Lecture 1: Problem Identification

Problem Identification

Identifying and clearly defining a research problem is one of the most important steps in the research process. While there are countless problems in fields such as family planning and health programs, not every problem is suitable for research. A well-defined research problem leads naturally to the statement of objectives, formulation of hypotheses, identification of key variables, and selection of appropriate methodology. Conversely, a poorly defined problem leads to confusion and weakens the entire research effort.

All research is set in motion by the existence of a problem. A problem is a perceived difficulty, a sense of discomfort with the way things are, or a discrepancy between what is and what should be. However, not all problems require research. A potential research situation arises when the following three conditions exist:

1. A perceived discrepancy between what is and what should be.
2. A question about why the discrepancy exists.
3. At least two possible and plausible answers to the question.

If there is only one possible and plausible answer, then a research situation does not exist.

Example of a Nonresearch Problem

Problem Situation: A recent survey in District A found that 1,000 women were continuous users of contraceptive pills. Last month's service statistics indicate that none of these women are using contraceptive pills.

Discrepancy: All 1,000 women should be using contraceptive pills, but none are.

Problem Question: What factor or factors are responsible for 1,000 women discontinuing their use of contraceptive pills?

Answer: A monsoon flood has prevented all new supplies of pills from reaching District A, and all old supplies have been exhausted.

In this example, the reason for the problem is already known, so there is no need to conduct research on pill discontinuation. However, research might be needed on why the supply system failed during the monsoon.

Example of a Research Problem

Problem Situation: District A is always flooded during the monsoon season. To address this, a new supply logistics system was established: each pill user is given a four-month supply before the monsoon, and small motorboats are available to transport new supplies to distribution centers. Despite these measures, this year there are no pill supplies in District A.

Discrepancy: The new logistics system should assure a continuous supply of pills, but this year there are no supplies.

Problem Question: Why has the new supply logistics system been incapable of delivering contraceptive pills to users?

Possible Answers:

1. An order for new pill supplies was not placed in time before the monsoon rains.
2. The riverboats used to transport the supplies are out of order.
3. Field-workers were not informed about the new system and failed to distribute a four-month supply before the monsoon.

Here, multiple plausible explanations exist, making this a suitable research problem.

Example of a Complex Research Problem

Problem Situation: A recent family planning survey revealed large differences between villages in contraceptive prevalence rates. Despite all villages receiving the same level of services, some have rates as high as 80%, while others are as low as 6%.

Discrepancy: All villages should have approximately the same contraceptive prevalence rate, but there is great variation.

Problem Question: What factors are responsible for the areal variation in contraceptive prevalence rates?

Possible Answers:

1. Socioeconomic differences between villages (e.g., type of community, religion, access to markets, availability of facilities).
2. Differences in institutional support for family planning (e.g., support from local leaders, presence of Mothers' Clubs).
3. Differences in effectiveness and motivation of village-level health and family planning workers.

In such cases, the researcher must devote considerable time and attention to identifying and defining the problem, focusing the research on the most important aspects.

Lecture 2: An Introduction to Research Methodology

Meaning of Research

Research is the systematic gathering and analysis of data and information for the advancement of knowledge in any subject. It aims to answer intellectual and practical questions through the application of systematic methods.

According to Webster's Collegiate Dictionary, research is "studious inquiry or examination; especially: investigation or experimentation aimed at the discovery and interpretation of facts, revision of accepted theories or laws in the light of new facts, or practical application of such new or revised theories or laws."

Research is often described as a movement from the known to the unknown, driven by human inquisitiveness. This instinct leads us to probe and understand what is not yet known, making research the mother of all knowledge.

Clifford Woody defines research as the process of defining and redefining problems, formulating hypotheses or suggested solutions, collecting, organizing, and evaluating data, making deductions and reaching conclusions, and carefully testing those conclusions to determine whether they fit the formulated hypothesis.

According to D. Steiner and M. Stephenson, research is "the manipulation of things, concepts or symbols for the purpose of generalizing to extend, correct or verify knowledge, whether that knowledge aids in construction of theory or in the practice of an art."

In summary, research is an original contribution to the existing stock of knowledge, advancing it through study, observation, comparison, and experiment. It refers to the systematic method of enunciating the problem, formulating a hypothesis, collecting and analyzing data, and reaching conclusions either as solutions to the problem or as generalizations for theoretical formulation.

Objectives of Research

The main purpose of research is to discover answers to questions through scientific procedures. The objectives of research can be broadly grouped as follows:

1. To gain familiarity with a phenomenon or achieve new insights (exploratory or formulative research).
2. To portray accurately the characteristics of a particular individual, situation, or group (descriptive research).
3. To determine the frequency with which something occurs or is associated with something else (diagnostic research).
4. To test a hypothesis of a causal relationship between variables (hypothesis-testing research).

Utility of Research

Research is extensively useful for managers in planning, forecasting, coordinating, motivating, controlling, and decision-making. Academic research supports academic objectives, while social research aids in social prediction, enlightenment, welfare, growth, cohesion, and control.

Key utilities of research include:

- Aiding decision-making and problem identification.
- Facilitating analysis, evaluation, and interpretation of business or social environments.
- Providing a basis for innovation and product development.
- Establishing relationships between variables and functional areas.
- Aiding forecasting and all managerial functions.
- Supporting economic utilization of resources.
- Assisting in market analysis and management information systems.
- Informing policy and strategy formulation.

Research Methods

Research methods are all the techniques used for conducting research. They include:

1. Methods concerned with data collection, used when available data is insufficient.
2. Statistical techniques for establishing relationships between data and unknowns.
3. Methods for evaluating the accuracy of results obtained.

Research methodology is the science of studying how research is done scientifically. It involves understanding the steps adopted by researchers, the logic behind them, and not just the techniques themselves. Researchers must know not only how to use certain methods but also when and why to use them, and what assumptions underlie them.

Need of Research Methodology

It is necessary for a researcher to design a research methodology for the chosen problem. Even if research methods are the same for two problems, the methodology may differ.

A good methodology addresses:

1. Why a particular research study is undertaken.
2. How the research problem is formulated.
3. What types of data are collected.
4. What particular methods are used.
5. Why a particular technique of data analysis is chosen.

Studying research methods trains one to apply them to problems. Studying research methodology provides the necessary training to choose appropriate methods, materials, and scientific tools for the problem at hand, and to justify those choices.

Lecture 3: Approaches and Types of Research

Deduction: Testing Theory

Deduction is a research approach used to test a theory. It involves developing a theory and then subjecting it to a rigorous test through the following stages:

1. Deducing a hypothesis from the theory
2. Expressing the hypothesis in operational terms
3. Testing the operational hypothesis
4. Examining the specific outcomes of the enquiry
5. Modifying the theory in light of the findings, if necessary

This process is cyclical: the revised theory is verified by returning to the first step and repeating the cycle.

Induction: Building Theory

Inductive research seeks to build theory from empirical evidence. The researcher tries to understand the nature of a problem, gathers quantitative and qualitative data, and analyzes them to draw conclusions. The result of this analysis is the formulation of a theory. Thus, in an inductive approach, theory is built from the data, whereas in deduction, data is collected to test a theory.

Characteristics of a Scientific Method

The main characteristics of the scientific method are:

1. **Verifiability:** Conclusions can be verified at any time. The phenomenon under investigation must be observable or measurable, even if only indirectly (e.g., via interviews).
2. **Generality:** Laws derived are universal in their applications and not limited to individual cases.
3. **Predictability:** Results can be predicted with sufficient accuracy, based on established cause-and-effect relationships and the stability of causative factors.

4. **Objectivity:** Results should be free from the investigator's own views or biases; conclusions should be the same for all observers.
5. **System:** Scientific studies follow an accepted, systematic mode of investigation; results from haphazard methods, even if true, are not considered scientific.

Types of Research

Research can be divided into two broad types relative to its purposes: **applied** and **fundamental** research.

1. Applied Research

Applied research is conducted in response to a specific problem that requires a solution. Its major purpose is to answer practical and useful questions about policies, programs, projects, procedures, or organizations. Business executives often commission applied research to find solutions that can be implemented to rectify a problem situation.

Applied research is also called *decisional research* because it is concerned with knowledge that has immediate applications.

Example: The Dairy Development Corporation (DDC) needs to improve productivity to remain competitive. Two strategies are considered: (a) improving all existing brands, or (b) focusing on new brand development. Each has pros and cons. Research is needed to determine which strategy best fits DDC's capabilities and resources. This illustrates the need for applied research to work out a strategy based on strengths and weaknesses, with the goal of making decisions and formulating policies.

The defining quality of applied research is that its results can be applied directly to a specific situation, often resulting in actionable recommendations.

2. Fundamental Research

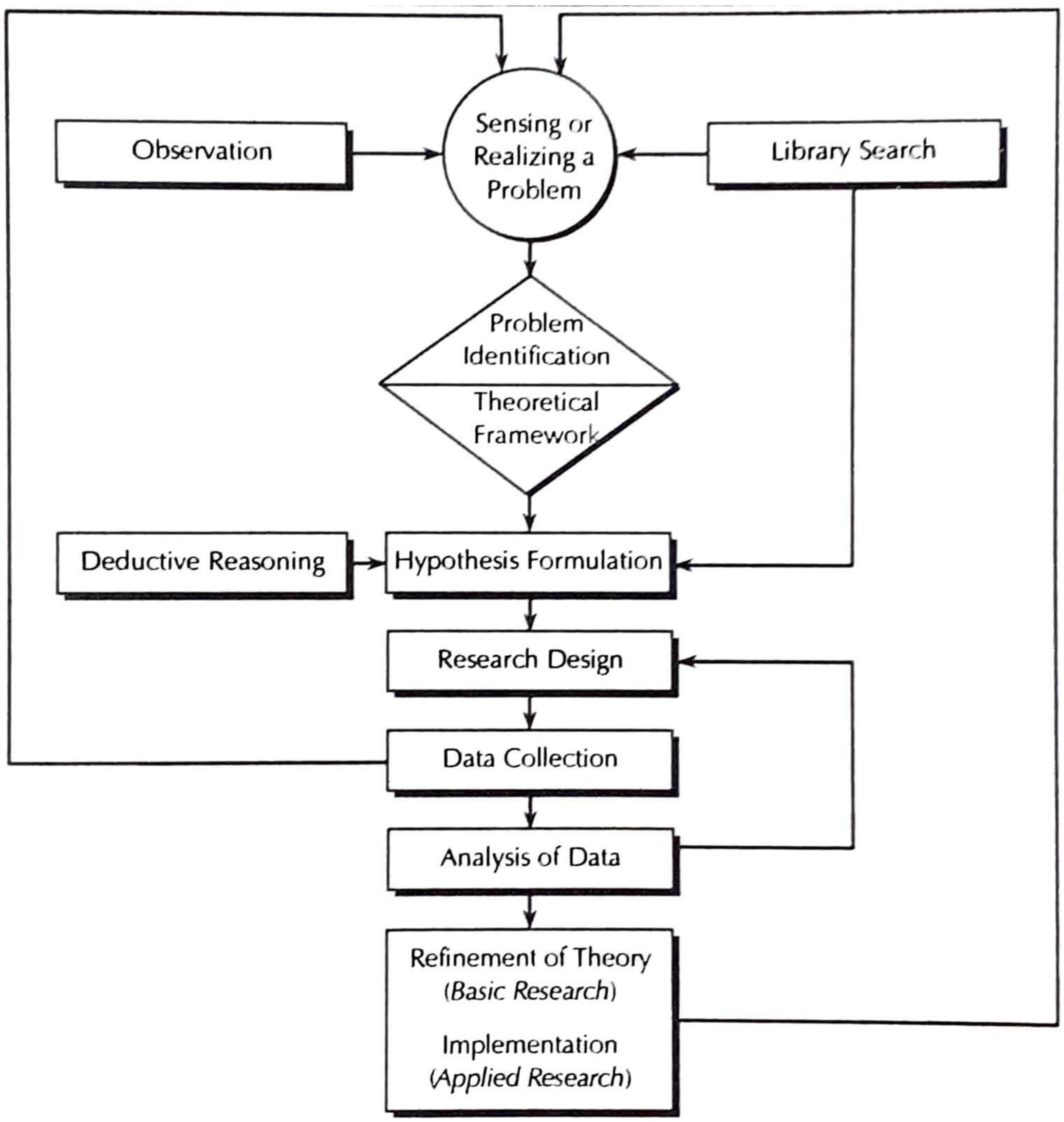
Fundamental (or basic/pure) research is undertaken to improve understanding of problems that commonly occur in organizational settings, without any practical end-use in mind. Its purpose is to add to generalizable knowledge and to build theories based on research results.

