11/28/22

# INTRODUCTION TO RESEARCH

By: Mr. Wanyama

# MODULE OBJECTIVE

- To enable the learner to acquire knowledge and skills in research

# Module units

- Concept of research process
- Application of research in health care delivery system

# Module learning outcomes

By the end of the learning sessions, the learner should be able to:

- Demonstrate understanding of the concept of research and its application in nursing
- Explain the research process
- Demonstrate knowledge in writing a research proposal
- Demonstrate knowledge in basic research statistics
- Participate in carrying out research and publicize findings in local, regional and international journals

# COURSE OUTLINE

- Introduction: Definition of Research
- Concept of Research and Purpose of Research
- Types of Research
- The Research process
- Research Designs
- Proposal writing
- Sampling process and sample size determination
- Basic statistics
- Research publications
- Citation and Referencing using the APA style

# REFERENCE MATERIALS

Mugenda, A & Mugenda O. (2003). *Research Methods*: Nairobi, Kenya.

Bassavanthapa, B. (1998). *Nursing Research*: Jaypee Brothers.

# INTRODUCTION: SCIENTIFIC DEFINITION OF RESEARCH

- According to Burns and Grove (1999), the word '**research**' means '**to search again**' or 'to examine carefully'.
- Research is diligent, systematic inquiry or study to validate and refine existing knowledge and develop new knowledge.

**OR**

- A process of arriving at an effective solution to problems through systematic collection, analysis and interpretation of data.
- **Nursing research** is defined as research into those aspects of professional activity, which are predominantly and appropriately the concern and responsibility of nurses.

# Nursing research cont’

- Research can be nursing research only when it addresses issues that are relevant to the nursing profession.
- If the topic is not going to have an impact on the nursing practice, it is not nursing research. Thus, it is important to understand that in order to define your research as nursing research; your study must be geared towards improving certain aspects of the nursing practice, education and administration.

# Meaning of research

- Carrying out a diligent inquiry or a critical examination of a given phenomenon. It implies an exhaustive study, investigation or examination following some logical sequence
- Involves a critical analysis of an existing conclusion or theories with regard to newly discovered facts due to ever changing world with advances in technology continually creating new possibilities. It means a continued search for new knowledge and understanding of the world around us.

# Purposes of research

- To discover new knowledge. This involves discovering of new facts, their correct interpretation and practical application
- To describe a phenomen. Accurate identification of any event which involves thorough description, e.g. size, shape, age, weight, colour, height , change over time and others. Description then provides knowledge of the basis of other research purposes

# Cont' purpose

- To enable prediction. This is the ability to estimate phenomenon A given phenomenon B.
- To enable control.

✓In scientific research control is concerned with the ability to regulate the phenomenon under study.

✓Usually one phenomen is manipulated in order to exert control over another.

✓The control and predictive functions of science are closely related.

# Cont’d purpose

- ✓ In predictive inquiry, the researcher is interested in studying naturally occurring associations between phenomenon to estimate its impact on another phenomenon.
- ✓ This is experimental research and leads to more powerful statements about associations compared to predictive inquiry.
- ✓ But the nature of some phenomenon may not permit experimental research, only predictive inquiry is possible. e.g. studying socio-class structures, race, mob behaviour, and personality cannot manipulate these phenomen hence must be content with predictive inquiry.

# Cont’d purpose

- To enable explanation of phenomena.

✓This involves accurate observation and measurement of a given phenomenon.

✓In order to explain a phenomenon one should be able to describe it, predict its occurrence and observe factors that cause its occurrence with certainty and accuracy.

# cont'd purpose

- To enable theory development.

✓This involves formulating concepts, laws and generalizations about a given phenomenon.

✓Research is also conducted in an attempt to confirm or validate existing theories.

# Sources of knowledge

- Research itself.

✓ Is an important source of knowledge because it is objective and involves systematic procedures.

- Experience is a common mode of obtaining knowledge.

✓ May be being present when an event happens which becomes a primary source.

✓ Can be direct personal experience, logical observation or evidence theory

# Cont’d sources of knowledge

- Tradition.

✓ All human beings inherit a culture. Culture is a reflection of an adopted system of rules, norms, standards and values.

✓ The socio-cultural system embodies accepted knowledge of how things are or should be.

✓ Acquisition of this knowledge for any member of the society is through indoctrination and socialization.

✓ Transmitters of cultural knowledge in the society are parents, peers, adults, teachers, and clergy

# Cont’d source of knowledge

- Authority.

✓This takes the form of an expert in specialized area, giving his/her opinion on a given issue.

✓The acceptance of such an opinion by others depends on the status of the person giving the opinion.

- Intuition.

✓Is the perception or explanation or insight into phenomena by instinct.

✓Or the ability to gain knew knowledge without conscious reasoning or rational process.

# Characteristics of research

- **Systematic**.

Follows specific steps and process, i. e. problem identification, sampling, data collection, interpretation and analysis, evaluation and reporting.

- **Logical**.

Procedures used should flow and link in a sensible manner that can allow evaluation. Scientific concepts must be employed.

- **Empirical**.

It relies on observation and measurement devoid of external influence or personal position

# Cont' characteristics

- **Perceptibility (understandability).**

Scientific explanations must be perceptible. People must be able to understand them if they are to be accepted and included in the body of knowledge

- **Limitations.**

Limits in the process must be acknowledged as protected and provided by research ethics e.g. financial constraints, lack of cooperation from subjects, time constraints etc

- **Replicable.**

The process must recorded to enable others to test/verify findings by the repeating the research or builds on it in future

# CHARACTERISTICS OF GOOD RESEARCH

- Makes a comprehensive statement of the problem and justification of the research problem
- Has clearly stated objectives and research questions
- Measures what it sought to from the onset. If for any reasons this is not possible, the researcher offers plausible explanations for the anomaly
- Has a clearly stated and relevant purpose. For example, if the research focuses on community health, then it must address health areas relevant to the target community and population

- Utilizes an intensive and extensive review of relevant literature to broaden the perspective of the research project
- Clearly and systematically explains the methodology utilized
- Offers plausible interpretations and explanations of the results
- Draws valid evidence-based conclusions deduced from the research results
- Generates plausible evidence based on the scientific methods applied
- Examines the implications of the evidence adduced with other relevant important areas of the target community, such as social, economic, health, policy and other aspects

# Answers that a good research should provide

- What are the perceived health needs of different groups of people?
- To what extent do the present health interactions cover the health priority needs of the people?
- Are the interventions that have been initiated acceptable to the people in terms of culture and cost (especially the poor)? Are these interventions provided as cost-effectively as possible?
- Is it possible to cover more needs given the available resources? Is it possible to expand cost-sharing through insurance? Would this help reduce most of the unexpected high costs?

- Is it possible to better control environmental factors which influence health and health care?
- Can other factors such as education, agricultural, public works or others, be of help?

# Basic terms in research

- **Population**

✓ Population is defined as 'the entire group of persons or set of objects and events the researcher wants to study' (Rensburg 2000:147).

✓ It is also referred to as all the possible entities or individuals that have the characteristic(s) of interest to the study.

✓ The most important aspect is that the population must posses all the characteristics the researcher is interested in.

# Cont' terms

- **Sample**

✓ A sample is a group of people, or records or a number of observations from a larger population.

✓ It is a representative group of individuals selected from a population.

✓ A sample aids the researcher to get access to the general population.

✓ A sample that is selected appropriately generates data that reflects the true status of the population in relation to the characteristic(s) or variables under study

# Cont’d terms

- **Sampling**. The process of selecting a number of individuals for a study in such a way that the individuals selected represent the large group from which they were selected is the population. The purpose of sampling is to secure a representative group which will enable the researcher to gain information about a population.
- **Variable**. Is a measurable characteristic that assumes different values among the subjects. Some variables are attributes

# Cont’ terms

- **Problem identification**. It is a process of generating a well articulated statement or question of a problem under study. It produces a problem statement that is logical with structure, sequence, substance and rationale
- **Hypothesis**. It is a tentative, testable statement of the relationship between two or more variables. An informed/learned guess that indicates what expectations from the study.

- **DATA**. Refers to all the information that a researcher gathers for his/her study. There are two types of data: primary data and secondary data. Primary data-the information that researcher obtains from the field i.e. from the subjects in the sample. Secondary data- the information that a researcher obtains from research articles, books, casual interviews e.t.c.
- **PARAMETER-** A characteristic that is measurable and can assume different values in the population

- **Statistics-** The science of organizing, describing and analyzing quantitative data
- **Descriptive statistics**: are indices that describe a given sample. Examples of descriptive statistics are: measures of central tendency (mean, mode and median), measures of dispersion (Range, standard deviation, variance), Distributions (percentages, frequencies) and relationships (correlation)
- **Inferential statistics:** A branch of statistics which researchers use to draw inferences about a given phenomenon in the population. They are used to test Hypothesis.

# THE RESEARCH PROCESS

- Has 5 phases;
  1. The conceptual phase, also called the thinking or planning phase= **selection & formulation of the research problem.**
  2. Design and planning phase= **choice of research design**
     - ❑**Research format and research design**
     - ❑**Description of the samples**
     - ❑**Sampling procedures**
  3. The empirical phase, also called the doing phase= **data collection**
     - ❑ **construction of the research instrument**
     - ❑**Actual data gathering**
  4. The interpretive phase (analytic), or the phase where the researcher looks at the meaning of it all= **analyzing & interpretation of data/results**
     - ❑**Data processing**
     - ❑**Statistical analysis**
  5. The communication phase or the phase of writing and disseminating the research report= **conclusion, recommendations & writing the research report**

**The 5 phases have been expanded to the following 10 steps**

# THE RESEARCH PROCESS

1. Identify the research topic
2. Problem statement
3. Rationale/ justification/purpose of the study
4. Formulate research objectives, questions & hypothesis CHAPTER 1
5. Literature review>>>> CHAPTER 2
6. Research methodology
7. Describe the methods of measurement >>>CHAPTER 3
8. Data collection and presentation>>> CHAPTER 4
9. Data analysis, Discussion & interpretation>>>> CHAPTER 4
10. Communicating research findings, conclusion, recommendations, limitations, report writing & appendices >>> CHAPTER 5

# Steps involved in selection & formulation of the Research problem

1. Selecting a topic for Research: identification of the problem
2. Formulating the research problem
3. Acquiring knowledge on current theories and researches: literature review
4. Identifying and labelling the variables
5. Defining concepts and establishing operational definitions
6. Formulating the hypothesis

# 1. IDENTIFICATION OF RESEARCH TOPIC

- This is the first step in developing community health research,
- It provides the focus of the study
- It becomes the title of your Report

**The characteristics of a good research topic are:**

- ❖It should be clear and concise.
- ❖Should be appealing i.e. attracting to the reader
- ❖It should not contain more than 20 words **(including full stops and comas**)
- ❖You can have a title that contains two parts. However, in this case the two parts have to be separated by a colon (:)

# Example of a title consisting of two parts

**FACTORS INFLUENCING IMPLEMENTATION OF THE HIV/AIDS EDUCATION CURRICULUM IN PRIMARY SCHOOLS: A CASE STUDY OF THARAKA NITHI COUNTY.**

# CRITERIA FOR TITLE SELECTION

- Should be of your ***own choice***
- One that provides a clear focus for the study you want to undertake

- A researcher may draw research topics from:
  - Existing professional knowledge and experience.
  - Socially significant issues i.e., the research topic has practical relevance and significance to the society. For instance, the study could be in a position to solve an existing and pressing social problem.
  - A study that has scientific relevance and significance as well. It is important to show that the research will, at the end, have academic and/or scientific relevance.

# Criteria for a researchable topic

The following criteria enables the researcher to determine whether a problem is suitable for research. These include:

- **Frequency of occurrence of the problem** - If the nature of the problem is such that it rarely occurs, then research is not advisable for that problem.
- **The degree of discomfort caused by the problem** – The problem may be causing harm to people, or even may have led to harm or risk to life, in which case it is expedient that research could lead to a solution.
- **Amount of resources affected by the problem** – a consideration is given to the nature of the problem, and whether a lot of financial or material resources are lost or jeopardized because of the problem e.g. is the non-availability of resources affecting retention of trained nurses in the Hospital?

