as the National Council of Educational Research & Training (NCERT), National University of Educational Planning & Administration (NUEPA), National Family Health Survey (NFHS) and National Sample Survey organizations (NSSO) have been consistently pointing to poor attendance rates, low completion rates, particularly among the vulnerable population segments including girls in primary schools. More importantly in recent years, surveys by independent agencies such as the PROBE (1998) and Annual Survey of Educational Reports (2005, 2006 and 2007) have been releasing report cards about learning attainments of children studying in government primary schools, which have been causing public uproar. Government primary schools and teachers in them are coming under increased public scanning and civil society organizations and general public are beginning to demand greater accountability from teachers in government schools.

In order to address the above challenges, the Indian government has launched the national flagship programme of Sarva Shiksha Abhiyana Mission in 2001 to bring all children in the age group 6-14 years into the fold of elementary education by 2010 with a focus on quality and equity concerns. Huge budgetary support has been provided to every district in the country on annual basis to start new schools, appoint new teachers and also to enhance quality of teaching and learning in schools. A number of activities have been planned and implemented at field level in this direction. Hence the progress of these

activities requires to be monitored from time to time not only to assess the optimal utilization of the resources but also to periodically effect midcourse corrective steps, if any and help the system to perform more effectively in realizing the goal of UEE.

India is a signatory to the World Education Forum and has committed itself to providing education for all. The UN Millennium Development Goal requires that India has to achieve universal primary education by 2010. Besides India has its own constitutional commitment to provide elementary education to every child after having declared elementary education as a fundamental right in 2002. The international donor agencies which are providing financial support to elementary education have made it mandatory to establish regular monitoring systems to assess timely progress made by individual districts across the country. Thus, within these mandatory frameworks and deadlines, it is becoming increasingly important to undertake regular monitoring of the educational activities. More importantly, education system is becoming increasingly accountable to several agencies and actors who have a stake in the same. Therefore having a monitoring system in place and checking the progress periodically enhances accountability of the system thereby increasing the efficiency of education delivery.

Monitoring activity helps us to know the level of progress achieved on different set of parameters, at different points of time, in different school contexts, across different geographical areas. The monitoring will also help us to identify bottlenecks or impediments if any which come in the way of smooth implementation of activities in the field and accordingly help us to take necessary corrective steps so that in the next course of action smooth progress can be ensured.

To elaborate on the above, let us take the same example of monitoring attendance in schools. On checking attendance in schools during the first quarter of the school year we discover that in certain schools and in certain classes, certain category of children are regularly absenting from school. Once we discover this it is possible for us to identify which type of schools, which classes and which groups of children reveal this problem. Accordingly we can further probe and find out the reasons for absenteeism. Having identified reasons for absenteeism, necessary measures can be initiated by the concerned school to address this problem. In some cases the teacher/school may be in a position to address this problem by talking to parents and persuading them to send children to schools. In some other cases, it may require effort of other teachers, head teacher, community representatives and even local PRI members. With these measures, the schools may be in a position to improve attendance. Subsequent monitoring in the second quarter will help us to find out what gains have been recorded and what is the kind of improvement made and which strategies have yielded the desired results. Based on the results, further course of action to enhance progress may be chalked out.

Thus periodic monitoring will not only help us to benchmark achievement with respect to several parameters in diverse situations, it will also help us to set the achievable targets, estimate the resources required in terms of men and materials, the kind of capacities needed to achieve the desired targets etc,.

## 11.4 When should monitoring be done?

Monitoring cannot be done abruptly in an ad-hoc manner. Generally, it is done at regular intervals to track the progress and make mid course corrections. Depending on the need and resources available, it can be done on quarterly basis or half yearly basis or on an annual basis. However for achieving best results in education, it is desirable that first quarterly monitoring is very essential as it will indicate the gaps that need to be bridged and also it would provide good opportunity for altering our strategies, if need be, in order to reach the end goal as indicated in the plan. It may be noted that formal education has a built in time frame and schedule of learning activities in order to progress in a sequential pattern. Hence progress at every stage is contingent upon the inputs received in the previous stage. In order to know whether adequate inputs are provided as required, monitoring becomes very essential.

There are certain activities, which are to be monitored quite frequently, such as the attendance of pupils, teachers, functioning of schools, mid day meals, learning outcomes etc. Certain other things need to be monitored in the beginning of the school year itself such as the supply of text books, uniforms, teacher deployment, teacher training, teaching-learning equipment and aids, all of which are essential elements for conducting regular school activities. Monitoring being part of the evaluation of education needs to be done at regular intervals so that the expected end goals can be achieved more effectively.