Fundamental research is also concerned with developing, examining, verifying, and refining research methods and techniques. It aims to advance knowledge and to identify and explain relationships between variables.

Example: At an HRD Managers' Conference in Kathmandu, participants discussed employee socialization, training, and commitment. It was observed that productivity of workers over 40 does not improve with socialization and training, though their organizational commitment is higher. Questions arise: Why does this occur? What factors are responsible? What type of training is effective for different age groups? To answer these, basic research is needed to increase knowledge about employee commitment and training, not necessarily to solve an immediate problem. The researcher would design studies in different work settings to observe the effects of socialization and training on productivity and commitment across age groups.

The main purpose of fundamental research is to advance the level of scientific knowledge.

Scientific Research Process



Scientific research is systematic and follows the steps of the scientific method. From the inception of a research idea to the final report of results, the research process has several crucial steps. However, these steps do not provide a rigid pattern into which you must force your thinking. Thinking simply cannot be scheduled. An investigator does not tackle one step at a time, complete that process and then move on to the next step. Some steps can go simultaneously. Others need proper sequencing and logical arrangement.

Lecture 4: The Scientific Research Process

Scientific Research Process

Scientific research is a systematic process that follows the steps of the scientific method. From the inception of a research idea to the final report of results, the research process involves several crucial steps. However, these steps are not rigidly sequential; some may occur simultaneously or require iteration and logical arrangement. The steps may vary depending on the subject and the researcher, and are often interdependent.

There are eight main steps in the scientific method, covering the full spectrum of a research endeavor—from problem formulation to the refinement of theory or practice. These steps are:

1. Sensing or Realizing a Problem

The first step in any scientific inquiry is to identify an issue to study. This can arise from observing a situation or sensing that something is not functioning as it should. At this stage, the exact nature of the problem may not be clear, but there is an awareness that something is amiss.

2. Problem Identification

With increased awareness, the researcher focuses on the problem and its associated factors by searching for more information. The aim is to identify what exactly are the issues in the situation. As the saying goes, "a problem well defined is a problem half solved."

3. Theoretical Framework

The next step is to logically integrate the information so that the reasons for the problem can be conceptualized. Critical variables are examined, and their associations are identified. By putting all variables and their relationships together, a theoretical framework is developed.

4. Hypothesis Formulation

Hypotheses are formulated as logical conjectures about the relationships between two or more variables, expressed as testable statements. Hypotheses are drawn from the theoretical framework and provide specific answers to the research questions. They are especially useful in quantitative research where statistical analysis is performed.

5. Research Design

At this stage, the researcher devises a plan or strategy for conducting the research. The research design outlines what information is needed and how it will be obtained and analyzed. The choice of design depends on the nature of the research problem and objectives.

6. Collection of Data

Data collection, or fieldwork, involves administering research instruments (such as questionnaires, interview schedules, or observation schedules) to gather data. The procedures used depend on the research design and data sources. This step is crucial to the success of the research project.

7. Data Analysis

After data collection, the data must be summarized and analyzed. Data analysis includes editing, coding, tabulating, and applying statistical techniques. Descriptive statistics provide a summary, while inferential statistics assess the reliability of the data and test the hypotheses.

8. Refinement of Theory or Practice

The final step involves interpreting and generalizing the findings, contributing to the broader body of knowledge. In applied research, this may involve proposing specific strategies to solve the identified problem. Through research, existing theories or practices are refined and modified.

Cyclical Nature of Scientific Research

The scientific research process is cyclical. Research often starts with a problem and ends with a tentative empirical generalization, which then becomes the starting point for further research. This cycle continues indefinitely, reflecting the ongoing accumulation and self-correction of scientific knowledge.

Example: A Case Study

A car dealer received complaints from users about rattling sounds in the dashboard and rear passenger seat after a few thousand kilometers of driving. The dealer:

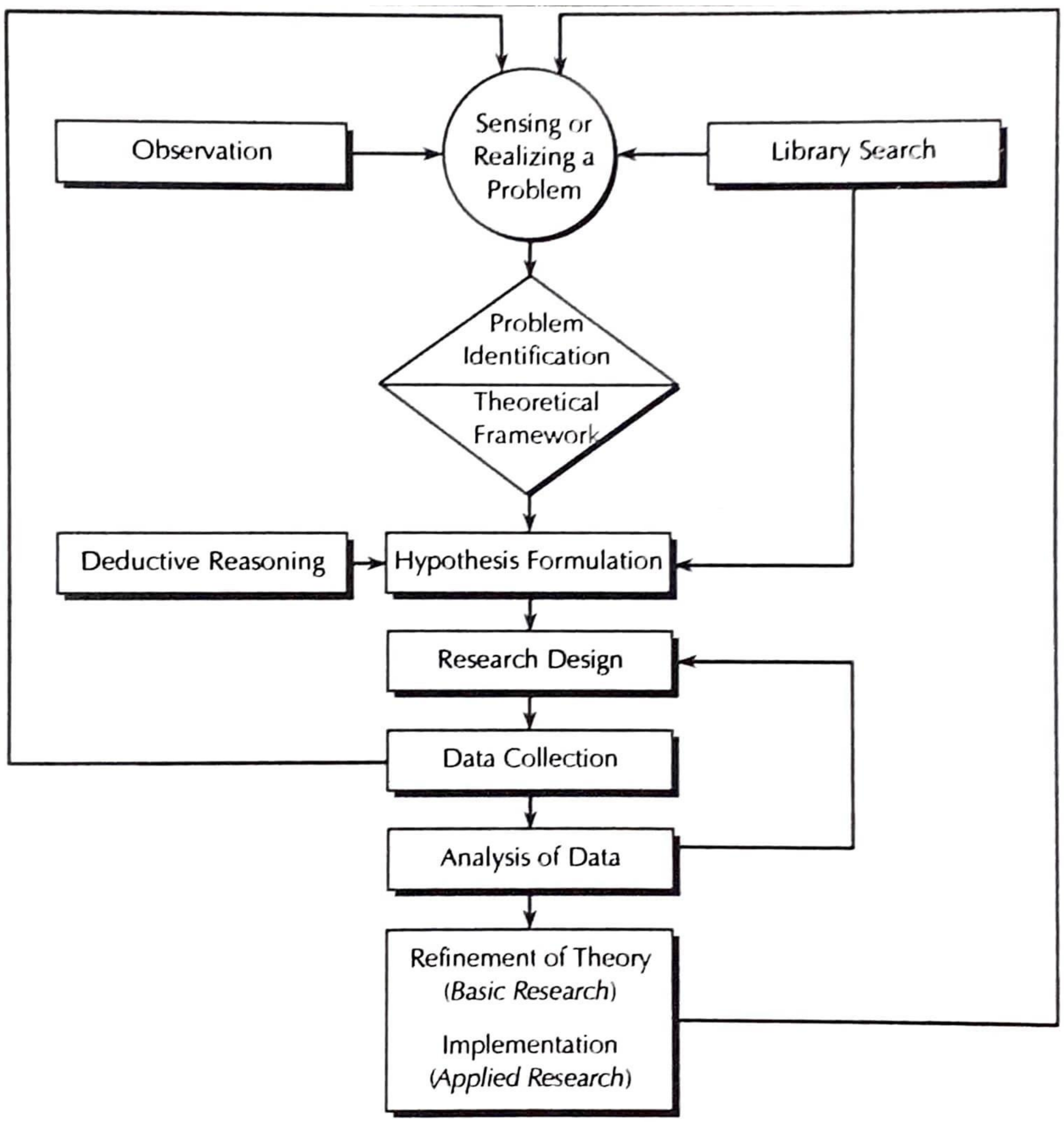
- Gathered information from company workers to identify influencing factors.
- Formulated the problem and generated hypotheses.
- Constructed a checklist and collected data from a representative sample of cars.
- Analyzed the data, interpreted the results in light of the hypotheses, and reached conclusions.

This example illustrates a systematic, step-wise scientific method of inquiry, leading to reliable conclusions.

Key Features

- The research process is systematic and follows a logical sequence.
- Steps may be revisited or occur simultaneously.
- The process is cyclical and self-correcting, with each generalization leading to further research.

Scientific Research Process



Scientific research is systematic and follows the steps of the scientific method. From the inception of a research idea to the final report of results, the research process has several crucial steps. However, these steps do not provide a rigid pattern into which you must force your thinking. Thinking simply cannot be scheduled. An investigator does not tackle one step at a time, complete that process and then move on to the next step. Some steps can go simultaneously. Others need proper sequencing and logical arrangement.

Lecture 5: Case Study, Research Language, and Problem Identification

Case Study: The CIO Dilemma

Observation: The Chief Information Officer (CIO) of a firm notices that the newly installed Management Information System (MIS) is not being used by middle managers as much as expected. Managers often seek help from the CIO or other "computer experts" or make decisions without adequate information. The CIO recognizes a problem.

Information Gathering through Informal Interviews: Informal discussions with middle managers reveal that many lack understanding of what MIS is, what information it provides, and how to access or use it.

Literature Survey: The CIO searches the Internet and finds that many middle managers, especially older ones, are not familiar with computers and experience "computer anxiety." Lack of knowledge about MIS offerings is a main reason for underuse.

Formulating a Theory: Based on these insights, the CIO develops a theory incorporating factors contributing to the lack of MIS use among managers.

Hypothesizing: From the theory, the CIO generates hypotheses for testing, such as: *Knowledge of the usefulness of MIS would help managers to put it to greater use.*

Data Collection: A short questionnaire is developed to assess factors influencing MIS use, including knowledge of MIS, types of information provided, access methods, comfort with computers, and frequency of MIS use in the past three months.

Data Analysis: The CIO analyzes the questionnaire data to identify factors preventing managers from using MIS.

Deduction: Based on the analysis, the CIO concludes that certain factors hinder MIS use. Actions such as organizing training seminars for managers on computers and MIS are recommended.

Research Language: Key Terms

- **Theory**
- **Conceptualization**

- **Operationalization**
- **Variables**
- **Hypotheses**
- **Assumptions**
- **Population**
- **Sample**
- **Validity**
- **Reliability**
- **Data**
- **Research process**
- **Summary**
- **Methodological queries**
- **References**

Concepts and Constructs

- **Constructs** are abstract concepts such as happiness, satisfaction, IQ, or morality. They represent real phenomena that cannot be directly observed but are useful for study.
- **Concepts** are abstractions formed by generalization from particulars (e.g., "weight" as a concept for heaviness or lightness).
- Constructs can be conceptually defined and are meaningful in theoretical terms. They are often not directly observable.
- Examples of concepts: Income, Age, Education Level, Number of Siblings.
- Concepts can be measured through direct observation (e.g., weight, height, hair color) or indirect observation.

Research Problem Identification and Formulation

Formulating and selecting a research problem is the first and foremost step in research. The entire research work can be conducted only if the problem is precisely identified, which may require:

- Basic knowledge in the field, developed through discussion with experts, literature review, or continued activity in the area.
- Careful selection of the research problem, often with guidance from a supervisor.

- Recognizing and defining the exact nature and dimensions of the problem.

Steps in Identifying a Research Problem

1. Determine the field of research of interest.
2. Develop mastery in the area or field of specialization.
3. Review recent research conducted in the selected area.
4. Based on the review, select the priority field of study.
5. Use analogy, insight, or personal experience to identify a problem; consult with supervisors or experts as needed.
6. Pinpoint the specific aspect of the problem to be investigated.

Ways of Understanding a Research Problem

Selecting a suitable research problem is not easy and requires serious consideration. Approaches include:

- Discussion with colleagues, research guides, or experts.
- Intensive reading of available literature (research abstracts, journals, handbooks, international abstracts).
- Personal experience in the field.
- Studying the field intensively to identify specific problems.
- Considering new innovations, technological changes, and curricular developments for new research opportunities.
- Consulting supervisors and experienced individuals in the field.
- Reviewing suggestions for further research found in previous research reports.