# Criteria for a researchable topic ctd’

- **Researchability of the problem** – What to consider is to establish that the problem actually merits research and the entire community affected.
- **Genuine interest of the researcher** - We should recognize that interest spurs people on for research; the researcher must be fully interested for any meaningful engagement in the research. Without interest, the commitment in the research would be poor, and of non effect.

# 2. PROBLEM STATEMENT

**INTRODUCTION**

- The next step after topic selection is the statement of the problem.
- This provides the context for your research
- It defines clearly the problem you propose to examine
- It explains why it is important to carry out the research
- A brief review of literature will be required in delineating and defining the research problem

# DEFINITION OF PROBLEM STATEMENT

- The problem statement is a specific statement that clearly conveys the **scope, magnitude and purpose** of the research study.

**Reasons why it is Important to State and Define the Problem Well**

- Is the foundation for further development of the research proposal (research objectives, methodology, work plan, budget)
- Makes it easier to find information and reports of similar studies from which your own study design can benefit

❖Enables you to systematically point out:

- why the proposed research on the problem should be undertaken and
- what you hope to achieve with the study result. This is especially important when you have to present your project to community members, health staff, relevant ministries and donor agencies that need to support your study or give their consent
- It important when you have to present your project to community members, health staff, relevant ministries and donor agencies that need to support your study or give their consent

# Chronological order of problem statement

i) Identify the problem situation

ii) This will be followed by the process of problem definition.

iii) The identified problem must now be defined in terms of its:

- Occurrence,
- Intensity,
- Distribution, and other measures for which data are already available

- The aim is to determine all that is currently known about the problem and the reason it exists

iv) Concerted efforts must be made to establish

- **How** wide spread is the problem?
- **Who** is affected by the problem?
- **What** is its distribution?
- How often does the problem occur?
- What social or cultural practices are associated with the problem?
- What costs are associated with the problem?

v) A thorough literature review would assist in determining the following:

- Incidence and prevalence of the problem;
- Geographic areas affected by the problem;
- Population affected by the problem;
- Probable reasons for the problem;
- Possible solutions;
- Unanswered questions that need to be researched.

# SUMMARY OF A WELL STATED PROBLEM

A well-defined research problem statement leads naturally to the statement of:

- Research objectives,
- Hypotheses;
- Definition of key variables, and
- Selection of a methodology for measuring the variables

**NOTE:** A poorly defined research problem leads to confusion

# Procedure for Identifying and Defining a Research Problem: Step 1

- Problem situation: Start with a simple statement of the problem situation. Write a small simple paragraph that identifies the problem.
- Discrepancy: State what the discrepancy is between what is and what should be.
- Problem Question: Write down the central problem question.
- Possible answers: Write two or more plausible answers to the problem question

# STEP 2

- Add details as you review the literature, review theoretical concepts and investigate the problem in greater depth. Try to answer the following questions:
  - What is the incidence and prevalence of the problem?
  - What geographic areas are affected by the problem?
  - Which population groups are affected by the problem?
  - How was the problem studied in the past?
  - What are the findings from other research studies?
  - What has been done to overcome the problem?
  - What seems to be the major unanswered questions about the problem?

# STEP 3

- Simplify the focus by identifying the most important aspects of the problem that are researchable.

## STEP 4

- Let one or more colleagues review your final statement identifying and defining the problem. Revise your statement if necessary in the light of the comments and suggestions received.

# Information to be Included in the Statement of the Problem

- A brief description of the socio-economic and cultural characteristics and an overview of the health status and health care system in the study area in as far as these are relevant to the research problem. Include a few illustrative statistics, if available to help describe the context in which the problem occurs

- A concise description of the nature of the problem (the discrepancy between what it is and what it should be) and its magnitude, distribution, and severity (who is affected, where, since when, and what are the consequences for those affected and for the services?)
- An analysis of the major factors that may influence the problem and a convincing argument that available information and knowledge is not sufficient to solve the problem.
- A brief description of any solutions that have been tried in the past, how well they have worked and why the further research is needed.

- A description of the type of information expected to result from the research and how this information will be used to help solve the problem

## CHARACTERISTICS OF A GOOD PROBLEM STATEMENT

- It is written clearly and attracts the reader's interest immediately
- It has identified a specific problem which is researchable
- It spells out the scope of the research problem
- The importance of the study in adding new knowledge is stated clearly.

# STEP 3: RATIONALE/JUSTIFICATION OF THE RESEARCH PROBLEM

- After stating the research problem is to justify why you chose to study this problem.
- The words ‘rationale’ and ‘justification’ are commonly used interchangeably.
- This is the section of the study that outlines reasons for carrying out the study.

- Justifications of the study should address some of the following questions( **QUESTIONS THAT YOU MUST ADDRESS):**

❖Is the problem you wish to study a **current and timely** one? Does it exist now?

❖How widespread is the problem? Are many areas and many people affected?

❖Does the problem affect key populations such as the adolescents, youth, expectant mothers or children?

- Does the problem relate to ongoing intervention activities?
- Does the problem relate to broad social, economic and health issues such as poverty, status of women, or education?
- Who else is concerned about the problem? Are top government officials concerned? Are health and other professionals concerned?
- What gaps in knowledge do you want to fill in and why is it important to generate information to fill those gaps?

- It is important to state the justification convincingly so as to rationalize the utilization of resources such as time, money materials and manpower.
- The rationale of the study should describe the utility and importance of the problem in health care services in general and the nursing profession in particular.

# AREAS YOU NEED TO JUSTIFY

- You will also need to justify the location or site(s) in which to conduct your research.
- The location or site must be described in some detail, paying attention to its appropriateness to the research proposed.

# STEP 4: Formulation of Research questions, Objectives and Hypothesis

**What are RESEARCH QUESTIONS?**

- These are questions whose answers should help us come up with a solution to the problem stated.
- Answers to these questions will be obtained through the research you will conduct

**HOW TO FORMULATE RESEARCH QUESTIONS**

- The research question (s) should develop logically from the problem and context you have defined under the introduction and problem statement section and
- should describe the core objective of your research.
- These are the most important questions whose answers you are seeking to find through your research

Research question assists the researcher to:

a) Focus on the study by narrowing it down to the essentials
b) Avoid collection of data that are not necessary
c) Organize the study in clearly defined parts or phases

## Good research questions should be **"FINER":-**

**F** - Feasible, allowing one to appreciate the practical limitations.

**I** - Interesting, sustaining the research process.

**N** - Novel, able to provide new findings.

**E** - Ethical.

**R** - Relevant, advancing science or influencing clinical care, health care policy among others.

- **Example of a FINER Research Question**
  - How do nurses in Kalala hospital practice the hand washing procedure as stipulated by the hospital infection control handbook?
- **Example of a non FINER Research Question**
  - Are nurses washing hands?

# Example of a research question

- An example of a research question might be: **"What is the relationship between Allergic conjunctivitis and age"?**

# RESEARCH OBJECTIVES

**Def:** A **research objective** is a clear, concise, declarative, statement expressed to direct a study. It focuses on identification and description of variables and/or determination of the relationships among variables.

- Objectives describe the expected results arising from the study
- Usually **broad and specific objectives** are stated

## ***Importance of research objectives***

- Research objectives help to:
  a) Bridge the gap between the research purpose and the study design.
  b) Guide on planning for data collection and analysis.

c) Summarize what is to be achieved by the study.

d) Build a close link with the statement of the problem.

e) Keep the researcher within the scope of study by defining the area of focus.

- Research objectives are sub-divided into broad and specific objectives. When formulating good research objectives, the objectives should have the following characteristics, using the acronym ‘SMART’

- **S** - Specific; clearly identifies the item at hand for investigation.
  **M** - Measurable; being quantifiable
  **A** - Achievable; acquire the set objectives
  **R** - Realistic
  **T** - Time bound; in form of human, financial and material resources
- **Example of a SMART Objective**
  - To establish the number of children born at home within the last two years in Ganga village
- **Example of a Non SMART Objective**
  - To find out the level of home deliveries.

# CATEGORIES OF RESEARCH OBJECTIVES

- Broad
- Specific

**Broad Objective**

- Describes the expected contributions arising from the study.
- Relate the reasonable and expected contributions of the study to broad social, economic, or health concerns
- Contribute to the justification of why research on the problem is required

**NOTE**

- Note that broad objectives are the expected contributions.
- The investigator does not promise that the contribution will occur and therefore, usually does not try to measure them.

# Guidelines for writing research objectives

The writing of the broad objectives will be guided by the following questions:

- How will the results from the study help improve service delivery, improve training programs, or assist in the design of educational materials?
- In other words, what are the anticipated contributions of the study

# EXAMPLES OF BROAD OBJECTIVES

- The broad objective of this study is to contribute towards increasing utilization of RH services among Kenyan adolescents.
- The broad objective of this study is to contribute towards reduction of the prevalence of Malaria among the study population.
- The broad objective of this study is to contribute to a better understanding of the factors that affect the use of maternal health care services in the study area

# SPECIFIC OBJECTIVES

- In contrast to broad objectives that state what is expected to happen, specific objectives relate directly to the research problem situation.
- These are the outputs or deliverables of the study for which the researcher is responsible
- They indicate the variables that will be examined and measured.
- An immediate objective represents a promise by the investigator that certain specific variables will be examined
- Specific objectives are expressed in **measurable terms**

# EXAMPLES OF SPECIFIC OBJECTIVES

- To establish the influence of education on the use of treated mosquito nets in the study area.
- To establish the association between the attitude of health workers and client satisfaction in the study area.
- To determine the effect of public health education campaign on the uptake of modern maternal health care services.
- To establish the effect of staff training on the quality of care provided and client satisfaction

# HYPOTHESIS OF THE STUDY

**Description:**

- This is a **predictive** statement about an expected relationship between two or more variables that permits empirical testing
- It is the **researcher's prediction** or explanation of the relationships between two variables. E.g.

  -Persons with Type II diabetes mellitus who have greater knowledge of their disease will have a higher rate of adherence to treatment regimen than those with less knowledge.

# Difference between Specific objectives, Broad objectives and hypothesis

- While broad objectives identify the **anticipated contributions** arising from a study,
- and specific objectives **specify what will be done or measured** in the study,
- hypotheses specify the **expected relationship** among the variables

# Where are hypothesis required?

- They are most appropriate for field intervention or evaluative studies.
- Diagnostic or exploratory studies **do not normally require hypotheses because they generally do not test relationships between variables.**

# Purposes of Hypotheses

- Hypotheses provide direction. They bridge the gap btwn the problem and the evidence needed for its solution.
- Hypotheses ensure collection of the evidence necessary to answer the question posed in the problem statement.
- It enables the researcher to assess the information he or she has collected from the standpoint of both relevance and organization.
- Sensitizes the investigator to certain aspects of the situation that are relevant regarding the problem at hand.
- Enables the researcher to understand the problem with greater clarity.

# The characteristics of good hypotheses

- ✓ Must state clearly and briefly the expected relationship between variables.
- ✓ Based on a sound rationale derived from theory, or research, or professional experience
- ✓ Consistent with common sense or generally accepted truths.

# Characteristics cont'd

- ✓ Testable, i.e. data can be collected to support or fail to support the hypotheses. This then implies that the variables stated in the hypotheses can be operationalized.
- ✓ Must be related to empirical phenomena. Words like 'ought', 'bad' reflect moral judgment should be avoided
- ✓ Be testable within a reasonable time.

# Characteristics cont'd

- ✓ Variables tested in the hypotheses must be consisted with purpose statement, objectives and the operationalized variables in the method section
- ✓ A good hypothesis must be as simple and as concise as the complexity of the concepts involved allows.
- ✓ A good hypothesis must be stated in such a way that its implications can be deduced in the form of empirical operations with respect to which relationship can be validated or refuted.