## 11.5 How to monitor?

Monitoring cannot be done in an arbitrary manner. It has to be done in a very systematic and methodical manner. Hence developing a good monitoring plan is the first and foremost task. Project managers usually prepare annual work plans that translate the project document into concrete tasks. The work plans will describe in detail the delivery of inputs, the activities to be conducted and the expected results. They will also clearly indicate schedules and the persons responsible for providing the inputs and producing results. The work plans are being used as the basis for monitoring the progress of program/project implementation.

The monitoring plan should indicate the resources needed to carry out project monitoring. Needed funds and staff time should be allocated to ensure effective implementation. There are certain norms which need to be kept in mind for conducting monitoring activities. A good monitoring system will have a clear laid down format and procedure. It will also identify the instruments which need to be used in order to gather data and evidences for checking the progress.

A UNICEF (1991) publication summarizes the plan for evaluating an activity. The same can be applied to the monitoring plan also. They are as follows:

- **Why** - The purposes of the evaluation - who can/will use the results.

- **When** - The timing of evaluation in the program cycle.
- **What** - The scope and focus of evaluation and questions for the evaluation to answer.
- **Who** - Those responsible for managing and those responsible for carrying out the evaluation, specifying whether the evaluation team will be internal or external or a combination of both.
- **How** - The methods of gathering data to answer the questions.
- **Resources** - The supplies and materials, infrastructure and logistics needed for the evaluation.

As a management tool, monitoring should be organized at each level of management. Monitoring systems should be linked to annual plans. A first step in designing a monitoring plan is to identify *who* needs what information, *for what purpose,* how *frequently,* and in *what form.*

To develop an effective monitoring system, the following steps might be followed:

1. A first step towards developing a good monitoring system is to decide what should be monitored. The careful selection of monitoring indicators organizes and focuses the data collection process.
2. The next question would be how to gather information, i.e. to select methods to track indicators and report on progress (observation, interviews, stakeholder meetings, routine reporting through proforma filling up, field visits, etc.).
3. When to gather information and by whom. The monitoring plan should include who will gather the information and how often. Project staff at various levels will do most data collection, analysis and reporting. In case where monitoring is done by educational authorities in the system, it becomes necessary to identify at what level, with what periodicity, in what form and who will collect information. At this point it becomes necessary to have clear laid out procedures.
4. It is also necessary to review the progress reports either by the project staff in case of development projects or by education planners and administrators and major stakeholders. Collecting regular feedback by the project managers or others concerned with monitoring on a regular basis is most desirable.

Monitoring can be done both on small scale and on large scale. Small scale monitoring will essentially be done at individual school level either by the teacher or head teacher concerned. This may include details relating to specific areas which reveal problem so that corrective steps can be initiated at the school level itself. However, in

case of large scale monitoring, the sample coverage may include a larger units of children, schools, teachers etc, in order to obtain a much broader picture about the education situation in a given area, say a cluster, block or district. This requires resources both in terms of men and materials for purpose of collecting, collating, tabulating and consolidating the data with reference to particular set of indicators. In both these cases, if the monitoring has to be more effective, lot of preparatory activities are required as discussed below.

In the first place, the parameters that need to be monitored have to be clearly stated. *That is what is to be monitored?*

For instance if we are required to monitor attendance in schools, we need to be clear about the area, school, class and students, whose attendance need to be monitored.

Let us discuss each one of them. Area refers to the geographical boundary within which monitoring needs to be undertaken. That is, whether monitoring of attendance needs to be done at cluster level, block level or district level or state, national level has to be clearly defined. This should also indicate which cluster, which block, which district, which state wherein monitoring will be undertaken.

School may refer to the stage/level or category/type. When we say level or stage, we need to clearly state, whether monitoring of attendance will be done at the primary level/stage (I-V) or higher primary level/stage (VI-VIII) or total elementary level/stage (I-VIII). By category we mean the type of schools either by management, location or other criteria. We need to specify in which type of schools, the attendance will be monitored - government, private aided, private unaided or schools in urban areas, rural areas, slums, hilly or tribal belts or in single sex schools or in mixed schools, minority schools etc,.

For monitoring children's attendance, we may consider various criteria such as the age, sex, class, caste (S.T; S.T; B.T; BCM), religion, and other economic and social factors.

In the next step, it is to be specified *how attendance is to be monitored.* Whether the information relating to this should be collected on daily or weekly or monthly or quarterly or half-yearly or annual basis. While actual percentage can be computed in case of daily attendance rates, in case of other durations, the average may be worked out per week, per month, per quarter, per half-year, per annum etc,. This can be done separately for different category of children as well.