Lecture 6: Steps and Criteria in Research Problem Formulation

Steps in Research Problem Formulation

In scientific inquiry, turning a general topic into a specific research problem is called research problem formulation. The researcher must thoroughly understand the problem and rephrase it in meaningful, analytical terms. The following are suggested steps for formulating a research problem:

- Statement of the problem in a general way
- Understanding the nature of the problem
- Surveying the available literature
- Developing ideas through discussion
- Rephrasing the research problem

Criteria of a Good Problem

Fred N. Kerlinger defines a good problem as an inquisitive sentence that asks what relation exists between two or more variables. The research questions, objectives, and hypotheses all rest on the problem statement. An ideal research problem should meet these three main criteria:

1. **Expresses a relation between two or more variables:** The topic should reflect relationships to be examined.
2. **Stated rigidly and unambiguously:** Ambiguity in the problem statement can affect research design, process, and results.
3. **Implies possibilities of empirical testing:** The problem must be testable within the research design and available resources.

A good research problem should be linked with the research design and consider available facilities and capabilities.

Problem Statement

Developing a problem statement involves generating and exploring an issue, then determining worthwhile research questions. A problem statement should:

- Raise a question about a relationship between variables
- Clearly state and explain the relationship between variables
- Suggest a method for researching the question

A problem statement can be written in either declarative or interrogative form.

Examples:

- *Declarative:* Factors contributing to the excessive absenteeism among Nepalese workers.
- *Interrogative:* Why is absenteeism so high among workers in Nepalese organizations?

Other examples of well-defined research problems (interrogative form):

- To what extent do age, education, length of service, level of earning, and place of residence of employees predict occupational aspirations?
- Do long work hours, lack of development opportunities, and discrimination account for the lack of inward mobility of women in civil service?
- Can cultural differences account for the differences in the nature of hierarchical relationships between supervisors and subordinates?

Example from Research Literature

A review of research on leadership and age (Kabacoff & Stoffey, 2001) highlights the importance and controversy in examining the relationship between these variables. Older workers remain employed longer and work alongside younger members in various roles. Modern organizations now have multigenerational teams, and both age groups contribute unique strengths. Leadership is no longer the exclusive domain of older people; flatter organizations encourage interaction between all ages. Understanding and leveraging multigenerational diversity is necessary for high performance in organizations.

Kabacoff & Stoffey (2001) and Kakabedse et al. (1998) found that age and time-related dimensions shape the attitudes and behaviors of senior leaders. In Australia, three leader profiles were identified: radicals (youngest, 26-35 years), bureaucrats (45-55 years), and team players (oldest, 56+ years). Older workers were mature and long-term oriented, while younger workers were competitive, energetic, and open in management style.

Research Questions

Research is intended to help us learn something new. Research questions are the most important element for effective execution of research. They are formulated in interrogative form and describe the ideas in the research objectives. Research questions guide the data to be collected, the research design, the methods, the analysis tools, interpretation, and reporting. They should address:

- **What?** (description)
- **Why?** (explanation)
- **How?** (intervention or change)

Without clear research questions, the research process may lack direction and face challenges.

Types of Research Questions

Generally, there are three types of research questions:

- **What questions:** Concerned with description.
  Examples: What are the types of community involved in skill transformation? What are the socio-economic characteristics of the community? What are the needs of the community?
- **Why questions:** Concerned with explaining causes or relationships.
  Examples: Why do drug abusers commit thefts? Why does stressful living result in heart attacks? Why do some people use a product while others do not?
- **How questions:** Concerned with bringing about change and outcomes.
  Examples: How has the caste system changed in Nepal in the last century? How does technology create unemployment? How do maternal and child health services affect infant mortality?

Other types proposed by Lin (1993) include Who, Where, How many, and How much.

Identification of Research Question

The main purpose of formulating research questions is to define the scope of the research—what is to be studied and to what extent. Neuman (1997) suggests the following techniques for developing research questions:

- Record all questions that arise after literature review, discussions, or reflection.
- Review and delete questions beyond the study's scope to remove overlap.
- Classify questions by their nature: What, Why, How, etc.
- Examine the scope of questions for feasibility within available time and resources.
- Separate major/key questions (core of the research) from subsidiary questions.

Detailed Summary of Lecture 7: Hypothesis in Research

1. Definition of Hypothesis

A **hypothesis** is a proposed relationship between two or more variables. It is essentially a provisional idea or working assumption that requires evaluation through investigation.

Example: *Political participation increases with education.* This implies a direct relationship between two variables.

Scholarly definitions:

- **Fred N. Kerlinger and H.B. Lee (2000):** "A hypothesis is a conjectured statement that implies or states a relationship between two or more variables."
- **John W. Creswell (2014):** "A hypothesis is a formal statement that presents the expected relationship between independent and dependent variables."

A good hypothesis:

- Clearly states the expected relationship or difference between variables.
- Defines variables operationally and in measurable terms.

2. Functions of a Hypothesis (Kumar, 2011)

- Provides focus by specifying what aspects of the research problem to investigate.
- Guides data collection by indicating what to collect and what to ignore.
- Enhances objectivity in the study.
- Contributes to theory formulation by determining what is true or false.

3. Hypothesis Formulation

Hypotheses can be derived from:

- General culture
- Scientific theory or past research
- Personal experience

- Discussions or intuition

Two justifications:

- **Logical:** Based on theoretical reasoning.
- **Empirical:** Based on reference to existing research.

A hypothesis is more operational than the research problem and is usually derived directly from it.

4. Deductive and Inductive Reasoning

4.1. Deductive Reasoning

Deduction moves from general theory to specific case.

Example 1:

- All books have pages.
- This is a book.
- Therefore, it has pages.

Example 2:

- Lung cancer is caused by smoking.
- John has lung cancer.
- Therefore, John was a smoker.

4.2. Inductive Reasoning

Induction moves from specific observations to general conclusions.

Researchers use both deduction and induction to:

- Organize facts
- Describe results
- Develop relationships
- Suggest new research

5. Hypothesis vs. Research Problem

- A research problem is posed as a question.
- A hypothesis is a testable solution to that question.
- Problems can't be directly tested; hypotheses can be.

Example:

- **Problem:** What is the relationship between population growth in Kathmandu before and after the introduction of family planning?
- **Hypothesis:** There is a significant difference in the population growth in Kathmandu between when family planning was first introduced and five years later.

6. One-sided vs. Two-sided Hypotheses

- **One-sided hypothesis:** Predicts a specific direction (e.g., increase or decrease).
- **Two-sided hypothesis:** Suggests a difference but not the direction.

7. Formats for Stating Hypotheses

7.1. Correlation

- "There is a significant relationship between Variable A and Variable B for Group 1."

7.2. Difference Between Means

- "There is a significant difference between mean levels of Variable A for Group 1 and Group 2."

7.3. Difference Between Frequencies

- "There is a significant relationship between Group 1 and Group 2 for Variable A."
- "There is a significant difference between Group 1, 2, and 3 for Variables A and B."

8. Conclusion

This lecture provides a comprehensive understanding of the concept of hypothesis in research. It discusses its definition, function, formulation, types of reasoning used in its development, and proper formats for stating hypotheses. Mastery of these elements is essential for constructing valid, testable research.

Detailed Summary of Lecture 8: Types of Hypothesis

1. Types of Hypothesis

1.1. 1. Descriptive and Relational Hypotheses

Descriptive Hypotheses state the existence, size, form, or distribution of a single variable. They do not show relationships between variables and thus are not ideal for scientific research hypotheses.

Examples:

- Tribhuvan University is experiencing budget difficulties.
- The Hetauda–Narayangadh sector of the East-West Highway has higher-than-average accident rates.
- Stockholders of Nepal Development Bank favor bonus dividends.

These can be reframed as **research questions**:

- What is the extent of budget difficulties at Tribhuvan University?
- Why is the accident rate higher in the Hetauda–Narayangadh sector?
- Why do bank stockholders prefer bonus dividends?

Relational Hypotheses describe the relationship between two or more variables and include:

- **Correlational Hypotheses:** Describe associations.
- **Explanatory (Causal) Hypotheses:** Imply causation.

Examples:

- Families with higher incomes spend more on recreation.
- Political participation increases with education.

1.2. 2. Explanatory Hypotheses

These state how one variable causes a change in another. Causal relations can be:

- **Unidirectional:** A affects B, but not vice versa.
- **Bidirectional:** A and B influence each other.

Examples:

- Increase in age leads to decreased organizational commitment.
- Productivity increases when workers are given incentive pay.

1.3. 3. Directional and Non-directional Hypotheses

Directional Hypotheses predict the specific direction of the relationship and are tested using a one-tailed test.

Examples:

- Younger workers are less motivated than older workers.
- Higher workload leads to lower job satisfaction.

Non-directional Hypotheses do not specify direction and use two-tailed tests. They are used when prior findings are conflicting or unavailable.

Examples:

- There is a difference in work attitudes between industrial and agricultural workers.
- No relationship between education level and occupational commitment.

1.4. 4. Null and Alternative Hypotheses

Null Hypothesis (H_0): Suggests no relationship or difference exists. It is the default hypothesis tested statistically.

Alternative Hypothesis (H_A): Suggests a relationship or difference exists, contrary to H_0.

Example:

- **H_A:** Productivity increases with incentive pay.
- **H_0:** No significant difference exists between productivity with or without incentive pay.

Statistical Form:

- **H_0:** $\mu_1 = \mu_2$
- **H_A:** $\mu_1 > \mu_2$ or $\mu_1 \neq \mu_2$

More Examples:

- H_0: No relationship between working conditions and job satisfaction.
  H_A: Improved working conditions increase job satisfaction.
- H_0: No gender difference in organizational commitment.
  H_A: Males are more committed than females.
- H_0: Pay and productivity are not related.
  H_A: Pay and productivity are positively related.

2. Stating the Null Hypothesis

- There is no difference between the means of two populations.
- The population means from which the samples were drawn are equal.

3. Criteria of a Good Hypothesis Statement (Mason and Bramble, 1997)

- Should be in declarative form.
- Clearly describes a relationship between two or more variables.
- Should be testable empirically.
- Must be limited in scope.
- Should be clear and precise with no ambiguity.
- Must specify the conditions under which it applies.
- Should be based on prior research, facts, or identified needs.

4. Linkage Between Research and Statistical Hypotheses

Research Hypothesis:

- Focused and derived from research objectives.
- Example: "Female students have higher grades than male students."

Statistical Hypothesis:

- Formally testable using statistics.
- Example: "The mean height of men is 175 cm."

Statistical testing usually focuses on rejecting the null hypothesis (H_0), not directly proving the alternative (H_A).

Detailed Summary: Lecture 9 – Literature Review

1 Introduction to Literature Review

A literature review is a systematic and critical summary of published works on a specific research topic. It provides insight into what others have studied, their findings, and how these findings relate to the current research.

Definitions from scholars (Cardesco & Gatner, Haywood & Wragg, Walliman) emphasize the processes of **summarizing, analyzing, classifying, and comparing** past research. The review prevents duplication and helps identify knowledge gaps.

2 Purpose of Literature Review

The primary goals include:

- Understanding how concepts are defined and measured by others.
- Identifying previous data sources and hypotheses.
- Discovering theoretical frameworks and research approaches.
- Recognizing consistencies, contradictions, and gaps in existing literature.
- Formulating hypotheses and guiding study design.

3 Need for Literature Review

- Demonstrates familiarity with the field.
- Justifies the need for current study by identifying gaps.
- Helps define theoretical and methodological focus.
- Prevents unnecessary repetition.
- Showcases critical understanding of the topic.

4 Types of Literature Review

- **Historical Review**: Tracks development of ideas over time.
- **Methodological Review**: Examines research techniques and their strengths.
- **Theoretical Review**: Focuses on conceptual or theoretical frameworks.
- **Integrative Review**: Summarizes and synthesizes overall knowledge.

Often, these types are combined in academic research.

5 Functions of Literature Review

- Avoids "reinventing the wheel".
- Gives credit to foundational work.
- Shows mastery of the research problem.
- Demonstrates critical evaluation and synthesis.
- Positions new research as a valuable contribution.

6 Encyclopedias as Sources

Useful encyclopedias include:

- **Encyclopedia Britannica** & **Britannica Online**
- **Encyclopedia of Social Sciences** & **International Encyclopedia of Social Sciences**
- **Encyclopedia of Education**
- **McGraw-Hill Encyclopedia of Science and Technology**
- **Business Encyclopedia and Legal Adviser**

7 Internet as a Source

The Internet provides:

- Quick access to primary research and bibliographies.
- Access to professional and academic organizations.
- Updated information not always found in libraries.

Note: Internet sources are often not suitable for citation in formal research.

8 Difference Between Reference and Bibliography

Reference

- Cited sources within the research text.
- Supports specific arguments.
- Based on primary sources.
- Common in theses and dissertations.

Bibliography

- All materials consulted, whether cited or not.
- Includes primary and secondary sources.
- Used in journal papers and broader research.