# TYPES OF HYPOTHESES

1. **Directional Hypothesis:** it predicts an outcome in a specific direction.

Example;

*Persons with Type II diabetes mellitus and have greater knowledge of their diseases will have a higher rate of adherence to treatment regimen than those with less knowledge.*

1. **Non Directional Hypothesis:** it indicates there is a difference or correlation but does not specify which. For example:

   *Persons with Type II diabetes mellitus who follow a structured programme on their condition have a higher rate of adherence to treatment.*

- This does not indicate a directional relationship.

## 3. Null Hypothesis and Alternative Hypothesis

1) Null hypothesis (denoted as $H_0$): The null (statistical) hypothesis is used for statistical testing and interpretation. It states no difference exists between groups or no correlations between variables. Example;

*There is no difference in performance of national examinations between standard eight pupils from rural primary schools and standard eight students from urban primary school in Kenya*.

4. **Alternative hypothesis** (denoted as $H_1$): It states that there is a difference or correlation.

# Considerations in Writing the Hypothesis

- In writing study hypotheses, always think in terms of the expected relationship between variables

**HOW TO WRITE A HYPOTHESIS**

- Think first about the central problem your study will address (the dependent variable).
- Next, consider what factor or factors (the independent variables) might cause, determine, or influence the dependent variable
- Finally, ask yourself if the relationship between the independent and dependent variables is direct or indirect through a set of intervening variables.

## VARIABLES

**Def:** Variables are defined as quality, properties or characteristics of persons, things or situation that change or vary/ assume different values. For example: sex (male and female) age (20–25, 26–30 years) academic success, stress and pain.

### *Types of variables*

1) **Independent Variable(treatment/experimental) I**t's a variable that influences other variables.

   It is perceived as contributing to or enabling a particular outcome.

- Independent variables usually describe what the researcher **wishes to measure** in order to determine its effect on an observed phenomenon (the dependent variable)

- It is the **intervention** or treatment that the researcher performs to see the resulting change in the dependant variable. It is also referred to as the input.

**2) Dependent Variable:** This is the **outcome** variable. It reflects the effects (outcome) or response to the independent variable.

It is the variable that appears, disappears, diminishes or increases.

it describes the **problem under study**

**For example**, to determine the effects of salt intake on **hypertension**, the blood pressure is the dependant variable and salt intake is the independent variable.

3) **Extraneous Variables:** These are uncontrolled variables that influence the findings of the research study. They include **intervening, antecedent, suppressor**, and **distorter variables**.

They influence both the dependent and independent variables. These are called threats to internal and external validity of the study and may bias the selection, the time factor, and the instrument used.

4) **Demographic Variables** :These are demographic attributes. They are variables that cannot be manipulated or influenced by the researcher, for example, age, sex religious beliefs or educational level

**5) Control variables**: If a researcher suspects that a certain variable is likely to influence the research results, he or she should control for that variable (the extraneous) in the study.

- When an extraneous variable is built into the study, it is referred to as a **control variable**. Some researchers refer to control variable as **concomitant, covariate**, or **blocking variables**.
- The introduction of a control variable in research study increases the validity of the data and therefore it leads to more convincing generalizations.

- **For example**, if gender may also influence reaction time, we should add sex as an independent variable in our study. Using a statistical procedure such as regression, we can measure the effect of alcohol on reaction time, controlling for sex.

**6) Intervening variable:** The logical status of an intervening variable is that it is recognized as being caused by the independent variable and as being a determinant of the dependent variable. i.e.

Independent variable → intervening variable → dependent variable

- An intervening variable comes in between the independent and the dependent variables.

- When intervening variable are used as control variables, one must establish the dominant direction of influence.
- The independent variable influences the intervening variable and the intervening variable influences the dependent variable.

# STEP 5: LITERATURE REVIEW

**DESCRIPTION:** is a summary of theoretical and empirical sources to generate a picture of what is known and not known about a particular problem.

- This is a systematic review of existing information (literature) about a specific subject or topic.
- This is sometimes referred to as desk or library research.
- The information could be published such as journal articles and books or unpublished such as research reports or monographs.
- This review is usually undertaken by the researcher himself or herself or by a team.

**Literature review entails;**

- The history of the problem
- The magnitude & distribution of the problem, including the population being affected
- The severity of the problem
- Methodology (ies) used in the previous studies
- Theoretical and analytical frameworks used
- Hypotheses and variables used and their measurements
- Research Designs used
- Methods of data collection used
- Sampling procedures and sample sizes used
- Methods of data analysis used

## LITERURE REVIEW ENTAILS CNTD’

- Main findings of the previous researches or studies
- Main conclusions and recommendations of previous studies
- Past efforts to solve the problem

## THE MAIN PURPOSES OF LITERATURE REVIEW ARE TO:

- Determine what has been done already as regards the research problem under investigation.
- Identify gaps existing in the area of interest
- Identify strategies, procedures and measuring instruments that have been found useful in the investigation of the research problem.
- Help make the researcher familiar with previous studies and thus facilitate the interpretation of the study.
- Help the researcher to narrow the research topic.
- Establish whether it's feasible to conduct research in that area or if it's duplication
- Help determine new approaches and stimulate new ideas

When reviewing any literature, the researcher needs to:

1. Assess the strengths and weaknesses of past work on the subject
2. Report any inconsistent findings
3. Identify gaps in the knowledge
4. Determine the contribution of proposed study
5. Consider the possibility of unintentional duplication

**Steps in carrying a Literature Review**

- Be familiar with the library before beginning the literature review.
- Make a list of key words or phrases to guide the review
- With the key words/ phrases go to the library. Lib staff are always ready to help

- With the key words/ phrases go to the computer internet and do a search
- Summarize references on cards for easy organization
- Analyze, organize and report in an orderly manner
- Make an outline of main topics/ themes, headlines and sub-titles
- Studies with contrary views should not be ignored. Attempt to account for the differences in opinion
- The more general literature should be covered first before narrowing to what is more specific to the research problem –this paves way for identifying testable hypotheses

- 2 major sources of information for literature review:
  - Primary and Secondary sources.

**Primary Source.** A direct description of any occurrence by an individual who actually observed or witnessed the occurrence. This is the work written by the person who is actually involved in, or is responsible for, the generation of the idea published.

- It's the work of the original author
- The person who conducts empirical research and publishes it in a journal is usually regarded as the primary source of information

**Secondary Source** :involves any publication written by an author who was not a direct observer or participant in the events described

- This type of work is usually a paraphrase of the primary source. Often, it does not give the correct interpretation of the primary sources.
- It is usually information given by someone who was not a direct observer or participant of the events described.

## Sources of information

1. Scholarly Journals
2. Theses and Dissertations: research projects written by Masters and PhD students.
3. Govt. Documents; i.e. policy papers, research reports owned by the govt., annual reports of hospitals and government ministries

4. Papers Presented at Conferences
5. Books
6. The internet: Computers also have databases prepared for literature search. Examples of such databases are MEDLINE or INDEX MEDICUS, Pub Med

## TIPS ON GOOD REVIEWING OF LITERATURE

- Do not conduct a hurried literature review. You are likely to overlook important studies.
- Do not rely too heavily on secondary sources. Try to get some primary information from experts, opinion leaders, peers e.t.c.
- Do not concentrate only on findings. Read the methodology and measurement of variables also

# TIPS ON GOOD REVIEWING OF LITERATURE cntd’

- Check daily newspapers –educative and current information
- Copy references correctly so as to avoid frustration of trying to retrace a reference later

# Referencing within text

- There are two common methods of referencing within a text. This refers to accrediting a statement or finding to another author, to show that the statement quoted is by that author. The statement may be paraphrased, so need not to be in quotation marks.
- Method 1: Author's last name & year of document's publication are put **after** a paraphrased statement in a text. The name and year are put in brackets
  - *E.g. Among the economic factors affecting satisfaction with quality of life, income has been found to be positively related to satisfaction with quality of life (Berry and Williams, 2013).*

- Method two: Sometimes the author's name and the publication year come at the beginning of a sentence e.g:
  - *Berry and Williams (2013) found a positive relationship between income and satisfaction with quality of life.*

# PLAGIARISM

- “The substantial use, without acknowledgement and with intend to deceive the examiners or knowing that the examiners might be deceived, representing, whether by copying or paraphrase, the ideas or discoveries of another or of others as one's own work submitted for assessment.’’
- The mere inclusion of the source in a bibliography shall not be considered sufficient acknowledgement”

# HOW TO AVOID PLAGIARISM

- Summarize- In your own words and cite reference
- Paraphrase- Translate source into your own words and cite reference
- Quotation- Word-for-word and cite reference (using quotation marks- “……”)

# STEP 6: RESEARCH METHODOLOGY

**Definition**

- This describes the procedures that are followed in conducting research where techniques of obtaining data are developed to test hypotheses
- The steps involved in conducting the study should be described in detail
- As part of the research methodology, the researcher has to describe the research design.

- Every researcher has to identify an appropriate research design for use in the study. Identification of a specific study design will depend on :
  - Available information (state of knowledge) about the problem.
  - The nature of the problem and its environment.
  - The availability of resources for the study.
  - The skills and creativity of the researchers.

# Types of research designs

- Experimental Design: True Experimental and Quasi-experimental.
- Survey Design: Comparative and Correlation.
- Descriptive Design: Descriptive and Explorative.
- Case Study Design

1. **Experimental Research Design**

   Commonly used in clinical settings because of its accuracy and reliability.

It could be subdivided into true experimental and quasi-experimental.

- Experimental research designs are concerned with testing hypotheses and establishing causality.
- This design tests the hypothesis of relationships, that is, attempts to make predictions of future outcomes based on a causal model implementing strategies to control the predicted outcome. If 'X' occurs 'Y' follows and so on.
- In an experimental research design, the researcher controls or manipulates the action of the independent or causal variable(s) and observes and measures the action or outcome on the dependant variable. For example, *the effectiveness of a particular drug such as paracetamol in relieving moderate pain.*

## ***Characteristics of experimental research design***

**Manipulation:** the researcher controls the independent variable, which can be an event, intervention or treatment that is expected to have some effect on the dependant variable.

**Control :**The researcher exercises control over the experimental situation by eliminating the actions of other possible variables beyond the independent variable.

**Randomizing :** Every subject is given an equal chance to participate in the study. The researcher assigns the subjects to the experimental or control groups on a random basis.

- In a quasi-experimental design, some of the above rules are relaxed. For example, there might be no need for having a control group or at times randomization may not be included. The quasi-experimental design enables the search for knowledge and examination of causality in situations where complete control is not possible.

**Advantages of experimental research design**

1. Most powerful design for testing the hypothesis of cause-effect relationships between variables.
2. It is practical, feasible and can be generalized to some extent. This type of design introduces some control over certain extraneous variables.

**Disadvantages of experimental research design**

1. In most real situations, it is difficult to conduct a true experimental design, since some of the variables cannot be manipulated or controlled.
2. At times it becomes quite difficult to get randomized research subjects or even a control group.
3. As a result of the need for randomization, control and manipulation with the aim of establishing cause-effect relationships, the design becomes very expensive, both in terms of time and money.

# 2. Survey Research Design

- Is the systematic gathering of information.
- Survey studies are concerned with gathering information **from a sample of population.**
- The purpose of the study is usually to identify general trends or patterns in the collected data.
- It's designed to obtain information from the population regarding the prevalence, distribution, and interrelations of variables within those populations.
- Survey studies primarily yield **quantitative data**. They are mainly cross sectional in design, but can also be longitudinal.

- They mainly deal with (investigate) what people do, for example, how or what they eat, how they meet their health needs, what kind of family planning behaviour they engage in and so on.