This is so far as quantitative parameters are considered. In case where we need to monitor certain qualitative indicators, say for example, we are interested in knowing the quality of text book supply. Some of the indicators which may be identified for this are; timeliness, adequacy, quality of books etc. whether text books are supplied on time, to all children, for all subjects and whether good quality books have been supplied could be

ascertained by direct observation as well as gathering views of children, teachers, parents & other stakeholders such as SDMC, CAC, GPs etc,.

# B: Documentation of Monitoring Work

Having gathered different set of data after monitoring a particular activity, it is necessary that the same are analysed and reported in readable formats so that various stakeholders can use it with ease and efficiency. Before we discuss about the reporting format a few points with respect to data organization, tabulation and analyses are necessary.

## 11.6 Data Organization

Organizing monitoring data is an important step for ensuring effective analysis and reporting. It can be done both manually and/or using computer, depending on the quantity of data. Although, small scale monitoring data can be organized manually, yet for purpose of computation and sophisticated analysis, it is desirable that the data are entered into a computer programme on an excel spreadsheet so that it lends itself for any kind of cross tabulation and statistical analysis. Besides, data entry on computer also helps in storage of data as well as in periodic updating so that it is easy to track progress on observable parameters. However, in order to do this certain technical expertise and skill are necessary. Most importantly, after data entry, it is necessary to recheck the data for ensuring validity and reliability of the same.

## 11.7 Tabulation and Analysis of Data

Tabulation and analysis of data gathered through monitoring will require a comprehensive data analysis plan. The plan should be well aligned with the goals and objectives of monitoring. For instance, if our intention is to identify strengths and weaknesses of a particular educational intervention, we can organize data into program strengths, weaknesses and suggestions to improve the program. Data analysis often involves the desegregation of data into categories to provide evidence about achievements of certain programmes and to identify areas in which a program is succeeding and/or needs improvement. Data can be broken down by gender, social and economic situation, class, school type, location (urban or rural), age, etc. However the choice of desegregation form will depend upon the objectives and type of indicators used.

Generally use of both quantitative and qualitative analysis has become the preferred model for many of the monitoring activities. Qualitative data includes detailed descriptions, direct statements in response to open-ended questions, analysis of case studies, the transcript of opinion of groups, and observations of different types. Qualitative analysis is best done in conjunction with the statistical analysis of related quantitative data. Analysis of qualitative methods may produce descriptions (patterns, themes, tendencies, trends, etc.), and interpretations and explanations of these patterns. The data analysis should include efforts to assess the reliability and validity of findings.

## 11.8 Format of Monitoring Report

There is no prescribed format for reporting monitoring activity. However, there are certain ground rules that need to be followed so that the report becomes a useful tool for taking further action by the authorities concerned. The following formats may be useful for documenting the monitoring activity.

**1. Title page/ Cover page**

The title page should clearly indicate what is being monitored, who has done the monitoring and the year in which the same was done.

**2. Content page**

This page would indicate the titles of contents chapter wise with sections or sub-titles in the report and the corresponding page numbers. Titles of Annexure and appendices or any enclosures should also figure in this list.

A separate list of tables/graphs can also be added to the content list.

**3. Preface**

This page would give brief introduction about the monitoring activity undertaken by the authors, terms of reference, acknowledgement to various agencies and people who helped in conducting the activity and the expected outcomes of the same.

**4. Body of the Report**

The body of the report should contain the following:

**i. Introduction:** which will inform the reader about the context, objectives and the framework (normative/criterion) with which the monitoring has been undertaken

**ii. Methodology:** will provide a brief description about the methods employed for gathering data, data types, source, instruments used, sampling frame, size and design and analysis mode.

**iii. Data presentation and analysis:** This chapter will describe the outcome of the empirical data gathered. The sequencing of this can be done objective wise so that there is logical coherence and appropriate sequencing of each monitoring component.

**iv. Summary & Conclusion:** In this chapter, it is necessary to highlight the outcomes of monitoring with respect to each component. The major findings can be presented in crisp tabular formats, which would reveal the strengths and weaknesses along with corrective actions to be initiated for each of the component monitored.

**A few Tips for Preparation of Monitoring Report:**

***a. Make the report short and concise***

It is indeed a great challenge for organizing the huge quantity of data gathered into a useful, concise and interesting report. The author should be skilful enough to sift the relevant and useful data from not so relevant, while making presentation. It is useful to remember that only a small and concise amount of tabulations prepared during the analysis phase should be reported. It is always better to abide by key questions, the indicators assessed and the type of information that the reader may be looking for.

***b. State recommendations clearly***

It is necessary to make recommendations clear, concise and direct by indicating where the programme has failed, where capacity building needed, what kind of actions needed to improve performance etc,.

***c. Make the presentation interesting***

Presentation should be clear and adjusted to the target clientele. It should be made in simple language that can be understood even by lay persons. The structure of the report should be simple. The text should be broken down in relatively small thematic or sequential parts, with simple and clear subtitles precisely identifying the topics discussed.