9 Key Differences Summary

Basis	**Reference**	**Bibliography**
Meaning	Cited sources in text	All consulted sources
Source Type	Primary	Primary and Secondary
Arrangement	Alphabetical/Numerical	Numerical
Use in Argument	Supports argument	Not used directly
Usage	Theses, Dissertations	Journal Papers, Research Work

Detailed Summary: Lecture 10 – Research Design

1 Introduction

A research design is a systematic plan or blueprint outlining how a research study will be conducted. It addresses the approach to the problem, data collection methods, strategies for analysis, and tools used.

Figure 1: Structure of a Research Design

2 Definitions

- **Fred N. Kerlinger (1986):** Research design is the plan, structure, and strategy of investigation to obtain answers to research questions.
- **John W. Creswell (2011):** Research designs are plans and procedures that span from assumptions to methods of data collection and analysis.
- **William Zikmund (2013):** Research design is a master plan specifying methods and procedures for collecting and analyzing needed information.

3 Essential Elements of a Research Design

- Provides a structured plan for the research process.
- Guides data collection, instrumentation, and sampling.
- Acts as a strategy to generalize findings from a sample to the population.
- Involves decisions about strategy (experimental/non-experimental), setting, instruments, and statistical methods.

4 Core Elements of Research Design

(a) Problem definition

(b) Methodology

(c) Data gathering

(d) Data analysis

(e) Report writing

research_process_cycle.png

Figure 2: Cycle of Core Elements in Research Design

5 Preparation of Research Design

According to Oliver (2011), a good design answers:

- What type of data is needed?
- Where and how will it be collected?
- What instruments and procedures will be used?
- Who are the data sources?
- Is permission required to collect data?
- When will data collection occur?

- How will data be analyzed?
- Will a theoretical framework be used?

6 Exploratory Research Design

Exploratory research is conducted when prior knowledge is limited. It helps in hypothesis formulation by clarifying problems and discovering new insights.

Purposes

- Diagnose a situation
- Screen alternatives
- Discover new ideas

Characteristics

- No fixed method—flexibility and creativity are key.
- Informal and loosely structured.
- Useful for clarifying understanding and guiding further research.
- Offers low-risk initial inquiry.

`exploratory_research.png`

Figure 3: Exploratory Research Framework

7 Descriptive Research Design

Descriptive research describes phenomena systematically, aiming to provide accurate portrayals of conditions or behaviors.

Categories

- Historical
- Descriptive
- Developmental
- Survey
- Case Study

`descriptive_types.png`

Figure 4: Types of Descriptive Research

8 Historical Research

- Concerned with past events and phenomena.
- Involves evaluation and synthesis of evidence to establish facts.

- Aims to provide relevance of past to the present.
- Depends on primary (firsthand) and secondary (reported) sources.

Characteristics

- Involves critical analysis of source material.
- Systematic and exhaustive in approach.
- Similar to literature reviews but broader in scope.

9 Descriptive Research (Type b)

- Describes current situations, behaviors, and opinions.
- Often used for collecting factual data.
- Can be both quantitative and qualitative.
- Focuses on what exists, not why it exists.

Characteristics

- Literal description of events/situations.
- Builds a factual database without testing hypotheses.
- May serve as groundwork for future predictive or explanatory research.

Purposes

- Gather factual data about existing phenomena.
- Identify problems and justify practices.
- Make evaluations and comparisons.
- Learn from similar cases to inform future planning.

Lecture 11: Developmental, Survey, and Case Study Research

Developmental Research

Developmental research aims to predict future trends by studying variables, their rates of change, directions, sequences, and interrelated factors over time.

(1) **Longitudinal Study:**

- Studies phenomena over time, collecting data from the same group at two or more points.
- Mostly quantitative, not cross-sectional.
- **Trend Study:** Collects data at intervals to establish patterns and predict future conditions, often using regression analysis.
- **Cohort Study:** Follows a group sharing a common characteristic (e.g., birth year) at different times; rare due to difficulty maintaining contact.
- **Panel Study:** Studies the same individuals repeatedly to understand why changes occur.

(2) **Cross-sectional Study:**

- Observes items from a population at a single point in time.
- Measures rates of change by sampling different subjects across groups.
- Often employs surveys; faster and less expensive than longitudinal studies.

Characteristics:

- Focuses on patterns, rates, directions, and sequences of growth.
- Longitudinal studies face sampling and attrition challenges and require sustained resources.
- Cross-sectional studies involve more subjects but fewer growth factors; comparability across age groups can be an issue.

Survey Research

Survey research systematically gathers information from a population to understand or predict behaviors and attitudes.

Types:

- **Exploratory:** Early-stage, forms basis for deeper studies, sometimes uses previous data.

- **Confirmatory:** Tests specific theories or hypotheses.
- **Descriptive:** Describes distributions in a population, aiding theory development.

Surveys can be written, oral, or electronic, and distributed in various formats (face-to-face, mail, telephone, online). They allow data collection from large or small populations using standardized questionnaires.

Case Study Research

A case study is an in-depth, empirical investigation of a real-life phenomenon within its context, using multiple data sources.

Definition: "An empirical inquiry that investigates contemporary phenomenon within its real-life context" (Yin, 1994).

Types (Jensen and Rodgers, 2001):

- **Snapshot:** Detailed study of one unit at one time.
- **Longitudinal:** Studies the same unit at multiple time points.
- **Pre-post:** Two time points separated by a critical event.
- **Cross-cut:** Multiple cases for comparison.

Sampling:

- Often uses information-oriented sampling (e.g., extreme, critical, exemplar cases) rather than random sampling.

Characteristics:

- In-depth investigation of a social unit, resulting in a complete, well-organized picture.
- Examines a small number of units across many variables.
- Useful for hypothesis generation and exploring complex realities.
- Does not allow generalization to the population without further research.

Limitations:

- More expensive due to depth and exploratory nature.
- Generalizations from a single case are limited.
- Potential for subjectivity and bias.

Analytical Research Design

Analytical Research designs can be experimental or observational and each type has its own features. A study design is critical to the research study because it determines exactly how we will collect and analyze our data. If we aim to study the relationship between two variables, then an analytical study design is the right choice. It's necessary to have a clear plan before we begin data collection. Analytical study designs can be experimental or observational and each type has its own features.

A study design is a systematic plan, developed so we can carry out our research study effectively and efficiently. Having a design is important because it will determine the right methodologies for our study. Using the right study design makes our results more credible, valid, and coherent.

Descriptive vs. Analytical Research

Study designs can be broadly divided into either descriptive or analytical. Descriptive studies describe characteristics such as patterns or trends. They answer the questions of what, who, where, and when, and they generate hypotheses. They include case reports and qualitative studies.

Analytical study designs quantify a relationship between different variables. They're used to test hypotheses and make predictions.

Experimental and Observational

Analytical study designs can be either experimental or observational. In experimental studies, researchers manipulate something in a population of interest and examine its effects. These designs are used to establish a causal link between two variables.

In observational studies, in contrast, researchers observe the effects of a treatment or intervention without manipulating anything. Observational studies are most often used to study larger patterns over longer periods.

Experimental Research Method

The experimental method of research is used as the classical method in physical sciences. It is based on observation or experiments. It deals with actual experiments to determine the relationship between cause and effect of various experimental treatments. It is defined as "the research method in which a researcher objectively observes phenomenon which is made to occur in a strictly controlled situation where one or more variables are varied and others are kept constant".

The purpose of experimental research is to investigate cause and effect relationship by exposing one (or more) experimental groups to one (or more) treatment conditions and comparing the result to one (or more) control groups not receiving the treatments. In this method, the researcher undertakes control or manipulation (vary) of various variables under study. The usual approach is to hold all variables constant except one in controlled condition. By varying this one variable, the outputs (the effects) are studied and documented.

Actually, in social sciences, in natural sciences, in biological phenomena and the human behavior, control of variable is hardly possible. However, in physical sciences and experimental technology the investigation in controlled condition is highly acceptable.

Experiment is a test of a causal proposition, such as:

i) Do changes in variable 'A' cause changes in variable 'B' keeping other variables constant?

ii) How do the changes in the value of one variable (called independent variable) affect another variable (called dependent variable)?

The mathematical form of the experimental method is given below: If $x_1, x_2, x_3, x_4, \ldots, x_n$ are n independent variables taken as the inputs of the process and y is the output of the process (a dependent variable), then y is defined as a function of x and denoted by,

$$y = f(x)$$

where x means $x_1, x_2, x_3, x_4, \ldots, x_n$ and f denotes the function.

Suppose for example, yield (y) of a product in an agricultural field is influenced by the following four different independent variables:

- x_1 – seed quality (qualitative variable say, S_1, S_2)
- x_2 – amount of fertilizer (quantitative variable, in kg)
- x_3 – irrigation scheme (categorical variable say, I_1, I_2, I_3)
- x_4 – labor input (quantitative variable, in number)

The production or yield, which depends upon these four variables, can be related mathematically as

$$Y = f(x)$$

or

$$Y = f(x_1, x_2, x_3, x_4)$$

By taking any three (say x_1, x_2, x_3) constant one can observe the effect of x_4 in Y; x_4 may vary as researchers will, so it is said to be a controlled variable.

The various factors in an experiment are divided into two groups: independent variables and dependent variables. The first set of factors are called an experimental group and the second set of factors are called control group. Control group is also known as a group of individuals, items or objects used as a standard for comparison or accepted norm.

To make the experimental method of research effective and distinct from normal activity, the method of local control (blocking) and statistical control methods are used. Control is necessary to reduce variations. In some experiments, some variables may be eliminated.

Types of Experiment

Experiment is the scientific investigation in which an investigator manipulates and controls one or more independent variables and observes the dependent variables for variation concomitant to the manipulation of the independent variable. There are four different types of experiments:

1. **Positive and Negative Experiment:** If the subject of an experiment is such that (i) the phenomenon and (ii) its cause both are present, the experiment is said to be positive. For example, a bell rings in the air. Here both the sound and cause of its propagation are present; but if, on the other hand, a bell is rung in a vacuum, there will be no phenomenon of a sound because the cause of the propagation, namely air, is absent. Such experiment is called negative experiment.

2. **Natural Experiment:** These experiments are to be observed in natural phenomenon. In most of the natural state experiments, the controlling of variable is unnecessary to obtain the real information about the phenomena. In such case, the whole phenomenon is divided into control group and experimental group to study the effects of seen and unseen variables.

3. **Laboratory Experiment:** These are the experiments performed in physical sciences with full control of external conditions. A laboratory experiment is an artificially created situation in which the researcher controls one or more variables while manipulating other variable at will. The method of lab experiment is used in the experiments, mainly related to the physical, chemical, microbiological, clinical and such other sciences. If it is difficult to conduct an experiment outside or in the field or in the society then one tries to carry out it in the laboratories.

4. **Field Experiment:** Field experiments are the experiments conducted in the field or in natural setting. Research study in a realistic situation in which one or more independent variables are manipulated by the experimenter under carefully controlled conditions as the situation will permit. In the social, managerial, agricultural, environmental researches, the method of field experiment is widely used. Some of the field experiments like agricultural or business field the controlling of the variable is possible but in the careful condition.

Purposes of Experimental Method

- To determine the effect of various treatments and to compare the differences of effects as significant or non-significant.

- To estimate the interaction effects of various treatments and to compare them.

- To establish the mathematical relationship between various treatments and their effects.

Problems in Experimentation

- **To single out one factor from the phenomena:** It is always difficult to single out one factor from a social phenomenon for the purpose of measurement, because in any event there may be many factors interacted.

- **Controlling the factors:** Control of factors sometimes is not possible, because some factors may be unknown and uncontrollable. It is better to select several random samples as experimental and control groups. One solution here is the adoption of the control group technique.

- **To get data from the control groups:** There are difficulties in getting data from the control groups. The remedy may be found in matching the control and experimental groups on as many points as possible.

- **To assign the level of significance:** The determination of the required level of significance of the differences between the experimental and control groups is also fraught with difficulty. What difference can be taken as significant? There is the problem of value judgment. But the scientific criterion is the determination of the statistical test of significance. However, this requires a reliable and valid socio-metric scale.

- **Change in response of people:** In field experiment related with human behaviors (society and clinical setup), when data collected through human interaction due to changes in time, situation, environment and types of questions to be asked people often change their responses.

- **Change in theme of trialing:** Due to change in behavior of the respondents and unsatisfactory management of the investigator, theme of trialing of the area under experiment (in social and clinical setup) may change at the end of the experiment from what it was started. Because of the changes made by experimentation may give different responses which may lead to wrong conclusion.