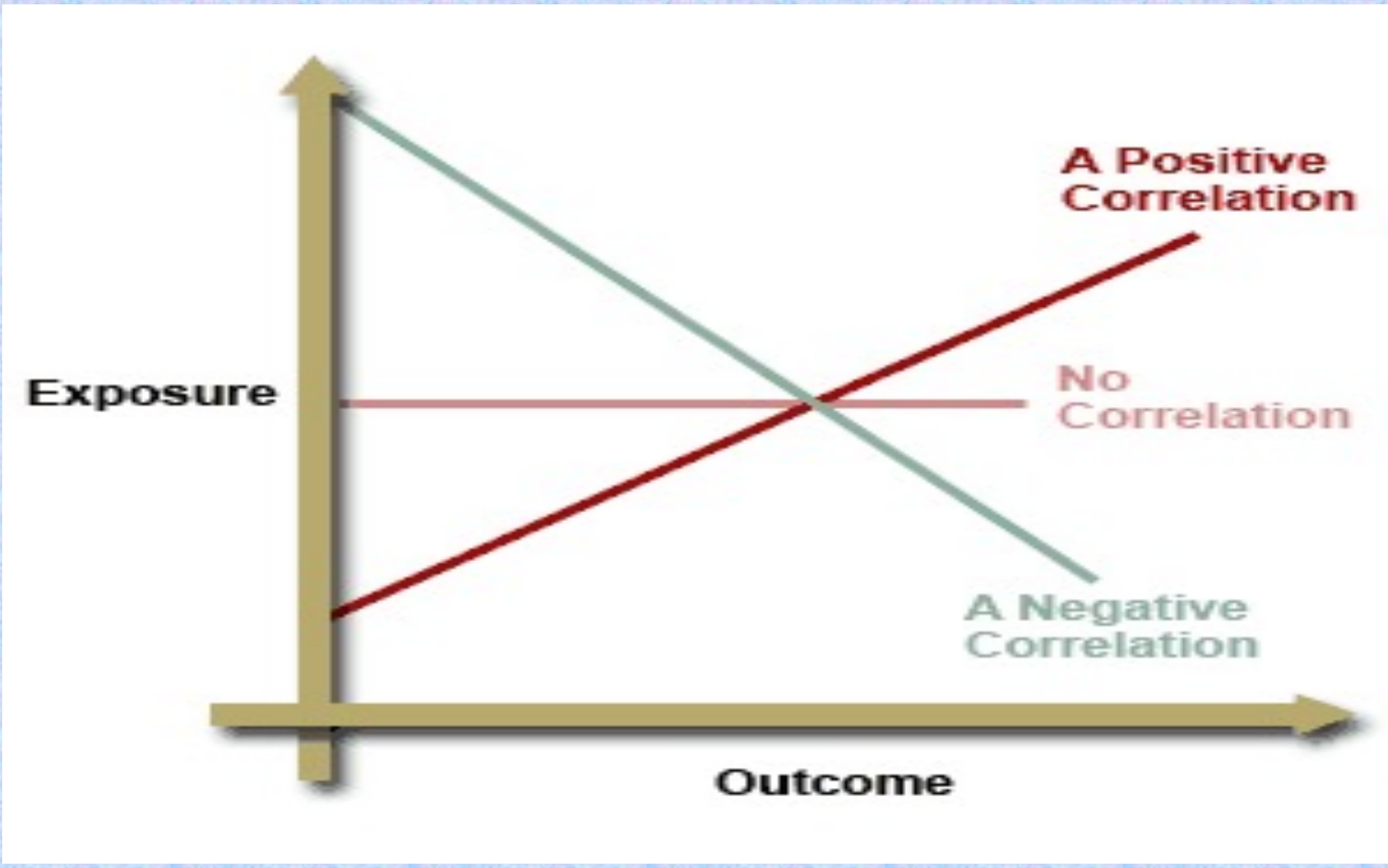
***Advantages***

1. It is flexible and broad in scope.
2. It can be applied to many people
3. It can focus on wide range of topics

The survey design is better suited for extensive rather than intensive analysis of a situation.

It is usually descriptive and specific based on the situation that needs intervention, for planning purposes, monitoring and evaluation of services.

- In a survey the researcher designs the phenomenon and study but does not manipulate any variables nor do they make any efforts to determine the relationships between variables.
- In the **correlation survey** meanwhile, the researcher attempts to determine and describe what relationship exists between variables.
- One independent variable is correlated with one or more dependent variables. Then statistical methods are applied to describe if the variables relate at all and what kind of relationship they have, that is, positive correlation or, negative correlation


A Positive
Correlation
Exposure
No
Correlation
A Negative
Correlation
Outcome

- When there is a positive correlation is an indication that the more the exposure the high the outcome of interest, for example smoking exposure and lung disease, which is the outcome.
- When there is no correlation one would conclude that the exposure is not related to the outcome, for example teething and diarrhoeal episodes.
- When there is a negative correlation it means that the more the exposure the less the outcome, for example tetanus vaccination and tetanus infection.

## 3. Descriptive or Explorative Research Design

- Is the systematic collection and presentation of data to give clear picture of a particular situation.
- It involves the systematic collection of information and aims to discover and describe new facts about a situation, people, activities, or events.
- Its main purposes include observing, describing and documenting all aspects of a situation **as it naturally occurs, without any external manipulation**.
- At times, descriptive designs are used as a starting point for hypothesis generation or theory testing.

- In the **exploratory descriptive design**, the main purpose is to explore the dimensions of a phenomenon (problem) as well as the major characteristics or facts that influence the phenomenon.
- In descriptive design, no manipulation of variables is involved as opposed to experimental design.
- Similarly, no dependent or independent variables are used because no attempt is made to establish causality.
- The overall purpose of descriptive research is to provide a picture of a phenomenon as it naturally occurs, as opposed to studying the impacts of the phenomenon or intervention.

# Categories of descriptive research design

1. **Explorative Descriptive Design:** the researcher explores a particular problem to discover what is there and if it could be solved. The study focuses on new events, evidence, or practices.
2. **Simple Design:** Is mainly a follow up of an exploratory design. The variable of interest has already been identified. It is used when the researcher intends to examine only a single problem.
3. **Comparative Descriptive Design:** Is mainly used when the researcher intends to examine and describe particular variables in two or more groups. The concept here is to compare the groups and how they differ or how similar they are in relation to the variable of interest.

4. **Time Dimensional Designs:** Are used in epidemiological studies and are further sub-divided into longitudinal that is when it examines changes in a group for a long period and it is cross-sectional where the data is collected at one point in time.
5. **Retrospective Study Design:** Also known as 'export facts'.
   It is a study design aimed at a looking back in order to link the present with the past or what happened in the past.
6. **Prospective Study Design:** Is similar to the longitudinal study as it starts from the present and ends by looking into the future. It is further divided into two categories: descriptive and explanatory.

- The subjects for the study are recruited based on presence or absence of an exposure of interest (workers in x-ray department) and followed up over many years to establish if they will develop outcomes of interest, for example cancer of the skin or reproductive health complications.

***Features used to diff. types of descriptive res. design***

1. Representativeness of the study data sources, for example, whether random, stratified, non probability.
2. Time frame of measurement, i.e., whether short, cross sectional or longitudinal.
3. Whether the study involves any comparisons, for example, with another group.
4. Whether the design is focused on a simple descriptive question or more complex, correlative questions.

# 4. Case Study Research Design

- A case study is 'an in depth study of one individual, a group of individuals or an institution’
- It is a detailed account of a particular experience event or situation.
- It is meant to provide a description of people’s thoughts, feelings and perceptions.
- It doesn’t aim at providing a causal relationship. Neither does it attempt to test a hypothesis.

For example, a case study on why in a certain health centres mothers are not coming in for their antenatal services.

## Limitations of case studies

1. They require plenty of time.
2. They are costly.
3. Have high possibility of subject drop out.
4. Data analysis also calls for skills and experience, particularly if the study is carried over a long period of time

- Case study designs are used when:
  1. There is a need to demonstrate the effectiveness of a specific therapeutic technique
  2. Generating and testing hypotheses
  3. There is need to generate knowledge on a particular issue or situation that has not been adequately researched on

# QUALITATIVE RESEARCH

- Qualitative research is a field of inquiry that cross-cuts disciplines and subject matters.
- It involves **in-depth understanding** of human behaviour and the **reasons** that govern human behaviour.
- It relies on reasons behind various aspects of behaviour. It focuses on understanding, rather than predicting or controlling phenomena.
- It investigates the **'why' and 'how'** of decision making

# Distinction between qualitative and quantitative research

| Quantitative | Qualitative |
| --- | --- |
| -what | -why |
| -Where | -How |
| -When | |
| -How many | |

# BIAS IN RESEARCH

- Systematic, non-random deviation of results and inferences from the truth, or processes leading to such deviation. Any trend in the collection, analysis, interpretation, publication or review of data that can lead to conclusions which are systematically different from the truth.
  - (Dictionary of Epidemiology, 3rd Edition)
- Bias can either be conscious or unconscious.
- In epidemiology, bias does not imply, as in common usage, prejudice or deliberate deviation from the truth.

# TYPES OF BIAS IN RESEARCH

- There are many types of biases, some studies being particularly prone to one type or another.
- Two main types:

➢ **Selection bias:** this occurs when the subjects studied are not representative of the target population about which conclusions are to be drawn.

➢ **Information bias:** results from the different quality of information and errors in obtaining and classifying information.

# TYPES OF SELECTION BIAS

- **Migration bias**: results from migration of subjects from an exposed status to an unexposed status during the course of a study.
- **Response bias**: Those who agree to be in a study may be in some way different with those who refuse to participate.
  - ➢ Volunteers may be different from those who are enlisted.
- **Membership bias** (health worker effect or influencing the direction of the research)
- **Prevalence-incidence bias**
- **Berksonian (admission rate) bias-** a type of selection bias arising when both exposure & the disease under study affect the selection e.g. in facilities/institutions that are commonly known for a specific task

# Types of information bias

- **Interviewer bias**: An interviewer's knowledge may influence the structure of questions and the manner of presentation, which may influence responses.
- **Recall bias**: Those with a particular outcome or exposure may remember events more clearly or amplify their recollections.
- **Diagnostic suspicion bias**: when potentially exposed subjects are subjected to more and in-depth diagnostic procedures and tests (cohort studies?).

# Types of information bias cont'd

- **Loss to follow-up**: Those that are lost to follow-up or who withdraw from the study may be different from those who are followed for the entire study.
- **Surveillance bias**: The group with the known exposure or outcome may be followed more closely or longer than the comparison group.
- **Observer bias**:
- **Misclassification bias**

# BASIC TERMS IN RESEARCH METHODOLOGY

- **Population:** refers to an **entire group** of individuals, events or objects having a common observable characteristic. It is **an aggregate** of all that conforms to a given specification
- The researcher first defines the population to which she or he wants to generalize the results. This is referred to as the "**target population** or the "universe".
- **Sample**: is a smaller group obtained from the accessible/Target population. Each member or case in this sample is referred to as "**subject**". Sometimes, the terms "**respondents**" or **interviewees**" are used.

- **Sampling:** the process of selecting a number of individuals for a study in such a way that the individuals selected represent the large group from which they were selected.
  - The purpose of sampling is to secure a **representative group** which will enable the researcher to gain information about the population

## ADVANTAGES OF SAMPLING

- Economy in expenditure. if data are collected for the entire population, cost will be very high
- Economy in time. Sampling is less time consuming than census technique
- There's a greater scope & flexibility of studies when a sample is used
- Greater accuracy. Sampling ensures completeness & high degree of accuracy due to a limited area of operation
- It's possible to collect intensive & exhaustive data as the number is limited
- Allows the researcher to work even with limited resources

# DISADVANTAGES OF SAMPLING

- There are chances of bias as sampling may involve biased selection & thereby lead to draw erroneous conclusions
- Difficulty in getting a representative sample especially when phenomena under study are of a complex nature
- Requires specialized knowledge of sampling technique, statistical analysis & calculation of probable error

# Characteristics of a good sample

- One that reproduces the characteristics of the population with the greatest possible accuracy
- Should be free from error due to bias or due to deliberate selection of the unit of the sample
- There should not be any substitution of originally selected unit by some other more convenient way

# THE SAMPLING PROCESS

- Define the population: the population that the research wants to obtain information from according to the objectives of the study should be clearly defined
- Specify the sampling frame: it's the means of representing the elements of the population
- Specify the sampling unit
- Specify the sampling method
- Determine the sample size
- Specify the sampling plan: the operational procedure for selection of the sampling units is specified

- **Sampling Frame:** Is a comprehensive list of all the sampling elements in the target population, for example, the list of nurses working in a particular county, the number of under five children in a village or all the households in a village.
- **Representative Sample:** the sample that resembles the population from which it is drawn in all aspects. It should possess all the variables a researcher is interested in, for example, educational level, socioeconomic factors etc.
- **Sampling Bias:** Occurs when the researcher has not carefully selected the samples that are expected to represent the general target population.