Given below is a suggestive format given by UNICEF for writing an Evaluation report. This can also serve as a monitoring report format.

| **Suggested Contents of Evaluation Report** |
|---|
| *1. Title page* |
| *2. Table of Contents* |

*3. Acknowledgments (optional)*

- Identify those who contributed to the evaluation.

*4. Executive Summary*

- Summarize the program/project evaluated, the purpose of the evaluation and the methods used, the major findings, and the recommendations in priority order.
- Two to three pages (usually) that could be read independently without reference to the rest of the report.

*5. Introduction*

- Identify program/project description/background.
- Describe the program/project being evaluated (the setting and problem addressed, objectives and strategies, funding).
- Summarize the evaluation context (purposes, sponsors, composition of the team, duration).

*6. Evaluation Objectives and Methodology*

- List the evaluation objectives (the questions the evaluation was designed to answer).
- Describe fully the evaluation methods and instruments (e.g., what data were collected, specific methods used to gather and analyze them, rationale for visiting selected sites).
- Limitations of the evaluation.

*7. Findings and Conclusions*

- State findings clearly with data presented graphically in tables and figures. Include effects of the findings on achievement of program/project goals.

- Explain the comparisons made to judge whether adequate progress was made.
- Identify reasons for accomplishments and failures, especially continuing constraints.

*8. Recommendations*

- List the recommendations for different kinds of users in priority order. Include costs of implementing them, when possible.
- Link recommendations explicitly with the findings, discussing their implication for decision-makers.
- Include a proposed timetable for implementing/reviewing recommendations.

*9. Lessons Learned (optional)*

- Identify lessons learned from this evaluation for those planning, implementing or evaluating similar activities.

10. *Appendices*

- Terms of Reference.
- Instruments used to collect data/information (copies of questionnaires, surveys, etc.).
- List of persons interviewed and sites visited.
- Data collection instruments.
- Case studies.
- Abbreviations.
- Any related literature.
- Other data/ tables not included in the findings chapter.

Source: *UNICEF,* A UNICEF Guide for Monitoring and Evaluation: Making a Difference? *New York. 1991.*

## C: Follow up of Monitoring Reports

It is very necessary that the monitoring reports are followed up regularly. An important factor is that it needs to be done immediately once the outcomes are made known to the concerned authorities. Sharing the report with concerned field level authorities who are implementing the project and engaging in face to face dialogue to sort out minor issues which can be dealt at the local level itself would greatly benefit the authorities concerned. Issuing circulars by the superior executive authorities to the local level field officers for removing impediments would facilitate accomplishment of programme objectives. Often time lag will render the monitoring report useless, if it is not immediately followed up. Hence, the report should clearly state what action needs to be initiated and who should do the same.

In case of development projects, the project staff themselves monitor the implementation of the project and make suitable changes based on the outcome and feed back of the Project staff and leader. However, when an external agency monitors an educational activity and there are multiple stakeholders (donors, national sponsors, regional authority, sub-state level authorities who implement the programmes in the field) are interested in obtaining feedback, it becomes rather complicated to coordinate the follow up between and among multiple agencies. Hence it is necessary that proper timeframe is maintained in release of money, issue of guidelines so that immediate actions can be initiated.

Immediate decision making and streamlining the procedural hurdles can be done at the local level by the immediate authorities. There should be built in flexibility to allow for some minor alterations for speedy and effective implementation of the programme.

Monitoring can also throw up more complicated policy related issues, which defy easy solutions. It often demands systematic reforms to change the deeply entrenched practices and traditions, which throw up major problems to bring about improvements. In such an event more systematic research needs to be undertaken to identify critical factors and processes, which can facilitate reform initiatives.

## References


IRC International Water and Sanitation Centre, 2004, **Monitoring Millennium Developments Goals** – A Review of Experiences and Challenges

Livingstone, I.D. 1985. "Standards, National: Monitoring" In: **The International Encyclopedia of Education.** (Eds). Postlethwaite T.N. & Ebel, The Pergamon Press, NY, pp. 4786-4791.

Summer R (ed) 1977. "Monitoring National Standards of Attainments in Schools", **National Foundation for Educational Research**, Slough.

The UN WWAP, 2003. **World Water Development Report.**

UNICEF, 1991-**A UNICEF Guide for Monitoring and Evaluation**: Making a Difference? New York, p.3

UNICEF, 2003 (Revised) -Programme Policy & Procedures Manual: Programme Operations, UNICEF, New York, pp. 109-120

World Bank. 1974-81. **Assessment Studies by Education Department Staff**, WB, Washington (mimeo).