- **Problem of handling or operation:** In social setup and to the medical trials, if the people under study area is not aware, attentive and responsive about the inquiry, true response cannot be possible.

Steps in Experimental Methods

1. **Statement of the problem, research questions and the objectives:** The first step in the application of field techniques is related to mentioning of problem, research questions and specific objective. The hypothesis, at this stage, should be stated explicitly in general terms.

2. **Examination of possible outcomes and events through literature:** The second step consists in setting up the field experiment by thorough reading of the available literature. The factors to be controlled must be assessed, the cooperation between the researcher and the subject must be set up.

3. **Design of experiment:** The next step is the choice of experimental design regarding its size, material, control groups etc. The choice of material should be based on the criterion of maximum possible accuracy. The basic problem of design relates to control. Control and experimental groups should be matched on all important factors.

4. **Performing experiment:** The next step of this method is to perform the experiment in predefined circumstances. The principles of randomization, replication and blocking should be implemented as much as possible. The sensitiveness of experiments can be augmented by neutralizing the biases through random choice, by increasing the replication, improving the quantitative technique and by refinements of techniques.

5. **Analysis of experimental outcomes statistically:** The analysis of the experimental data should be done starting from stating the descriptive nature of the data, measuring relationship between them and modeling data into some mathematical models. The analysis of variance permits a study of complex interrelationship, which is not possible by simpler designs. It permits more reliable conclusions about more hypotheses with fewer cases than if hypotheses were tested in separate design.

6. **Drawing conclusions by measuring reliability:** For an experimental research the conclusions are drawn based on the statistical significance testing. The tests can be performed as required level of design

by the use of different statistical techniques. The results obtained then are put to test their reliability and the conclusions are made.

7. **Testing the validity of the conclusion:** The validity of the results should be measured before disseminating the results and reports. The validity of the experimental results is checked by comparing with other similar phenomenon or to the standards.

8. **Evaluation of the entire investigation through practice:** The success of the experimental study can be measured only through putting into practice the experiments many times. If the repeated experiments give similar or better results, then the experimental results may be considered satisfactory.

Ethical Issues in Experimental Research Design

The following practices are considered unethical:

- Putting pressure on individuals to participate in experiments through coercion, or applying social pressure.
- Deceiving subjects by deliberately misleading them as to the true purpose of the research.
- Exposing participants to physical or mental stress.
- Not allowing subjects to withdraw from the research when they want to.
- Using the research results to disadvantage the participants, or for purposes not to their liking.
- Not explaining the procedures to be followed in the experiment.
- Not debriefing participants fully and accurately after the experiment is over.
- Not preserving the privacy and confidentiality of the information given by the participants.

- Withholding benefits from control groups.

Research Guides

Research Guides are librarian-curated pathways to information, videos, databases, and other resources for your discipline. That is, they pull many different types of resources on a subject or topic together in one place.

Hand Book

A handbook is a compilation of miscellaneous information in a compact and handy form. It contains data, procedures, principles, etc. Tables, graphs, diagrams, and illustrations are provided. Scientists and technologists use handbooks in their fields.

A treatise on a special subject; often nowadays a simple but all-embracing treatment, containing concise information, and being small enough to be held in the hand; but strictly, a book written primarily for practitioners and saving for constant revision or reference. Also called a 'Manual'.

Examples:

- Britain, 1948/49-, an official handbook, London, Stationery Office, 1948-, Annual.
- Handbook of Chemistry and Physics: A ready reference book of chemistry and physical data, 52nd ed, Cleveland, Ohio, Chemical Rubber, 1971.

Citation

A "citation" is the way you tell your readers that certain material in your work came from another source. It also gives your readers the information necessary to find the location details of that source on the reference or Works Cited page. A citation must include a set of parentheses.

APA

APA is the style of documentation of sources used by the American Psychological Association. This form of writing research papers is used mainly in the social sciences, like psychology, anthropology, sociology, as well as education and other fields.

IEEE

The Institute for Electrical and Electronics Engineers (IEEE) is a professional organization supporting many branches of engineering, computer science, and information technology. In addition to publishing journals, magazines, and conference proceedings, IEEE also makes many standards for a wide variety of industries.

IEEE citation style includes in-text citations, numbered in square brackets, which refer to the full citation listed in the reference list at the end of the paper. The reference list is organized numerically, not alphabetically.

Citation Index

Citation indexes allow researchers to trace the impact of an article upon later publications. Besides including the bibliographic information about an article (author, article title, journal title, date, etc.), citation indexes also provide each article's references or bibliography (the list of sources cited).

SCIFinder

SciFinder is a database focused on the literature in chemistry. It is produced and published by CAS: Chemical Abstracts Service, a division of the American Chemical Society. CAS has, as its objective, "to find, collect and organize all publicly disclosed chemical substance information."

SCOPUS

Scopus Indexed Journals are considered better sources for citation as compared to other databases. Scopus publications enjoy a good reputation among peer researchers due to their rigid selection procedure that ensures high-quality content and reliable data. In addition, the journal database is recognized by scholars in research and academia.

ScienceDirect

ScienceDirect is a website which provides subscription-based access to a large database of scientific and medical research. It contains the world's largest electronic collection of full-text and bibliographic information on science, technology and medicine.

Impact Factor

In any given year, the two-year journal impact factor is the ratio between the number of citations received in that year for publications in that journal that were published in the two preceding years and the total number of "citable items" published in that journal during the two preceding years:

$$\mathbf{IF}_y = \frac{\text{Citations}_y}{\text{Publications}_{y-1} + \text{Publications}_{y-2}}$$

For example, Nature had an impact factor of 41.577 in 2017:

$$\text{IF}_{2017} = \frac{\text{Citations}_{2017}}{\text{Publications}_{2016} + \text{Publications}_{2015}} = \frac{74090}{880 + 902} = 41.577$$

H-Index

The h-index is defined as the maximum value of h such that the given author/journal has published at least h papers that have each been cited at least h times.

Detailed Summary of Lecture 14

Data Analysis and Interpretation: Measurements and Scales

1. Introduction to Measurement

Measurement refers to the assignment of numerals to objects or events according to well-defined rules. The primary objective of measurement in research is to make empirical phenomena quantifiable for statistical analysis.

Key Concepts:

- **Numerals** are symbols used to distinguish objects. They lack inherent quantitative meaning.
- When quantitative meaning is assigned to numerals, they become **numbers**.
- Measurement involves a process of **mapping** objects from one set onto another set using a function or rule.
- The rule of assigning numerals to objects is defined as a **scale**.

Example: Measuring the gender of family members by assigning 1 for male and 0 for female. If a family $A = \{a_1, a_2, a_3, a_4, a_5\}$ contains three males, the mapping $f(a_i) \to \{1, 0\}$ allows symbolic representation of gender.

2. Steps in Measurement Procedure

1. Define the objects in the universe of discourse (e.g., all 10th grade pupils in a school).
2. Identify the measurable properties of these objects (e.g., sex, income, education).
3. Partition the universe into mutually exclusive and exhaustive subsets based on the property.
4. Count the number of members in each subset, assuming internal homogeneity within a subset.

3. Types of Measurement Scales

There are four fundamental scales of measurement:

3.1. 1. Nominal Scale (Categorical)

Definition: Classifies data into distinct categories without any order or quantitative value.

Examples:

- **Gender:** Male, Female
- **Religion:** Hindu, Buddhist, Muslim, Christian
- **Occupation:** Teacher, Manager, Doctor
- **Department:** Sales, Finance, Production

Properties:

- Mutually exclusive and collectively exhaustive categories.
- No implied order or ranking.
- Statistical operations: Frequencies, percentages.
- Test: **Chi-square test** is suitable for nominal data.
- Associated with exploratory research and qualitative variables.

Example Use Case: If 150 students are coded as 1 for male and 2 for female, frequency analysis may reveal 100 males (66.6%) and 50 females (33.4%).

3.2. 2. Ordinal Scale (Ranking)

Definition: Represents ranked order of variables without assuming equal intervals between them.

Examples:

- Ranking jobs by social status: Doctor, Engineer, Professor, etc.
- Ranking cities by suitability for opening a bank branch.

Properties:

- Captures order or preference.

- Interval between ranks is unknown or non-uniform.
- Measures of central tendency: **Median**, **percentiles**.
- Measures of dispersion: **Quartiles**, **range**.
- Suitable statistical tools: **Non-parametric tests**, rank-order correlation.

Application: Frequently used in surveys, customer satisfaction ratings, and preference studies.

3.3. 3. Interval Scale (Equal Intervals without True Zero)

Definition: A numeric scale where both order and exact differences between values are meaningful, but the zero point is arbitrary.

Example:

- Age differences among siblings: Radha is 4 years older than Rabina, Rambha is 3 years older, etc.
- Attitude measurement using Likert scales (e.g., 1 to 5 scale).

Properties:

- Equal spacing between scale points.
- No absolute zero; zero does not indicate absence of quantity.
- Appropriate statistics: **Mean**, **Standard Deviation**, **Pearson correlation**, t-test, F-test.

Application: Behavioral and psychological research (e.g., intelligence tests, attitude scales).

3.4. 4. Ratio Scale (True Zero)

Definition: The highest level of measurement that includes all properties of the interval scale and has an absolute zero.

Examples:

- Income (e.g., Rs. 50,000/year)
- Age (e.g., 25 years)
- Number of children, years of schooling, hours worked

Properties:

- Allows comparison of magnitudes (e.g., "twice as much").
- Absolute zero point implies absence of the variable.
- All mathematical operations are possible.
- Measures: **Geometric mean**, **harmonic mean**, **coefficient of variation**.

Example Use Cases:

- *How many children do you have?*
- *What is your annual income?*
- *How many employees work in your factory?*

Note: Ratio scales are more common in physical sciences than in social sciences.

4. Comparative Summary of Scales

Scale	**Order**	**Equal Interval**	**True Zero**	**Examples**
Nominal	No	No	No	Gender, Nationality
Ordinal	Yes	No	No	Rank, Preferences
Interval	Yes	Yes	No	Temperature (C), IQ
Ratio	Yes	Yes	Yes	Age, Income, Height

5. Conclusion

Measurement scales are foundational in determining how data can be analyzed statistically. The choice of scale impacts the types of statistical methods that are appropriate. Understanding the properties and limitations of each scale is crucial for effective research design and data interpretation.

Detailed Summary of Lecture 15

Measurement of Scale

1. Needs of Scaling

Scaling is essential in research as it converts qualitative facts into quantitative data, enabling scientific analysis. The main needs of scaling are:

- To achieve scientific measurement of qualitative facts.
- To ensure objective and reliable measurement in technical and social studies.
- To improve and develop more precise measuring instruments through scientific use of existing scales.

2. Characteristics of a Good Scale

A good measurement scale should possess the following characteristics:

- **Continuum:** Represents a continuous series of points, with interrelated factors, defined by the nature of the phenomenon.
- **Reliability:** Produces consistent, stable, and accurate results. A reliable scale gives the same result for the same object under the same conditions, indicating absence of measurement error.
- **Validity:** Measures what it is intended to measure. A valid scale reflects true differences among subjects. Validity implies reliability, but not vice versa.
- **Practicability:** The scale should be economical, convenient, easy to administer, and interpretable. It should come with clear instructions, scoring keys, utilization guides, and evidence of reliability.

3. Methods of Estimating Reliability

Reliability is the consistency of a scale. The main methods of estimating reliability are:

1. **Test-Retest Method:**
   - The same test is administered multiple times to the same subjects.
   - The correlation between the two sets of scores is calculated.
   - High correlation indicates high reliability.
2. **Parallel (Alternate) Forms Method:**
   - Two equivalent forms of a test are administered to the same group.
   - The correlation between the two sets of scores measures reliability.
3. **Split-Half Method:**
   - The test is split into two halves (e.g., odd vs. even items).
   - The correlation between the halves is calculated.
   - The Spearman-Brown formula is used to adjust the reliability estimate for the full test:
     $$R_{xy} = \frac{2r_{xy}}{1 + r_{xy}}$$
     where r_{xy} is the correlation between the two halves.
4. **Rational Equivalence (Kuder-Richardson) Method:**
   - Measures internal consistency by analyzing the correlation among items within the same test.
   - The Kuder-Richardson formula (KR-1) is:
     $$KR_1 = R_w = \frac{n}{n-1}\left[1 - \frac{\Sigma pq}{\sigma^2}\right]$$
     where n is the number of items, p is the proportion of correct answers, $q = 1-p$, and σ^2 is the variance of test scores.