- **Sampling Error:** Refers to the difference between population parameters (for example, the average age of the population) and the sample statistics (for example, the average age of the sample group). It is the degree of deviation of the sample from the population from which it was drawn.

**Sampling Techniques** 2 types

- Probability or random sampling
- Non probability sampling

- The **nature of your research** will determine which type of sampling you should use.
- ✓ Large-scale exploratory/descriptive studies almost always use probability sampling techniques.
- ✓ Intervention studies sometimes use probability sampling but also frequently use non probability sampling.
- ✓ Qualitative studies almost always use non-probability samples

# FACTORS TO CONSIDER WHEN SELECTING A SAMPLING TECHNIQUE

- **Purpose of the survey:** if researcher intends to generalize the findings based on the sample survey to the population, then appropriate probability sampling method must be selected. This depends on the geographical area of the survey, the size & nature of population under study
- **Measurability:** where research objectives require statistical inference, sample should be drawn by applying s.r.s. method or stratified sampling depending n whether the population is homogenous or heterogeneous

- **Information about population:** where no information about a population is available, it's difficult to apply probability sampling method. Exploratory study with non probability sampling may be made to gain a better idea of the population then adopting appropriate probability sampling method after gaining sufficient knowledge about the population
- **Nature of the population:** variables to be studied. If the population is homogenous, even s.r.s. will give a representative sample. If the population is heterogeneous, stratified sampling is appropriate
- **geographical area of the study & the size of the population:** if the area covered by survey is very large & the size of population is large, multistage & cluster sampling may be appropriate

- **financial resources**: if limited, choose a less costly sampling method like quota sampling
- **time limitation:** choose a less time consuming method like s.r.s. instead of stratified sampling where time is limited

# METHODS OF SAMPLING:

## 1. PROBABILITY (RANDOM) SAMPLING

- Allows for a much more a representative sample of the population and enables the estimation of sampling error.
- In probability sampling, each study unit has an **equal** or at least a known **non-zero probability** of being selected in the sample.

- It also enables the calculation of differential statistics and allows the study to be generalized to other areas.
- It involves a random selection procedure to ensure that each unit of the sample is chosen on the basis of chance.
- All units of the study population have an equal or at least a known chance of being included in the sample from list provided

## PROBABILITY SAMPLING TECHNIQUES

1. **Simple Random Sampling:** one of the commonest and the simplest methods of sampling.

- Each unit (subject) has the chance to be selected.
- It involves one stage selection. It also allows the researcher access to the study population
- There are several ways of selecting a random sample using this technique. These include;
  - The lottery method
  - Use of random tables, or tossing a coin to help you decide where and how to start.

# SIMPLE RANDOM SAMPLING CT'

LOTTERY METHOD

- Here, each member or item of the population at hand is assigned a unique number.
- Then, the numbers are then thoroughly mixed, like if you put them in a bowl or jar and shook it.
- Then, without looking, the researcher selects *n* numbers. The population members or items that are assigned that number are then included in the sample.

# SIMPLE RANDOM SAMPLING CT'

BY USE OF RANDOM NUMBER TABLE

- Most statistics books and many research methods books contain a table of random numbers as a part of the appendices.
- A random number table typically contains 10,000 random digits between 0 and 9 that are arranged in groups of 5 and given in rows.
- In the table, all digits are equally probable and the probability of any given digit is unaffected by the digits that precede it.

## 2. Systematic or Interval Sampling:

- This is a modification of simple random sampling, which is ordinarily less time-consuming and easier to implement.
- The estimated number of elements in the larger population is divided by the desired sample size, yielding a sampling interval, **n.** It is thre4 the selection of every **$n^{th}$ element** from a sampling frame, where n, the sampling interval, is calculated as;
  - **n= number in population÷ desired number in sample**
- This involves listing the population elements in an arbitrary order and selecting every nth case, starting with a randomly selected number between 1 and n. The selected study units constitute the sample

- It is much more efficient and much less expensive to do than simple random sampling.

**3. Stratified Random Sampling:**

- A useful method of sampling if the population is **heterogeneous.**
- When sub-populations vary considerably, it is advantageous to sample each subpopulation (stratum) independently.
- Stratification is the process of grouping members of the population into **relatively homogeneous subgroups**, e.g., by education level, age etc before sampling.
- The strata should be mutually exclusive, i.e., every element in the population must be assigned to only one stratum.

- The strata should also be collectively exhaustive, that is, no population element can be excluded.
- Simple Random sampling is then applied within each stratum to select the subjects. This often improves the representativeness of the sample by reducing sampling error.

**4. Cluster Sampling:**

- Is a sampling technique used when **'natural' groupings** but relatively heterogeneous groupings are evident in a statistical population.
- The total population is divided into these groups (or clusters), and a sample of the groups is selected. The required information is then collected from the elements within each selected group.

- This may be done for every element in these groups, or a subsample of elements may be selected within each of these groups.
- Elements within a cluster should ideally be as homogeneous as possible. However, there should be heterogeneity between clusters.
- Each cluster should be a small scale version of the total population. The clusters should be mutually exclusive and collectively exhaustive.
- A random sampling technique is then used on any relevant clusters to choose which clusters to include in the study.

## 2. NON PROBABILITY SAMPLING METHOD

- This type of sampling is used for interventional studies.
- Non-probability sampling refers to the selection of a sample that is not based on known probabilities.
- It is distinguished from probability sampling by the fact that subjective judgment plays a role in selecting the sampling elements.
- When the probabilities of selection are not known, there is no precise way to adjust for such distortions.
- Despite these drawbacks, there are many instances in which obtaining a truly representative probability sample may be too difficult or too expensive.

- In fact, much of health research uses some kind of non-probability sampling.
- For example, it is usually necessary to use non-probability samples when studying sex workers and their clients, injecting drug users, gay men and lesbians

# TYPES OF NON-PROBABILITY SAMPLING METHODS

1. **Purposive or Judgmental Sampling**: it is where the researcher selects a particular group or groups based on certain criteria.

- In this method the researcher determines who should be included in the study. It is, in fact, the researcher's opinion that the sample is representative of the target population. Commonly used in qualitative studies.

**Advantage:** it gives the researcher a free hand to respond according to their judgment.

**Disadvantages:**

- Sampling biases are relatively common with this method of sampling.
- The possibility of unrepresentative samples.
- Lack of generalizations of the study findings due to perceived lack of representativeness.

**2. Quota Sampling**: The population is first segmented into mutually exclusive sub-groups, just as in stratified sampling.

- Then judgment is used to select the subjects or units from each segment based on a specified proportion.

**3. Convenience or Accidental/Availability Sampling**

- The researcher is unable to control bias at all.
- The study units that happen to be available at the time of data collection are selected and used as a sample. This type of sampling allows the utilization of any available target population.

e.g., if you need to assess the BP of females using depo provera who are above 40 years of age, you will need to check the blood pressure of any female patient above this age irrespective of her parity or other characteristics.

4. **Snow Ball Sampling:** it is a technique for developing a research sample where existing study subjects recruit future subjects from among their acquaintances. Thus, the sample group appears to grow like a rolling snowball.

- This sampling technique is often used in hidden populations which are difficult for researchers to access. E.g. populations include drug users and commercial sex workers.
- Sample members are not selected from a sampling frame, therefore, snowball samples are subject to numerous biases.
- Other terms used in this case are;
  - sister to sister approach
  - brother to brother approach
  - child to child approach

# Difference btwn probability & non-probability sampling

- Probability sampling involves random selection, non-probability doesn't involve random selection of the study participants.
- Probability samples can depend upon the rationale of probability theory, non-probability cannot
- Probability sampling represents the ppln well, non-probability samples may or may not represent the ppln well (they lack representatives)

# Determination of the sample size

- Sample size determination is the technique of selecting the number of observations to include in a sample.
- The sample size is an important feature of any study or investigation in which the aim is to make inferences about the population from a sample.
- In general, the sample size used in a study is determined based on the cost of data collection, and based on sufficient statistical power

There are many approaches to determining the sample size. These include:

- **Using a census for small populations,**
- **Imitating a sample size of similar studies,**
- **Using published tables, and**
- **Applying formulas to calculate a sample size.**

# i) Using census for small populations

- One approach is to use the **entire population** as the sample
- A census approach is more convenient to use for small populations e.g. 200 or less
- A census eliminates sampling error and provides data on all the individuals in the population.
- In addition, some costs such as questionnaire design and developing the sampling frame are "fixed," that is, they will be the same for samples of 50 or 200

# ii) Using the sample size of a similar study

- Another approach is to use the same sample size as those of studies similar to the plan.
- Without reviewing the methods used in these studies may run the risk of repeating errors that were made in determining the sample size for another study, thus, not ideally recommended to rely solely on this method

# iii) Using Published tables

- A third way to determine sample size is to rely on published tables, which provide the sample size for a given set of criteria, given combinations of precision, confidence level and variability.
- Most studies use a confidence level of 95% (P=5%) or 90% (P=10%)
- Glenn (1992), presented two tables for the selection of sample size as shown below:

| Size of Population | Sample Size (n) for precision (e) | |
|---|---|---|
| | ±5% | ±10% |
| 500 | 222 | 83 |
| 1,000 | 286 | 91 |
| 2,000 | 333 | 95 |
| 3,000 | 353 | 97 |
| 4,000 | 364 | 98 |
| 5,000 | 370 | 98 |
| 7,000 | 378 | 99 |
| 20,000 | 392 | 100 |
| 25,000 | 394 | 100 |
| 50,000 | 397 | 100 |
| 100,000 | 398 | 100 |
| >100,000 | 400 | 100 |

# Table 2

| Size of Population | Sample Size (n) for Precision (e) of: | |
|---|---|---|
| | ±5% | ±10% |
| 100 | 81 | 51 |
| 125 | 96 | 56 |
| 150 | 110 | 61 |
| 200 | 134 | 67 |
| 250 | 154 | 72 |
| 300 | 172 | 76 |
| 350 | 187 | 78 |
| 400 | 201 | 81 |
| 450 | 212 | 82 |

# iv) Using formulas to calculate the sample size

- There are many formulas used in calculation of the sample size. One and the most common is the Fischer's statistical formulae as described below:

# SAMPLE SIZE DETERMINATION USING FISCHER'S STATISTICAL FORMULA

- In social science research, the following formula can be used to determine the sample size? **FISCHER'S statistical formula (Mugenda A. & Mugenda O., 1994)**

$$n = \frac{Z^2 pq}{d^2}$$

Where:

**n** = the desired sample size (if the target population is greater than 10,000)

**z** = the standard normal deviate at the required confidence level, usually set at 1.96 which corresponds to the 95% confidence level

**p** = the proportion in the target population estimated to have characteristics being measured in the study.

**q** = the proportion in the target population estimated not to have characteristics being measured=1-p

**d** = the level of statistical significance/ degree of accuracy desired, usually set at the 0.05 level, corresponding with 95% confidence level.

- If there is no estimate available of the proportion in the target population assumed to have the characteristics of interest, 50% should be used as recommended by Fischer's *et al.*
- For example, if the proportion of a target population with a certain characteristic is 0.50, the Z statistic is 1.96, and we desire accuracy at the 0.05 level, then the sample size is

$$n = \frac{(1.96)^{2\times}\,(0.50)\times\,(0.50)}{(0.05)^2} = 384.16$$

- If the target population is less than 10,000, the required sample size will be smaller. In such cases, calculate a final sample estimate ($n_f$) using the following formula:

$$n_f = \frac{n}{\{1+ (n/N)\}}$$

Where:

nf = the desired sample size (when the population is less than 10,000)

n = the desired sample size (when the population is more than 10,000).