4. Examples

- **Split-Half Example:** If a test is split into odd and even items and the correlation between halves is 0.72, the reliability coefficient for the whole test is:

$$R_{xy} = \frac{2 \times 0.72}{1 + 0.72} = 0.8272$$

indicating 82.72% reliability.

- **Kuder-Richardson Example:** For a test with 60 questions, $p = 0.7$, $q = 0.3$, and standard deviation 10:

$$KR_1 = \frac{60}{59}\left[1 - \frac{60 \times 0.7 \times 0.3}{100}\right] = 0.8888$$

or 88.88% reliability.

5. Summary Table: Reliability Estimation Methods

Method	**Description**	**Formula/Approach**
Test-Retest	Same test, same subjects, different times	Correlation coefficient between sco
Parallel Forms	Two equivalent tests, same subjects	Correlation between forms
Split-Half	Divide test into two halves, correlate scores	Spearman-Brown formula
Kuder-Richardson	Internal consistency among test items	$KR_1 = \frac{n}{n-1}\left[1 - \frac{\Sigma pq}{\sigma^2}\right]$

6. Conclusion

Scaling and reliability are essential for scientific measurement in research. Good scales must be continuous, reliable, valid, and practical. Reliability can be estimated using several statistical methods, ensuring that measurement instruments yield consistent and meaningful results.

Summary of Lecture 16: Validity and Scaling in Social Sciences

1 Validity

Definition: A test possesses *validity* when it measures what it is intended to measure. Validity indicates the accuracy of predictions made from test scores.

1.1 Types of Validity

1.1.1 1. Content Validity

- Also known as *logical validity*.
- Ensures test items represent the entire domain of content.
- Often used in academic, vocational, and clinical contexts.
- Determined by expert judgment and statistical analysis.

1.1.2 2. Criterion-Related Validity

- Measures how well a test predicts outcomes on an external, validated criterion.
- Based on:
  - External criterion
  - Future or concurrent behavior
  - Logical and empirical methods
- Types:

  Predictive Validity: Forecasts future performance.

  Concurrent Validity: Correlates with current valid measures.
- Validity is measured using correlation coefficient r:

Range of r	Interpretation
$0.9 \leq r \leq 1.0$	Very high validity
$0.8 \leq r < 0.9$	High validity
$0.6 \leq r < 0.8$	Satisfactory validity
$0.4 \leq r < 0.6$	Moderate validity
$0.0 \leq r < 0.4$	Poor validity
$r < 0$	Negative validity

1.1.3 3. Construct Validity

- Most abstract form of validity.
- Involves theoretical constructs or psychological traits.
- Evaluated using statistical correlations and factor analysis.

2 Scaling

2.1 Scores

- **Raw Score:** Basic count of correct answers.
- Limited in comparative value.

2.2 Scales and Derived Scores

- Scales map raw scores into structured levels.
- Derived scores include:
  - Percentile Score
  - Z-Score (z-score)
  - T-Score
- These scores are arranged to form:
  - Percentile Scale
  - Sigma Scale
  - T-Scale

3 Difficulties in Scaling Social Phenomena

- Abstract nature of social values
- Heterogeneous customs and norms

- Variability of human behavior
- Absence of universal standards
- Inapplicability of laboratory methods

4 Scales in Social and Physical Sciences

1. Point Scale

- Assigns one point per criterion.
- Methods:
  - Tick favorable options.
  - Cross unfavorable options.
  - Indicate agreement/disagreement.

2. Social Distance Scale (Bogardus Scale)

- Measures willingness to associate with members of different social or ethnic groups.

3. Rating Scale

- Measures non-binary traits or attitudes.
- Examples:
  - 3-point scale: Very Good – Satisfactory – Poor
  - 5-point Likert scale: Strongly Agree – Agree – Neutral – Disagree – Strongly Disagree

4. Ranking Scale

- Respondents rank items based on preference.
- Lower score indicates higher preference.

5. Thurstone Scale

- Developed by Louis L. Thurstone.
- Measures attitudes on a continuum from favorable to neutral to unfavorable.
- Used in educational and psychological assessments.

Detailed Summary of Lecture 17: Sampling

1 Introduction to Sampling

- Sampling is the process of selecting a subset of individuals, items, or data points from a larger population for the purpose of estimating characteristics of the whole population.
- Studying an entire population is often not feasible due to constraints in time, cost, and effort. Sampling provides a practical alternative.
- A **sample** is a smaller, manageable representation of the larger population.
- A **population** refers to the complete set of individuals or items that possess some common characteristic under investigation.
- Types of populations:
  - **Finite Population**: Has a countable number of elements (e.g., students in a college).
  - **Infinite Population**: The total number of elements cannot be counted (e.g., coin toss outcomes).
- A well-selected sample ensures that generalizations made about the population are reliable and valid.

2 Sample Design and Terminologies

- **Sampling Frame**: A complete list or map that includes all members of the population from which the sample will be drawn (e.g., voter list, class roster).
- **Sampling Unit**: The basic element or group of elements considered for selection in a sampling process (e.g., person, household, farm).
- **Sample Size**: The number of units to be included in the sample. A larger sample size improves accuracy but increases cost and complexity.
- **Sample Statistics**: Values computed from the sample data, such as sample mean, variance, and proportion. These are used to estimate population parameters.
- **Population Parameters**: True values (e.g., population mean, variance) that describe the entire population. Often unknown and estimated using sample statistics.

3 Principal Steps in a Sample Survey

1. **Define Objectives**:
   - Clearly define the goals and purposes of the survey.
   - Objectives should be practical, specific, and aligned with available resources.
2. **Define the Population**:
   - Clearly specify what constitutes inclusion in the population.
   - Ambiguity in defining the population can lead to biased results.
3. **Frame and Sampling Units**:
   - Ensure that all elements in the population are covered.
   - Units must be distinct and non-overlapping.
   - A good sampling frame is essential for effective sample selection.
4. **Data to Collect**:
   - Decide in advance what data are necessary.
   - Avoid collecting irrelevant or excessive information.
   - Draft an outline of the tables you want to produce post-survey.
5. **Design Questionnaire or Schedule**:
   - Prepare clear, concise, and unbiased questions.
   - The questionnaire may be self-administered or interviewer-administered.
   - Include clear instructions and ensure logical flow of questions.
6. **Data Collection Methods**:
   - *Interview Method*: Investigator meets respondents directly and fills out the questionnaire.
   - *Mailed Questionnaire*: Sent to respondents to be filled out and returned.
   - Choose the method based on literacy level, cost, accuracy, and coverage.
7. **Non-response Handling**:
   - Some selected respondents may not respond due to absence or refusal.
   - Investigators must record reasons and take steps to minimize non-response bias.
8. **Select Sampling Design**:
   - Choose an appropriate sampling method (e.g., simple random, stratified, cluster).
   - Estimate an appropriate sample size for desired precision and confidence level.

- Consider cost and logistics before finalizing the design.

9. **Organize Field Work**:
   - Train enumerators to locate and interact with sampling units.
   - Pretesting helps to identify potential issues in questionnaire and methodology.
   - Proper supervision ensures data quality.
10. **Data Analysis**:
    - **Editing and Scrutiny**: Check for inconsistencies or incomplete responses.
    - **Tabulation**: Summarize data into tables for easier analysis.
    - **Statistical Analysis**: Apply estimation formulas and compute confidence intervals or significance levels.
    - **Reporting**: Document methodology, findings, and limitations clearly.
11. **Learning for Future Surveys**:
    - Record lessons learned regarding design, execution, and analysis.
    - Use these insights to improve accuracy and efficiency in future surveys.

Lecture 18: Parameter, Statistic, Sampling and Non-Sampling Errors

1 Parameter and Statistic

- **Parameter**: Statistical constants of the population, such as mean, variance, etc., are called parameters.
- **Statistic**: Measures computed from sample observations alone (e.g., sample mean, sample variance) are called statistics.
- In practice, parameter values are usually unknown and are estimated based on sample values. Thus, a statistic is an estimate of the parameter, derived solely from sample values.
- Since there are multiple possible samples from a population, a statistic varies from sample to sample.
- Characterizing the variation in the values of a statistic obtained from different samples (attributed to chance or sampling fluctuations) is a fundamental problem in sampling theory.

2 Sampling Distribution

- The number of possible samples of size n that can be drawn from a finite population of size N is ${}^{N}C_{n}$.
- For each sample, a statistic (e.g., mean, variance) can be computed, and these values will vary from sample to sample.
- The aggregate of all such values (one from each sample) forms a frequency distribution known as the **sampling distribution** of the statistic.
- Thus, we can have the sampling distribution of the sample mean $\bar{x}$, the sample variance, etc.

3 Standard Error

- The standard deviation of the sampling distribution of a statistic is called its **Standard Error (S.E.)**.
- The standard errors of some well-known statistics are commonly tabulated, where n is the sample size, σ^2 is the population variance, P is the population proportion, and $Q = 1 - P$.

4 Sampling and Non-Sampling Errors

The errors in data collection, processing, and analysis can be broadly classified as:

1. **Sampling Errors**
2. **Non-sampling Errors**

4.1 Sampling Errors

- Originate from the use of only a part of the population (sample) to estimate population parameters.
- Absent in a complete enumeration (census) survey.
- Sampling biases can arise due to:
  (1) **Faulty selection of the sample:** Use of defective sampling techniques (e.g., purposive or judgment sampling) may introduce bias. This can be minimized by using simple random sampling or random sampling with restrictions that do not introduce bias.
  (2) **Substitution:** Replacing a selected unit with a convenient member if difficulties arise in enumeration introduces bias, as the substitute may have different characteristics.
  (3) **Faulty demarcation of sampling units:** Especially significant in area surveys (e.g., agricultural experiments), where discretion in including borderline cases can introduce bias.
  (4) **Constant error due to improper choice of statistic:** For example, using the sample variance formula
  $$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$
  as an estimate of the population variance σ^2 is biased, whereas
  $$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$
  is an unbiased estimate.
- **Sample Size:** Increasing the sample size generally decreases sampling error.

4.2 Non-sampling Errors

- Arise at the stages of observation, ascertainment, and processing of data.
- Present in both sample surveys and complete enumeration surveys.
- Major sources include:

  (1) **Planning Errors:**
    - Inadequate or inconsistent data specification.
    - Errors in locating units, measurement, recording, or due to poorly designed questionnaires.
    - Lack of trained investigators or supervisory staff.

  (2) **Response Errors:**
    - *Accidental errors:* Misunderstanding questions.
    - *Prestige bias:* Upgrading or downgrading responses due to pride.
    - *Self-interest:* Providing incorrect information to protect one's interests.
    - *Interviewer bias:* Influence of interviewer's beliefs or recording style.
    - *Recall failure:* Inability to accurately remember past events.

  (3) **Non-response Biases:** Occur when full information is not obtained for all units (e.g., respondent not at home, refusal to answer).

  (4) **Coverage Errors:**
    - Inclusion of units that should not be included or exclusion of units that should be included due to unclear objectives or definitions.

  (5) **Compiling Errors:** Errors during data processing (editing, coding, tabulation, summarizing). Can be controlled by verification and consistency checks.

  (6) **Publication Errors:**
    - Mechanical errors in publication (proofing, printing).
    - Failure to point out limitations of the statistics.

Lecture 19: Sampling Frame and Sampling Techniques

Sampling Frame

The **sampling frame** (also known as the "sample frame" or "survey frame") is the actual collection of units from which a sample is drawn. A basic random sample gives all units in the frame an equal probability of being selected.

A sampling frame is a complete list or collection from which sample participants will be drawn in a predetermined manner. This list is organized so that each member of the population has an individual identity and a contact mechanism, allowing for categorization and coding of known segmentation features.

Having a sampling frame means we have a supply or list of all individuals in the target population, as well as a process for selecting the sample. Any resource that enables access to every individual in the targeted group qualifies as a sampling frame.

Characteristics of a Good Sampling Frame

When selecting lists, ensure the sample frame is large enough for the requirements. A good sampling frame should:

- Include everyone in the target demographic.
- Exclude anyone not in the target group.
- Contain factual information to reach specific people.
- Assign a unique identification to each member (e.g., a number code from 1 to 3000).
- Be free of duplicates.
- Be well organized (e.g., sorted alphabetically).
- Be up to date, with regular checks for address or contact changes.

Examples of the Sampling Frame

Studying every individual in a population is often impractical. For example, to learn about the opinions of Nepalese bankers about vehicle ownership, surveying every bank would be too time-consuming and expensive. In such cases, a sample is investigated.

Before choosing a sample, construct a sampling frame-a list of all units in the population of interest. Study findings can only be generalized to the population identified by the sampling frame.

Conclusion

A sampling frame is a researcher's list or device to specify the population of interest. A basic random sample gives all units an equal probability of selection. Units can be people, organizations, or records. It is critical to be as detailed as possible when describing the population.

Issues of Choosing Appropriate Sampling Technique(s)

Choosing a sampling strategy is essential to ensure data reliability and representativeness. Consider a survey on characteristics (tax, education, etc.) of residents in five towns, totaling 3,200 households. These households form the target population.

Step One: Define Sample and Target Population

Sometimes a survey covers the entire target population (a census survey), but this is often impractical. Instead, a smaller, representative sample is chosen to reflect the population's characteristics. A survey on a smaller number is called a *sample survey*; findings from this can be generalized to the whole population.

Step Two: Define Sample Size

There are no strict rules for sample size; it depends on objectives, time, budget, and desired precision. To select an appropriate sample size, determine the degree of accuracy (confidence interval and confidence level):

- **Confidence interval** (margin of error): The range within which the true value is expected to fall (e.g., ± 5).
- **Confidence level**: The probability that the sample reflects the population (e.g., 95%).

Example: If 65% of sampled households say "yes" to a question, with a ± 5 confidence interval and 95% confidence level, we can say that between 60% and 70% of all 3,200 households would also answer "yes."

Sample size can be determined using a standard calculator. For 3,200 households:

- **Option A:** 5% confidence interval, 95% confidence level $\rightarrow$ sample size = 345 households.

- **Option B:** 5% confidence interval, 99% confidence level → sample size = 551 households.

The improvement in accuracy diminishes as sample size increases, so the choice should balance objectives and resources.

Step Three: Define Sampling Technique

After choosing the sample size, select a sampling technique based on the project's nature and objectives. Sampling techniques are broadly divided into:

- **Random Sampling**
- **Non-random Sampling**

Random Sampling

Random sampling means selecting the sample randomly from a population, often from a list or at the survey location.

- **Simple random sampling without replacement:** Each unit can be selected only once.
- **Simple random sampling with replacement:** Units may be selected multiple times.
- **Systematic sampling:** Divide the population by the sample size to get the sampling interval. E.g., for 3,200 households and a sample of 345, the interval is 9 ($3200/345 \approx 9$); select every ninth household.
- **Stratified random sampling:** For heterogeneous populations, divide into strata and sample each in proportion to its size.

Example of stratified random sampling:

Location	Population size	Proportion (%)	Stratified sample size
Town 1	1200	38%	129
Town 2	900	28%	97
Town 3	800	25%	86
Town 4	180	6%	19
Town 5	120	4%	13
Total	3200	100%	345

Non-random Sampling

Non-random sampling selects samples based on specific characteristics. Used when the sample must meet certain criteria, such as owning a car or having young children. Methods include:

- Convenience sampling
- Judgment sampling
- Quota sampling

- Snowball sampling

Step Four: Minimize Sampling Error

Efforts should focus on reducing sampling error and making the sample as representative as possible. The robustness of the sample depends on minimizing error, which varies by technique.

For random samples, results are reported with the $\pm$ sampling error. In non-random samples, the sampling error is unknown.

Summary: Use random sampling when inferring proportions of the population. Use non-random sampling when targeting specific perceptions or characteristics, especially when speed or specificity is needed.

Note: Without a sampling strategy, data may be biased or unrepresentative, rendering it invalid.

Lecture 20: Sample Size Determination

Sample Size

The size of the sample is an important factor as it directly affects the accuracy, estimation, cost, and administration of a survey. Large samples have lower sampling error, whereas small samples have higher sampling error. However, unnecessarily large samples increase costs, so an optimum sample size should be chosen to ensure efficiency, representativeness, reliability, and flexibility.

Factors Affecting Sample Size

Sample size depends on several factors:

1. **Nature of population:** If the population is homogeneous, a small sample size suffices; if heterogeneous, a larger sample is needed for representativeness.
2. **Number of classes:** More classes in the classification require a larger sample.
3. **Nature of study:** Studies that take longer may benefit from smaller samples due to financial and analytical constraints.
4. **Types of sampling used:** Simple random sampling typically needs a larger sample, while stratified sampling can achieve representativeness with a smaller sample.
5. **Degree of accuracy:** Higher accuracy requires a larger sample.

Testing Reliability of the Sample

A sample is considered reliable if it is representative of the population. Reliability can be tested by:

1. **Drawing parallel samples:** Draw another sample parallel to the original and compare measures such as average, dispersion, skewness, and kurtosis. If measures are similar, the sample is reliable.
2. **Comparing sample with population:** Compare sample statistics with population parameters. If they match closely, the sample is reliable.

3. **Drawing sub-samples from the main sample:** Compare measures from sub-samples with the main sample to detect errors due to faulty selection.

Methods of Estimating Sample Size

Estimation Using the Mean

Let $\bar{x}$ be the sample mean from a random sample of size n drawn from a population with mean μ and standard deviation σ. For a confidence level $(1-\alpha)$ and margin of error d:

$$P\left(|\bar{x}-\mu| \leq Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

Setting $d = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$, we get:

$$n = \frac{Z_{\alpha/2}^2\sigma^2}{d^2}$$

If σ is unknown, use the sample standard deviation s.

For a finite population of size N:

$$n = \frac{Z_{\alpha/2}^2\sigma^2}{d^2 + \frac{Z_{\alpha/2}^2\sigma^2}{N}}$$

Estimation Using Proportion

Let p be the sample proportion from a random sample of size n drawn from a population with proportion P and $Q = 1-P$:

$$P\left(|p-P| \leq Z_{\alpha/2}\sqrt{\frac{PQ}{n}}\right) = 1-\alpha$$

Setting $d = Z_{\alpha/2}\sqrt{\frac{PQ}{n}}$, we get:

$$n = \frac{Z_{\alpha/2}^2 PQ}{d^2}$$

If P is unknown, use $P = p$.

For a finite population of size N:

$$n = \frac{Z_{\alpha/2}^2 PQ}{d^2 + \frac{Z_{\alpha/2}^2 PQ}{N}}$$

Examples

Example 1: Determine the minimum sample size required so that the sample estimate lies within 10% of the true value with 95% confidence, given a coefficient of variation (CV) of 60%.

Solution: CV $= 0.6$, margin of error $d = 0.1\mu$, $Z_{0.025} = 1.96$.

$$0.1\mu = 1.96 \times \frac{\sigma}{\sqrt{n}}$$

$$n = \left(\frac{1.96 \times \sigma}{0.1\mu}\right)^2 = 384.16 \times (0.6)^2 = 138.29 \approx 138$$

Example 2: A psychologist wants to be 99% confident that the error in estimating reaction time (standard deviation 0.05 seconds) does not exceed 0.01 seconds.

Solution: $Z_{0.005} = 2.58$, $d = 0.01$, $\sigma = 0.05$.

$$n = \frac{(2.58)^2 \times (0.05)^2}{(0.01)^2} = 166.4 \approx 167$$

Example 3: A survey in Kathmandu Valley wants to estimate the proportion of disabled persons (population proportion $p = 0.1$, desired error $d = 0.02$, 95% confidence).

Solution: $Z_{0.025} = 1.96$, $Q = 0.9$.

$$n = \frac{(1.96)^2 \times 0.1 \times 0.9}{(0.02)^2} = 864.36 \approx 865$$

Example 4: For $p = 0.2$, $d = 0.05$, $Z = 2$, find n. If $N = 1000$, find corrected n.

Solution:

$$n = \frac{4 \times 0.2 \times 0.8}{(0.05)^2} = 256$$

For $N = 1000$:

$$n_{corr} = \frac{256}{1 + \frac{256}{1000}} = 204$$

Example 5: Mean systolic blood pressure is 125 mm Hg, standard deviation 15 mm Hg. Find sample size for 5% significance and error not exceeding 2.

Solution: $Z_{0.025} = 1.96$, $d = 2$, $\sigma = 15$.

$$n = \frac{(1.96)^2 \times 15^2}{2^2} = 216.09 \approx 216$$

For $N = 500$:

$$n_{corr} = \frac{216}{1 + \frac{216}{500}} = 151$$

Yamane Formula

For a finite population, Yamane's formula is:

$$n = \frac{N}{1 + Ne^2}$$

where n = sample size, N = population size, e = acceptable sampling error.

Example: For $N = 10,000$, $e = 0.05$:

$$n = \frac{10,000}{1 + 10,000 \times (0.05)^2} = \frac{10,000}{26} = 385$$

Standard Error of the Mean

- For an infinite population:
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

- For a finite population of size N:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

Note

When the original sample is more than 5% of the population, use the finite population correction. The sample size should be manageable, cost-effective, and representative.

Lecture 21: Statistical Analysis and Data Preparation

Statistical Analysis

Analysis refers to the categorizing, ordering, manipulating, and summarizing of data to answer research questions. The purpose is to reduce data to an intelligible and interpretable form so that relationships between research problems can be studied and tested. Statistics are used to manipulate and summarize numerical data, and to compare results with chance expectations. The researcher should plan analysis paradigms or models when formulating problems and hypotheses.

Data Editing

Editing is the process of examining collected data for errors and omissions, and making necessary corrections. Editing occurs in two stages:

1. **Field editing:** Review of reporting forms by the enumerator or investigator for completing what was written in abbreviated form during data collection.
2. **Central editing:** Editing of obvious errors (e.g., entries in the wrong place, missing replies) by an editor after all forms have been returned to the office.

Data Coding

Coding is the process of assigning numerals or symbols to responses so they can be grouped into a limited number of classes or categories. Quantitative data collected via questionnaires or schedules is already numeric and may not require coding. For qualitative data, numeric codes are assigned before analysis. This allows qualitative responses to be converted into numerical figures suitable for statistical treatment.

Classification of Data

Classification arranges related facts or data into groups or classes according to similarities, making data easily understandable. Classification should be:

- According to the research problem

- Exhaustive
- Mutually exclusive
- Independent

Objectives of classification:

- To condense large amounts of data
- To facilitate comparison
- To highlight features of the data at a glance
- To enable statistical analysis

Types of classification:

1. **Chronological classification:** Data is arranged by time (years, months, weeks, etc.), usually in ascending order.
   Example:

Year	1970	1971	1972	1973	1974	1975	1976
Birth rate	36.8	36.9	36.6	34.6	34.5	35.2	34.2

2. **Geographical classification:** Data is classified by region or place.
   Example:

Country	America	China	Denmark	France	Nepal
Yield of wheat (kg/acre)	1925	893	225	439	862

3. **Qualitative classification:** Data is classified by attributes or qualities such as sex, literacy, religion, employment, etc. Attributes cannot be measured on a scale. When two or more attributes are considered, a manifold classification is formed.
   Example: Classifying population by sex and employment:
   - Male employed
   - Male unemployed
   - Female employed
   - Female unemployed

   Further attributes (e.g., marital status) can extend the classification.
4. **Quantitative classification:** Data is classified by measurable characteristics such as height, weight, etc.

Data Entry into Spreadsheet

A spreadsheet is an interactive computer application for organizing and analyzing data in tabular form. Spreadsheets simulate paper accounting worksheets, with data in cells organized as rows and columns. Each cell can contain numeric or text data, or formulas that automatically calculate values. Users can change values and instantly see the effects on calculations, making spreadsheets useful for "what-if" analysis. Modern spreadsheets can have multiple sheets and display data as text, numbers, or graphs.

Management of Missing and Inconsistent Information

Researchers often face:

- Missing data
- Impossible values
- Inconsistencies
- Transcription errors

Missing and inconsistent data are common in research. Strategies for handling them include:

- Careful questionnaire design and data entry
- Training data entry personnel
- Cross-checking responses
- Checking for impossible values
- Recording and creating composite variables
- Documenting any logical changes to raw data
- Using coding systems to reduce errors
- Data cleaning with standard software (e.g., SAS, SPSS)
- Labeling values and merging cells as needed
- Formatting data for analysis
- Maintaining a master dataset for all analysts

Listwise Deletion

Cases with any missing values are deleted from analysis. Only cases with complete data are retained. This is the default in many statistical programs but is generally not recommended.

Pairwise Deletion

The maximum amount of available data is retained. Cases are excluded only from analyses where required data is missing. For example, a case missing data on one variable is excluded from analyses involving that variable, but included elsewhere.