N = the estimate of the target population size

- For example: if n = 384 and now our N is 1000, what is our nf?

384/(1+384/1000)

= 384/1.384= 277 respondents

**Commonly used confidence coefficients & their z values**

| P | Z |
|---|---|
| 0.90 | 1.64 |
| 0.95 | 1.96 |
| 0.96 | 2.00 |
| 0.98 | 2.33 |
| 0.99 | 2.58 |

# STEP 7: METHODS OF MEASUREMENT

## 1. Review of Existing Records

***Advantages***.

a) Saves time and money

b) Limits respondent bias. Existing data utilizes records, which are unbiased, as the person who collected the data had no knowledge of the future use to which it would be put.

c) They cover a long period of time, which is particularly useful to the researcher who can only dedicate a short time to their research.

d) Saves the researcher from the worries and concerns of seeking the cooperation of the respondents.

***Disadvantages***

- Since the researcher is not responsible for the collection and recording of the data, limiting biases and even ascertaining the authenticity of the data would be difficult. This can often result in doubts about the validity of the data.

## 2. Structured Interview Schedule

It is a formal and written document where questions are asked orally in either face-to-face or telephone interviews.

The responses are then recorded by the researcher

***Advantages***

1. It provides comparability of responses and facilitates analysis.

2. Good for **measuring attitudes** and most other content of interest.
3. Allows probing and posing of follow-up questions by the interviewer.
4. Can provide in-depth information.
5. Can provide information about participants' internal meanings and ways of thinking.
6. Closed-ended interviews provide exact information needed by researcher.
7. Telephone and e-mail interviews provide very quick turnaround.
8. Moderately high measurement validity (i.e., high reliability and validity) for well constructed and tested interview protocols.

9. Also applicable to a special category of respondents, such as children, the elderly and the illiterate who may not be able to read and write.
10. Useful for exploration as well as confirmation.
11. The instrument is expected to have a high response rate, since the researcher administers it personally.
12. An interview schedule also has the advantage of capturing the respondent's own words.

***Disadvantages***

1. The instrument demands a much longer time to complete than other instruments, such as questionnaires.
2. Due to the presence of the researcher, respondents may withhold certain vital information or even change information to please the researcher- Reactive effects

3. In-person interviews usually are expensive and time consuming.
4. Investigator effects may occur (e.g., untrained interviewers may distort data because of personal biases and poor interviewing skills).
5. Interviewees may not recall important information and may lack self-awareness.
6. Perceived anonymity by respondents may be low.
7. Data analysis can be time consuming for open-ended items.
8. Measures need validation.

**Key informant**

- Def: a person (or group of persons) who has unique skills or professional background related to the issue/intervention being evaluated, is knowledgeable about the project participants, or has access to other information of interest to the evaluator.
- Can also be someone who has a way of communicating that represents or captures the essence of what the participants say and do.

**Advantages**

1. Information concerning causes, reasons, and/or best approaches from an "insider" point of view.

2. Advice/feedback increases credibility of study
3. Pipeline to pivotal groups
4. May have side benefit to solidify relationships between evaluators, clients, participants, and other stakeholders

- **Disadvantages**

1. Time required to select and get commitment may be substantial
2. Relationship between evaluator and informants may influence type of data obtained
3. Informants may interject own biases and impressions
4. May result in disagreements among individuals leading to frustration/ conflicts

**3. Focus Group Discussions**

- They are interviews with groups of 5-15 people whose opinions and experiences are solicited simultaneously.
- The composition of the group is usually limited to those with similar characteristics, such as socio- economic status, so that the members feel free in contributing to the issue at hand.
- It is efficient and can generate dialogue
- The instrument allows the members to share their views, experiences and opinions.
- The interpersonal interactions create a free and enjoyable environment.

***Advantages***

1. Allows a large number of respondents to be interviewed at one go, which saves time and money.
2. Useful for exploring ideas and concepts.
3. Provides window into participants' internal thinking.
4. Can obtain in-depth information.
5. Can examine how participants react to each other.
6. Allows probing.
7. Most content can be tapped.
8. Allows quick turnaround.

***Disadvantage***

1. Because of the number of respondents involved, calls for diligence and skill in ensuring that the process runs smoothly and yields the desired information.
2. Sometimes expensive.
3. May be difficult to find a focus group moderator with good facilitative and rapport building skills.
4. Reactive and investigator effects may occur if participants feel they are being watched or studied.
5. May be dominated by one or two participants.
6. Difficult to generalize results if small, unrepresentative samples of participants are used.
7. May include large amount of extra or unnecessary information.

8. Measurement validity may be low.
9. Usually should not be the only data collection methods used in a study.
10. Data analysis can be time consuming because of the open-ended nature of the data.

**4. In-depth Interviews (unstructured interview)**

It utilizes face-to-face in-depth interviews using semi-structured questionnaires for key informants.

Key informant interviews are defined as interviews with people who have special positions in the community and whose opinions and experiences are seen as representative of a whole group.

***Advantages***

1. Semi-structured questionnaires possess a flexibility that allows the researcher to gather in depth information.
2. The instrument enables the participants to give responses in a narrative form and is quite useful in qualitative research.
3. The respondent takes the lead and determines the flow of the conversation, which is of great importance when new areas are to be investigated.
4. Key informant interviews provide valuable and independent information about the research population within a short span of time and save time and money.

***Disadvantage***

It requires the researcher to be very articulate.

## 5. Non Participatory Structured Observation

- It is observing a given situation from the outside.
- The observer declares their intention to observe and goes ahead to watch the activities being carried out without asking questions or interfering in any way.
- As the activities to be observed progress, the observer remains in the background, keenly observing and noting down events without comment.
- It provides the in-depth and variety of information.
- The observers are used as measuring instruments and provide a uniquely sensitive and intelligent tool.

***Advantages***

1. Allows one to directly see what people do without having to rely on what they say they do.
2. Provides firsthand experience, especially if the observer participates in activities.
3. Can provide relatively objective measurement of behavior (especially for standardized observations).
4. Observer can determine what does not occur.
5. Observer may see things that escape the awareness of people in the setting.
6. Excellent way to discover what is occurring in a setting.
7. Helps in understanding importance of contextual factors.
8. Can be used with participants with weak verbal skills.

9. May provide information on things people would otherwise be unwilling to talk about.
10. Observer may move beyond selective perceptions of people in the setting.
11. Good for description.
12. Provides moderate degree of realism (when done outside of the laboratory).

***Disadvantages***

1. Reasons for observed behavior may be unclear.
2. Reactive effects may occur when respondents know they are being observed (e.g., people being observed may behave in atypical ways).
3. Investigator effects (e.g., personal biases and selective perception of observers)

4. Observer may “go native” (i.e., over-identifying with the group being studied).
5. Sampling of observed people and settings may be limited.
6. Cannot observe large or dispersed populations.
7. Some settings and content of interest cannot be observed.
8. Collection of unimportant material may be moderately high.
9. More expensive to conduct than questionnaires and tests.
10. Data analysis can be time consuming

## 6. QUESTIONNAIRES

- A questionnaire is a self-report data collection instrument that is filled out by research participants.

### MAIN MODES OF QUESTIONNAIRE ADMINISTRATION ARE:

- Face-to-face questionnaire administration, where an interviewer presents the items orally.
- Paper-and-pencil questionnaire administration, where the items are presented on paper.
- Computerized questionnaire administration, where the items are presented on the computer.
- Adaptive computerized questionnaire administration, where a selection of items is presented on the computer, and based on the answers on those items, the computer selects following items optimized for the testee's estimated ability or trait.

***15 principles of questionnaire construction***

Principle 1: Make sure the questionnaire items match your research objectives.

Principle 2: Understand your research participants.

- Your participants (not you!) will be filling out the questionnaire.
- Consider the demographic and cultural characteristics of your potential participants so that you can make it understandable to them.

Principle 3: Use natural and familiar language.

- Familiar language is comforting; jargon is not.

Principle 4: Write items that are clear, precise, and relatively short.

- If your participants don't understand the items, your data will be invalid (i.e., your research study will have the garbage in, garbage out, GIGO, syndrome).
- Short items are more easily understood and less stressful than long items.

Principle 5: Do not use "leading" or "loaded" questions.

- Leading questions lead the participant to where you want him or her to be.
- Loaded questions include loaded words (i.e., words that create an emotional reaction or response by your participants).

- Always remember that you do not want the participant's response to be the result of how you worded the question. Always use neutral wording.

Principle 6: Avoid double-barreled questions.

- A double-barreled question combines two or more issues in a single question (e.g., here is a double barreled question: “Do you elicit information from parents and other teachers?” It’s double barreled because if someone answered it, you would not know whether they were referring to parents or teachers or both).
- Does the question include the word "and"? If yes, it might be a double-barreled question.
- Answers to double-barreled questions are ambiguous because two or more ideas are confounded.

Principle 7: Avoid double negatives.

- Does the answer provided by the participant require combining two negatives? (e.g., "I disagree that teachers should not be required to supervise their students during library time"). If yes, rewrite it.

Principle 8: Determine whether an open-ended or a closed ended question is needed.

- Open-ended questions provide qualitative data in the participants' own words. Here is an open ended question: How can your principal improve the morale at your school?
- Closed-ended questions provide quantitative data based on the researcher's response categories.

Principle 9: Use mutually exclusive and exhaustive response categories for closed-ended questions.

- Mutually exclusive categories do not overlap (e.g., ages 0-10, 10-20, 20-30 are NOT mutually exclusive and should be rewritten as less than 10, 10-19, 20-29, 30-39, .
- Exhaustive categories include all possible responses (e.g., if you are doing a national survey of adult citizens (i.e., 18 or older) then the these categories (18-19, 20-29, 30-39, 40-49, 50-59, 60-69) are NOT exhaustive because there is no where to put someone who is 70 years old or older.

Principle 10: Consider the different types of response categories available for closed-ended questionnaire items.

# ESSENTIALS OF A GOOD QUESTIONNAIRE

- Should be comparatively short & simple. The size of the questionnaire should be kept to the minimum
- Questions should proceed in a logical sequence moving from easy to more difficult ones. Personal and intimate questions should be kept to the end
- Technical terms and vague expressions must be avoided to minimize misinterpretations

Rating scales are the most commonly used, including:

- Numerical rating scales (where the endpoints are anchored; sometimes the center point or area is also labeled).
- 1 2 3 4 5 6 7
- Very Low Very High

Fully anchored rating scales (where all the points on the scale are anchored).

1 2 3 4 5

- Strongly Agree Neutral Disagree Strongly
- Agree Disagree

Principle 11: Use multiple items to measure abstract constructs.

- This is required if you want your measures to have high reliability and validity.

Principle 12: Consider using multiple methods when measuring abstract constructs.

- The idea here is that if you only use one method of measurement, then your measurement may be an artifact of that method of measurement.

Principle 13: Use caution if you reverse the wording in some of the items to prevent response sets. (A response set is the tendency of a participant to respond in a specific direction to items regardless of the item content.)

- Reversing the wording of some items can help ensure that participants don't just "speed through" the instrument, checking "yes" or "strongly agree" for all the items.

Principle 14: Develop a questionnaire that is easy for the participant to use.

- The participant must not get confused or lost anywhere in the questionnaire.
- Make sure that the directions are clear and that any filter questions used are easy to follow.

Principle 15: Always pilot test your questionnaire.

- You will always find some problems that you have overlooked!
- The best pilot tests are with people similar to the ones to be included in your research study.

## Strengths of questionnaires

- Good for measuring attitudes and eliciting other content from research participants.
- Inexpensive (especially mail questionnaires and group administered questionnaires).
- Can provide information about participants' internal meanings and ways of thinking.
- Can administer to probability samples.
- Quick turnaround.
- Can be administered to groups.
- Perceived anonymity by respondent may be high.
- Moderately high measurement validity (i.e., high reliability and validity) for well constructed and validated questionnaires.

## Weaknesses of questionnaires

- Usually must be kept short.
- Reactive effects may occur (e.g., interviewees may try to show only what is socially desirable).
- Nonresponsive to selective items.
- People filling out questionnaires may not recall important information and may lack self-awareness.
- Response rate may be low for mail and email questionnaires.
- Open-ended items may reflect differences in verbal ability, obscuring the issues of interest.
- Data analysis can be time consuming for open-ended items.
- Measures need validation.

# 7. ACTION RESEARCH

**DESCRIPTION:**

- This is whereby the **researcher actually becomes part of the study** by participating in various interventions involved in the study
- Useful in collecting **qualitative data**
- Action Research is a framework that is:
  - Collaborative
  - There is a practical intervention made - i.e. you do something to make a change or intervention in a situation that you research
  - The researcher will be actively involved in the planned intervention(s).

**<<<……>>>>**

# CONDUCTING A PRE-TEST OR PILOT STUDY

- A pilot study may be defined as ‘a small version of a proposed study conducted to refine the methodology.
- It is developed and conducted in a manner similar to the proposed study, using similar research respondents and the same setting.
- A pilot study may also be defined as ‘the process of carrying out a preliminary study going through the entire research procedure with a small sample’.
  A pre-test usually refers to a small scale trial of a particular research component.

**purposes of a pre-test or pilot study**:

1) To determine whether the proposed study is feasible
2) Identify any problems with the research design
3) To ensure that items in the data collection instrument are stated clearly and have the same meaning to all research respondents
4) To assess the time taken to administer the research instrument
5) Determine whether the sample is representative of the population
6) To determine the effectiveness of the sampling technique used
7) Give the researcher the real experience in the field

8) Determine the human and financial resources requirement for the study
9) Determine the effectiveness of the training given to research assistants where necessary
10) Evaluate the procedure for data processing and analysis

- The advantages of conducting the pre-test before you finalize your proposal is that you can draft the work plan and budget based on realistic estimates as well as revise the data collection tools before you submit your proposal for approval.

# ETHICAL CONSIDERATIONS IN RESEARCH

**Ethics**: is defined as that branch of philosophy which deals with one's conduct and serves as a guide to one's behavior'.

- It is important to note that most professions, including nursing, have a code of conduct to which all members of the profession have to adhere.
- Plooy (2003) describes ethics in research as 'that which is morally justifiable'.
- Research ethics fundamentally consist of collecting, analyzing and interpreting data in a way that respects the rights & humanity of your participants and respondents.

## Basic Ethical Principles Underlining the Protection of Human Rights

### 1. Principle of Respect for Persons

It involves two main convictions:

a) Individuals are autonomous, that is, they have the right to self-determination and this right should be respected.

- This means the research respondents have the right to:

  a) Accept or decline to participate in the study without punishment or prejudice

  b) Withdraw from the study at any stage

  c) Withhold information

  d) Seek clarification concerning the purpose of the study where it's not clear

b) Individuals with diminished autonomy require protection. This group includes

- Children,
- The mentally impaired,
- Unconscious patients and
- Institutionalized persons.

- For this group of persons, you will need to seek the consent of their **legal guardian.**

**2. Principle of Justice**

This includes the subjects' right to fair selection and treatment and their right to privacy:

- Right to Privacy: This is the freedom of an individual to determine the time, extent and the circumstances under which private information will be shared with or withheld from others. The privacy of the subject is considered to be protected if the subject is informed and consents to participate in a study and voluntarily shares private information with a researcher.

- Right to Anonymity and Confidentiality:

Complete anonymity exists when the respondents' identity is not revealed and the information collected is not linked to the respondent.

Confidentiality refers to the researcher's responsibility to protect all data gathered within the scope of the project from being divulged or made available to any other person, which means the research data should never be shared with outsiders.

A breach of confidentiality can occur when a researcher allows an unauthorized person(s) to gain access to the raw data of a study or when the researcher accidentally reveals the identity of the research respondents.

**3. Principle of Beneficence**

This principle involves an effort to secure the well being of persons.

It is the right to protect respondents from discomfort and harm. This principle states that one should do what is good and above all should do no harm.

Discomfort and harm can be physical, emotional, spiritual, economic, social or legal.

# The Major Content of an Informed Consent

Informed consent revolves around the following three major elements:

- The type of information you need to obtain from the research subjects.
- The degree of understanding required of the subject in order to give consent.
- The fact that the subject has a free choice in giving consent.

Informed consent should be based on the following factors:

- The purpose of the research study
- Foreseen risks
- A guarantee of anonymity and confidentiality
- Identification of the researcher
- Number of subjects involved
- Benefits and compensation or lack of them

# Access to Research Population

- Prior to commencing the study, a formal application to the government of Kenya for permission to conduct the research must be written.
- This should include two or more copies of the research proposal, accompanied by a recommendation letter from the supervisor(s) as required by the Kenyan authorities.
- If your institution is authorized to conduct research, there may be a 'Research and Ethics Committee'. Such a committee is usually vested with the authority of granting research permits, which you could use.
- As a requirement, each research respondent should be requested to accept in writing and sign or affix a thumb print.

- In cases where a respondent can neither read nor write, a consent form should be completed and duly signed after you have clearly explained the purpose of the research.
- You should inform the respondents that their participation is absolutely voluntary and they may pull out of the study whenever they so wish.
- As part of the contract, you should guarantee the respondents absolute confidentiality during and after the study.

# Writing a Research Proposal

# Introduction

- A proposal is a document which details an intended activity.
- It's a contract in that once accepted by an institution or a funding agency, the writer of the proposal is obliged to follow the steps outlined in the proposal.

## Components of a research proposal:

- An overview. Consists of
  - An introduction section which includes the background to the problem, statement of the problem, the purpose, objectives & justification of the study.
  - A literature review section
  - A methodology section which includes description of research design, population, sample and sampling techniques, data collection procedures and methods of analysis.
  - A time schedule.
  - A budget.

# STEP 8: DATA COLLECTION AND PRESENTATION

**Def:** 'the precise, systematic gathering of information relevant to the research purpose or the specific objectives, questions or hypotheses of a study'.

There are three main stages in the data collection process:

- Stage One - Permission to Proceed(seeking consent)
- Stage Two - Data Collection
    - Listing tasks, training assistants, available data, time
- Stage Three - Data Handling

***Sources of data bias***

1. The subject being studied changes their behaviour as a consequence of the research.
2. The researcher may use non-standard measuring scales, imprecise or no guidelines for interviewing.
3. Researchers themselves vary in what they observe or measure, that is, observer variability. E.g., researchers may be selective in their observations, that is, observer bias and measure, question, or note down answers with varying accuracy or follow different interview approaches, with one researcher being more open, friendly and probing than the other.

Aspects of data collection that ensure data quality include;

- Guidelines on sampling procedures and what to do if respondents are not available or refuse to cooperate.
- A clear explanation of the purpose and procedures of the study. This should be used as an introduction before each interview.
- Instruction sheets on how to ask certain questions and how to record the answers.

members of the research team master techniques such as:

- Asking questions in a neutral manner.
- Not showing, by words or expressions, what answers one expects to hear.

- Not showing agreement, disagreement or surprise.
- Recording answers precisely as they are provided, without sifting through them or interpreting them.

**Data Handling guidelines**

- Check that the data gathered is complete and accurate.
- At some stage questionnaires will have to be numbered. Decide if this should be done at the time of the interview or when the questionnaires are stored.
- Identify the person responsible for storing data and the place where it will be stored.
- Decide how data should be stored. Record forms should be kept in the sequence in which they have been numbered.

# DATA ENTRY

- Decide on a format, that is, the way you will organize the data in a file.
- Next, design a code, that is, the rules by which the respondents' answers will be assigned values that can be processed by the computer.
- Then do the actual coding, that is, turn the responses into the standard categories you developed in your coding system.
- Data entry is the next step, which is keying the data into the computer so that you can process it.
- Finally, data cleaning is the final check you make on the data file for accuracy, completeness, and consistency prior to the onset of analysis.

# STEP 9: DATA ANALYSIS AND INTERPRETATION

**Data Presentation:** is the way in which data is displayed for viewing, interpreting & understanding

**methods of data presentation**

- Tables
- Charts
- Graphs
- Frequency distribution tables
- Histograms
- Narrative method

**Qualitative Data Presentation and Analysis**

The data presentation and analysis of qualitative research is quite different from that of presenting data collected when using the quantitative research method

This is because qualitative research uses words while quantitative uses numbers (numerical).

However, the principles are the same for both types. In both cases the researcher has to do the following:

a) Describe the sample population by providing a description of the:

- Respondents, for instance, key informants or focus group members.
- Age, sex, occupation, educational background etc.

b) Order, reduce and/or code the data (data processing).

c) Display the summaries of data for interpretation.

d) Draw conclusions.

e) Develop strategies for testing or confirming the findings to prove their validity.

# Measures of Central Tendency

- Referred to as 'average' measures. They describe how closely related the data is.

1. **Mode:** it is the numerical value or score that occurs most times. It is the most suitable measure of central tendency for nominal data.
2. **Median:** it is the score at the exact centre of a distribution; it is also called the 50th percentile. It is the most central value when raw data is arranged on a scale from the highest to the lowest.
3. **Mean:** It is the total sum of scores divided by the number of scores being summed. The mean is the most suitable measure of central tendency for interval and ratio level data.

# Measures of Dispersion

- Are used to measure the individual differences of scores in a sample. They give an indication of how scores in a sample are dispersed around the mean.
- They show how different the scores are or the extents to which individual scores deviate from one another.
- If the individual scores are similar, the measure of variability is small and the sample is relatively similar or homogeneous in terms of those scores. A wide variation in scores may indicate a heterogeneous sample.

1. **Range:** It is obtained by subtracting the lowest score from the highest score. The range is the difference between the highest and lowest score. It is not a very significant statistical measure.

2. **Variance**: it is a measure of how individual scores in a set of data vary in their distribution from one to the other.
3. **Standard Deviation:** It is calculated by finding the square root of variance - that means you have to calculate the variance first.

## MEASUREMENT SCALES

1. **Nominal scale:** the lowest level of measurement. It groups subjects or cases from the sample into categories. Variables which can only be measured at the nominal scale include: sex, race, marital status, color etc
2. **Ordinal scale:** it not only groups subjects into categories, but it also ranks them into some order, this could be in an increasing order. i.e., in an ordinal scale, numerals are used to represent relative position or order among the values of the variables. E.g.; social class, military rank etc.

3. **Interval scale:** The numerals are assigned to each measure and ranked in an order and the intervals between the numerals are equal.

Mathematical operations are limited to additions and deductions; multiplication and division are not applicable.

An interval scale does not have a true zero point. The minimum and maximum points o the scale are only arbitrary

4. **Ratio scale:** is the highest level of measurement. It is the most precise method of measuring variables. It has all the characteristics of the other scales. The only additional characteristic is that it has a true zero point and all the mathematical operations can be applied to yield meaningful values. Most physical objects can be measured at the ratio scale. E.g. height, weight, distance, age, area,

# Qualitative Data Presentation and Analysis

1. **Organizing the Data**: involves putting all the information in a simple format that can be understood. This is known as ‘cleaning’ the data.
2. **Creating Categories, Themes and Patterns**: The researcher needs to be very familiar with the data so as to establish relationships among these categories. One use the research questions or discussion topics.
3. **Analyzing and Interpreting the Data:** it involves evaluating the data to determine its usefulness
   and accuracy.
4. **Writing the Report:** Unlike the quantitative research where the report writing is done after analyzing the data, in qualitative research techniques, the writing and the analysis go hand in hand.

# Quantitative Data Presentation and Analysis

## 1. Tabular Method

- **Simple Table:** usually a single line of characters explaining a few columns of information.
- **Compound Table**: a single line of characters has been described by two or more components of information.

When constructing tables it is important that:

- You give correct consecutive numbers to the tables and indicate the title of the table at the top.
- Clearly indicate the totals and percentages.
- Show where the data is obtained from, that is the source of the data with the year when the data was collected clearly shown.

## 2. Graphic Method

Graphs are used to organize and describe data.

This enables the reader to see at a glance the trend of distribution of the data. Graphs have two axes: vertical and horizontal.

Scores are usually presented along the horizontal axis, while the frequency is placed along the vertical axis.

It is important to note that the intersection point between the vertical and horizontal axes is usually represented by a zero (0).

Graphs should be well labelled both along the vertical axis and the horizontal. There are three common graphic methods that you may use for presenting data, i.e. Bar Charts, histograms, frequency polygons

- **Bar Charts**: used when the data being presented is discrete or when the scale is nominal. Bar charts are quite similar to histograms with the exception that there are spaces in bar charts.
- **Histograms:** Unlike the bar chart, in the histogram there are no spaces between the bars. A histogram in many cases is used to represent continuous variables.
- **Frequency Polygons**
  Frequency polygons are drawn based on the frequencies of the observations along the vertical axis against the group or class midpoints.
  They form a polygon shape hence the name.

3. **Charts Method**: they include; pie charts, doughnut charts, and scatter charts. One of the commonest types of charts used is the pie chart. Pie charts are relatively easy to interpret.

Each portion represents a variable.

# STEP 10: COMMUNICATING THE RESEARCH FINDINGS

Ways of communicating the research findings include:

1) A written report for academic purposes, e.g., a dissertation or a thesis, which are a requirement in the obtaining of a certain academic level.
2) A written report prepared for managers and policy implementers.
3) A written report sent as an article for publication in refereed scholarly journals.
4) Presentations of the research findings in workshops, seminars and conferences.
5) Written clinical summary statements which are then distributed to the relevant departments
6) Presentations at local unit meetings & at various hospital committee meetings
7) Presentations at continuing educational inservices

# STUDY LIMITATIONS

Limitations may include:

1. Factors such as the inherent weakness in the sampling method, faulty designs and controls, weaknesses in the methods used to collect data and so on.
2. Time factor due to pressure of work.
3. Expenses involved if grant is not secured, where the nature of the research needs acquiring grants
4. Possibility that some of the respondents may not agree to participate in the study and others dropping out as the study progresses
5. Diverse spread of the target population, which may hinder easy access to respondents.

The researcher has the opportunity to recommend ways to minimize or eliminate the limitations of the current study or to offer alternative methodology or improvements of the methods of the study presented.

# REFERENCES, CITATIONS AND BIBLIOGRAPHY

**DESCRIPTION:**

- A reference or citation is a description of any document from which you have taken information, e.g. a complete book, a chapter from it, a journal article, a newspaper article, a web page, or DVD etc
- There are various ways of writing references, depending on the institution's policy. For instance the referencing style approved by the college is adopted from the American Psychological Association (**APA**) manual which is currently in its 6$^{th}$ edition. Therefore, this guideline is based on the APA style only.

**Assignment**: ***Check on Harvard referencing style and note the differences with APA referencing style.***

## What is “Citing”?

“Citing” a reference is the act of recording it. It is made in two places:

1. a brief entry for each source in the text of your work, which then leads your reader to …
2. your source, in full, at the end of your work, as an alphabetical reference list.

As the list is in alphabetical order, it is easy to pick out the required author's work.

# IMPORTANCE OF CITATION & REFERENCING

- It is both a legal requirement and academic practice to provide references to guide your reader to the sources you have used
- To support the arguments you are making
- To demonstrate the breadth of your research,
- To credit the established work of others.

**NOTE:**

- Failure to acknowledge your sources is likely to lead to a suspicion of **plagiarism – i.e. trying to pass off someone else's work as your own: it is a form of cheating.**
- Incomplete or inaccurate referencing also reflects badly on your work

# How it works

- There are 2 parts:

1. **Author + Date in your text.**
2. **Full reference in the Reference List**

- Whoever you cite in your **text has to match your reference list** as the list is in alphabetical order, normally by author. **It must be in alphabetical order**
- e.g. *In your text: …Marieb and Hoehn (2007)… leads to the reference list and finds:* Marieb, E. N., and Hoehn, K. (2007). Human anatomy and physiology. 7th International ed. San Francisco: Benjamin Cummings.

# SECONDARY REFERENCING

- If you refer to a document which you **DID NOT read,** but which was cited (referenced) by somebody else whose work you **DID read**, you must make this clear. When you compile your reference list you must only cite the work in which you read it. Try to avoid this type of reference as you cannot always check the original and are relying on interpretation by others.

**Examples:**

- Dunn (1988), as cited by Campbell and Muncer (1998), believed that …

or Dunn (1988) revealed that ….. (cited in Campbell and Muncer, 1998)

**NB: your reference list will include the full details of the Campbell and Muncer work, but no mention of Dunn‟s.**

# How to Put References into the Text of your Essay / Report

**Author/s and Date**

- ❑ For each reference you make in the text of your essay, you need to provide:
- The **authority:** usually surname (family name) of the author(s), maybe a corporate author
- The date it was published.

Example: (Nursing and Midwifery Council, 2008)

- If you include the author's name at the beginning of the sentence statement, only the date needs to be in brackets**.**
  - **Example: … Hartley (1999) declared that …**
- If it is not part of your sentence, both the name and date must be in brackets, separated by a comma.
  - **Example: … although other authors have denied this (Hartley, 1999).**

❑ If there are **two authors:**

- **Example:** In the much acclaimed work on the subject by **Martin and Frost (2001),** it is clear that …
- For **three authors or more,** it is usual to use the Latin ***et al*** *(meaning* ***"and others")*** after the name of the first author.
- **Example: … Anderson et al (2003)** concluded that …

**<u>Multiple references to the same author</u>**

- If you cite different documents by the same author which were published in the same year, to distinguish between them add the letters a, b, c, etc. in lower case after the year. Repeat in the reference list.
- **Example: … (Williamson, 2001a), (Williamson, 2001b) etc. …**

## Quotations in the Text

- If you quote the exact words directly from a text you must use **quotation marks** to indicate this**.**
- The author(s) and date must be stated, and if possible the page number (or at least the chapter heading e.g. Chapter 6) from which the quote is taken.

**NB:** Page numbers for books are not included in the Reference List

- **Example: … Jackson (2004) declared that "This is the finest example of postmodernism …"**
- If page numbers are in separate sequences and therefore duplicated e.g. different issues of a journal throughout the year, or sections of a book, you must include the issue or section number or name.

# Listing Your References at the End of Your Work

**NOTE:** Your list should have both printed and electronic sources in one single **alphabetical sequence.**

**Order of referencing textbooks**

- Surname of author(s), comma, initial(s), full stop
- Year of publication (in brackets)
- The title (**in italics** with only the first letter of first word capitalized), colon between short and secondary/sub title, full stop.
- The edition (if other than the first), full stop
- Place of publication (the first city or town) followed by a colon
- Publisher's name, full stop

**Example:**

- Macionis, J. J. and Plummer, J. (2008) *Sociology: a global introduction.* 4th ed. Harlow: Pearson, Prentice Hall.
- **Mugenda A. & Mugenda O.,(1994).Research Methodology. Oxford Publishers: Kenya.**

**E-books (Electronic Books)**

As above examples, except for certain additions. You need to include:

- Author (or editor) surname and initials
- Year (in brackets). Always use the publication date of the version being used.
- Title of book (and any subtitle) - *italics or underlined. Only initial letter capitalized.*
- Edition (other than the first)
- Place of publication (of printed original - if available) followed by a colon(:)

- Publisher's name.
- Available from: (i.e. the e-book service you used), URL (web address)
- (Date accessed).

**Example:**

- White, R. and Downs, T. E. (2005) How computers work, 8th ed. Indianapolis: Que. [Online]. Available from: Safari Tech Books Online. http://0-proquest.safaribooksonline.com [Accessed: 16 August 2007].

# JOURNAL ARTICLES

- For journals, details are normally on the contents page and usually at the top or bottom of every page of each article. You need to include:

- Surname of the author(s), comma, initial(s), full stop
- Year of publication in brackets
- Title of the ARTICLE ( only first word with capitalized initial letter, unless proper name), comma
- Title of the JOURNAL (*in italics), comma*
- Volume number, issue or part number (in brackets), comma
- First and last pages of the article separated by a hyphen and indicated by the abbreviation “pp.”

# EXAMPLE

- Smith, A. and Jack, K. (2005). Reflective practice: a meaningful task for students, Nursing Standard, 19 (26), 33-37.
- Morrison, C. and Jutting, J. (2005) Women's discrimination in developing countries: a new data set for better policies, World Development. July, 33 (7), 1065-1081. [Online]. Available from: Science Direct. http://sciencedirect.com [Accessed 31 July 2005].

**Papers presented at a conference**

Mugenda, O. (1999) Redefining and Actualizing the Research Mission in African Universities. Paper presented at the BOLESWA Educational Research Symposium, Maseru, Lesotho, July, (1999).

**Newspaper article**:

Ngw'eno, H.B. (1993, September). Multiply and fill the earth. The Weekly Review, pp 15 – 17.

# INTRODUCTION TO BIOSTATISTICS

- What is statistics?

✓ Statistics is the summary of information (data) in a meaningful fashion, and its appropriate presentation.

✓ Statistics is the postulation of a plausible model explaining the mechanism that generates the data, with the ultimate goal to extrapolate and predict data under circumstances beyond the current experiment

- Bio-statistics is the segment of statistics that deals with data arising from biological processes or medical experiments

- Two broad branches in statistics

**1. Descriptive statistics**

Once data has been collected, normally the step that follows is to summarize the data, if possible, with one or two summary statistics. Summary or descriptive statistics describe the original data set (the set of responses for each question) by using just one or two numbers – typically an average and a measure of dispersion.

**2. Inferential Statistics**

This is the branch of statistics that makes use of sample data to make generalization concerning the population parameters. Here theoretical distributions become handy.

# Errors in statistical inference

- **Type I error** (or, **error of the first kind**) and
- **Type II error** (or, **error of the second kind**)

Are precise technical terms used in statistics to describe particular flaws in a testing process, where a true null hypothesis was incorrectly rejected (**Type I error**) or where one fails to reject a false null hypothesis (**Type II error**).

# TYPE 1 ERROR

- Occurs when the null hypothesis ($H_0$) **is true**, **but is rejected**. It is **asserting something that is absent**, a **false hit**.
- Occurs when the researcher rejects the null hypothesis when it's actually true.
- A type I error may be compared with a so called *false positive* (a result that indicates that a given condition is present when it actually is not present) in tests where a single condition is tested for.
- A Type I error is committed when we fail to believe a truth. In terms of folk tales, an investigator may be "crying wolf" without a wolf in sight (raising a false alarm) ($H_0$: no wolf).

# TYPE 2 ERROR

- Occurs when the null hypothesis is false, but it is erroneously accepted as true. It is **missing to see what is present**, a **miss**.
- Occurs when the researcher fails to reject the null hypothesis when it’s actually false.
- A type II error may be compared with a so-called *false negative* (where an actual 'hit' was disregarded by the test and seen as a 'miss') in a test checking for a single condition with a definitive result of true or false.
- A Type II error occurs when we believe a falsehood. In terms of folk tales, an investigator may fail to see the wolf ("failing to raise an alarm“).

# RELIABILITY AND VALIDITY IN RESEARCH

## Reliability

- **Def:** is the extent to which an experiment, test, or any measuring procedure **yields the same result on repeated trials.**
- Without the agreement of independent observers able to replicate research procedures, or the ability to use research tools and procedures that yield consistent measurements, researchers would be unable to satisfactorily draw conclusions, formulate theories, or make claims about the generalizability of their research.

## Validity

- Validity refers to the degree to which a study **accurately reflects or assesses** the specific concept that the researcher is attempting to measure.
- While reliability is concerned with the accuracy of the actual measuring instrument or procedure, validity is concerned with the study's success at measuring what the researchers set out to measure.

# *The End*