Descriptive Statistical Measures

Types of average:

- Arithmetic Mean
- Geometric Mean
- Harmonic Mean
- Median
- Mode

Measures of Dispersion:

- Absolute and relative measures
- Range
- Quartile deviation
- Mean deviation
- Standard deviation
- Coefficient of Variation

Skewness and Kurtosis:

$$\text{Kurtosis } K = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})} \qquad K = 0.263$$

Correlation and Regression

Lecture 22: Inferential Statistics and Hypothesis Testing

Inferential Statistics

Inferential statistics involve drawing conclusions about a population based on sample data. A key component is hypothesis testing, which allows us to make decisions or inferences about population parameters.

Testing of Hypothesis

Hypothesis testing is a statistical method that uses sample data to evaluate a hypothesis about a population parameter. The general steps are:

1. Formulate the null hypothesis (H_0) and alternative hypothesis (H_1).
2. Select an appropriate test statistic.
3. Determine the level of significance (α).
4. Find the critical value(s) from statistical tables.
5. Make a decision: reject or fail to reject H_0.

Z Test

The Z test is an important parametric test based on the assumption of normality. It is used when the sample size is large ($n > 30$), or when the population variance is known. For large samples, the sampling distribution of the statistic is approximately normal, even if the population is not strictly normal.

The Z test statistic is defined as:

$$Z = \frac{t - E(t)}{SE(t)} \sim N(0, 1)$$

where t is a statistic, $E(t)$ its expected value, and $SE(t)$ its standard error.

Z tests are used for:

- Significance of a single mean
- Significance of difference between two means
- Significance of a single proportion
- Significance of difference between two proportions
- Significance of difference between sample and population correlations
- Significance of difference between independent sample correlations

Test of Significance of a Single Mean

Suppose a sample of size $n > 30$ is drawn from a normal population $N(\mu, \sigma^2)$. The sample mean $\bar{x}$ is approximately normally distributed.

Steps:

1. **State the hypotheses:**
    - $H_0 : \mu = \mu_0$ (sample is from population with mean μ_0)
    - $H_1 : \mu \neq \mu_0$ (two-tailed), or $H_1 : \mu > \mu_0$ (right-tailed), or $H_1 : \mu < \mu_0$ (left-tailed)
2. **Test statistic:**
$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \quad \text{(for known variance)}$$
$$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}} \quad \text{(for unknown variance, large } n\text{)}$$
   For finite population size N:
$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}}$$
3. **Level of significance:** Usually $\alpha = 0.05$ unless specified.
4. **Critical value:** Obtain from standard normal tables according to α and the alternative hypothesis.
5. **Decision:** Reject H_0 if $|Z| > Z_{\text{tabulated}}$; otherwise, accept H_0.

Example:

A sample of 400 students has a mean height of 170 cm. Can it be regarded as a sample from a large population with mean 169.5 cm and standard deviation 3.5 cm?

Solution:
$$Z = \frac{170 - 169.5}{3.5/\sqrt{400}} = \frac{0.5}{0.175} = 2.857$$

At $\alpha = 0.05$, $Z_{\text{tab}} = 1.96$. Since $2.857 > 1.96$, reject H_0. The sample cannot be regarded as from the population with mean 169.5 cm.

Test of Significance of Difference Between Two Means

Suppose two independent samples of sizes n_1 and n_2 are drawn from populations with means μ_1, μ_2 and variances σ_1^2, σ_2^2. Let $\bar{x}_1$, $\bar{x}_2$ be the sample means.

Steps:

1. **State the hypotheses:**
   - $H_0 : \mu_1 = \mu_2$
   - $H_1 : \mu_1 \neq \mu_2$ (two-tailed), or one-tailed as appropriate
2. **Test statistic:**
$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$
   If population variances are unknown (large samples), use sample variances s_1^2, s_2^2.
3. **Level of significance:** Usually $\alpha = 0.05$.
4. **Critical value:** Obtain from standard normal tables.
5. **Decision:** Reject H_0 if $|Z| > Z_{\text{tabulated}}$.

Example:

Sample 1: $n_1 = 500$, $\bar{x}_1 = 20$, $\sigma_1 = 4$
Sample 2: $n_2 = 400$, $\bar{x}_2 = 15$, $\sigma_2 = 4$

$$Z = \frac{20 - 15}{\sqrt{\frac{16}{500} + \frac{16}{400}}} = \frac{5}{0.27} = 18.51$$

At $\alpha = 0.05$, $Z_{\text{tab}} = 1.96$. Since $18.51 > 1.96$, reject H_0. The samples are not from the same population.

Test of Significance of Difference Between Two Proportions

Let P_1 and P_2 be the population proportions, p_1 and p_2 the sample proportions from independent samples of sizes n_1 and n_2.

Steps:

1. **State the hypotheses:**
   - $H_0 : P_1 = P_2$
   - $H_1 : P_1 \neq P_2$ (two-tailed), or one-tailed as appropriate
2. **Test statistic:**
$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$, $Q = 1 - P$.

3. **Level of significance:** Usually $\alpha = 0.05$.
4. **Critical value:** Obtain from standard normal tables.
5. **Decision:** Reject H_0 if $|Z| > Z_{\text{tabulated}}$.

Example:

Machine 1: $x_1 = 21$ defectives out of $n_1 = 500$
Machine 2: $x_2 = 3$ defectives out of $n_2 = 100$

$$p_1 = \frac{21}{500} = 0.042, \quad p_2 = \frac{3}{100} = 0.03$$
$$P = \frac{500 \times 0.042 + 100 \times 0.03}{600} = 0.04, \quad Q = 0.96$$
$$Z = \frac{0.042 - 0.03}{\sqrt{0.04 \times 0.96 \left(\frac{1}{500} + \frac{1}{100}\right)}} = 0.571$$

At $\alpha = 0.01$, $Z_{\text{tab}} = 2.58$. Since $0.571 < 2.58$, accept H_0. There is no significant difference in performance.

t Test

The t test is used when the sample size is small ($n \leq 30$), the sample is drawn from a normal population, and the population standard deviation is unknown. The t distribution is similar to the normal distribution but has heavier tails. As the sample size increases, the t distribution approaches the normal.

t tests are used for:

- Significance of a single mean
- Significance of difference between means
- Significance of correlation coefficient
- Significance of regression coefficient

Other Tests

- Chi-square test
- ANOVA (Analysis of Variance)
- Run test
- Sign test
- Mann-Whitney U test

- Kruskal-Wallis test

Lecture 23: Preparation of Research Report

Research Report

A research report is a concise, clear communication of the important findings of research work. It is a statement or description of things that have already occurred, compiling information as a result of research and data analysis. Reports focus on transmitting information with a clear purpose to a specific audience. Good reports are accurate, objective, complete, well-written, clearly structured, and expressed in a way that holds the reader's attention and meets their expectations.

Purpose and Importance of Research Reports

- To organize data, analysis, and conclusions in a form that can be used for academic or application purposes.
- The research report is the main tangible product of your work and is the primary basis for evaluation by examiners or committees.
- The quality of your research will largely be judged by the report, not by your fieldwork.
- A good report demonstrates your performance, skills, and thoughts, which are vital for assessment and grading.
- Reports are used by organizations, professors, researchers, and students for various purposes.
- Writing the report helps you:
  - Monitor progress and spot problems in time.
  - Reflect on progress, consolidate arguments, and identify gaps in knowledge, data, or methodology.
  - Develop an appreciation of standards and learn to monitor your own progress.
  - Practice academic report writing and discourse.
  - Form a basis for future project work or journal articles.

Research Report Process

The process of preparing a research report involves several steps, where raw data collected from different sources is gradually compressed and organized. The main stages are:

1. **Collection of Raw Data:** Gather data from various sources.
2. **Processing and Analysis:** Process and analyze the collected data to extract meaningful information.
3. **Interpretation:** Interpret the results in the context of the research objectives.
4. **Presentation:** Organize and present the findings in a structured report.
5. **Communication:** Communicate the findings effectively to the intended audience.

Qualities of a Good Research Report

A good research report should be:

- **Accurate:** Free from errors and faithfully represents the research findings.
- **Objective:** Unbiased and based on facts, not personal opinions.
- **Complete:** Covers all relevant aspects of the research.
- **Clear and Concise:** Written in simple, straightforward language.
- **Well-structured:** Logically organized with a clear flow from introduction to conclusion.
- **Relevant:** Focused on the research objectives and the needs of the audience.

General Structure of a Research Report

Although formats may vary, a typical research report includes the following sections:

1. **Title Page:** Title of the research, author's name, affiliation, and date.
2. **Abstract:** A brief summary of the research objectives, methods, findings, and conclusions.
3. **Table of Contents:** List of chapters and sections with page numbers.
4. **Introduction:** Background, statement of the problem, objectives, and significance.
5. **Literature Review:** Summary of previous research and theoretical background.
6. **Methodology:** Description of research design, sampling, data collection, and analysis methods.
7. **Results:** Presentation of findings using tables, charts, and descriptive text.
8. **Discussion:** Interpretation of results, implications, and comparison with previous studies.

9. **Conclusion and Recommendations:** Summary of findings and suggestions for future work or policy.
10. **References:** List of sources cited in the report.
11. **Appendices:** Supplementary material such as questionnaires, raw data, or additional tables.

Tips for Writing a Good Research Report

- Use clear, simple, and precise language.
- Be objective and avoid personal bias.
- Present data and analysis logically.
- Use tables, figures, and charts for clarity.
- Revise and proofread to eliminate errors.
- Ensure proper citation and referencing.
- Tailor the report to the needs and expectations of the audience.

Lecture 24: Conventions and Layout of Academic Research Writing

Conventions of Academic Writing

- Write direct, positive sentences using familiar words and short, simple constructions.
- Avoid unessential, overly technical, or unusual words and phrases.
- Well-constructed, natural, and direct sentences are a mark of skill in writing.

Presentation

- Label charts, sections, sub-sections, and tables adequately.
- Keep the system of headings and subheadings simple.
- Ensure the sequence of sections and subsections is logical and clear.
- Use past tense for introduction, data analysis, and findings; present tense for conclusions; and future tense for recommendations.

Use of the First Person

- Academic reports should be written in the third person.
- Avoid pronouns such as I, my, mine, our, ours, we, us, and me.
- If necessary, refer to yourself as "the writer" or "the investigator."

Use Gender-neutral Language

- Select terminology that treats all genders equally.
- Do not make assumptions about one gender over another.
- Use "he or she" or similar constructions when necessary.

Avoid Emotional Terms

- Prefer factual statements over emotional or subjective descriptions.
- For example, state the percentage increase in sales rather than using terms like "tremendous" or "fantastic."

Label Opinions

- Facts are preferred, but specialist opinions may be included when facts are unavailable.
- When presenting opinions, reveal the background and identity of the person if relevant.
- Opinions can substantiate explanations and conclusions, especially when data is inconclusive.

Use of Notes and Footnotes

- Supplementary material inappropriate for the main text may be included in footnotes (bottom of page) or notes (end of chapter).

Non-English Terms and Expressions

- Commonly used non-English terms in English do not need italics.
- Less familiar expressions (e.g., *chakka jam*) should be italicized.
- Consider whether the term will be understood by most readers.

English and American Spellings

- Use a consistent spelling format (either -ise or -ize endings) throughout the report.
- Both British and American spellings are acceptable, but consistency is essential.

Abbreviations

- Use abbreviations sparingly, as they can disrupt the flow and readability.
- On first use, provide the full term followed by the abbreviation in parentheses.
- Use the abbreviation alone thereafter.

Confidentiality and Anonymity

- Maintain confidentiality and anonymity as much as possible.
- Use fictional names for case studies if needed to protect identities.

Consistency

- Maintain consistency in spelling, abbreviations, style, and formatting throughout the report.
- Choose a format and adhere to it strictly.

Typing the Research Report

Paper

- Use white Xerox paper, size 8.6 by 11 inches.
- Type on only one side of the paper.

Chapter Page

- Center the chapter number about two inches from the top of the page.
- Place the chapter title in capital letters two spaces below the chapter number.
- Begin the first line of text four spaces below the title.

Margins

- Use 1-inch margins on all sides (top, bottom, left, right) as per APA guidelines.

Spacing

- Double-space the main text.
- Single-space indented quotations and footnotes.
- Use the same style and size of font throughout.

Page Number

- Place the page number at the top right corner, one inch from the top and right edges.
- The first line of text should be two spaces below the page number.

Pagination

- Number pages consecutively in Arabic numerals from the first page of text to the end, including appendices.
- Introductory sections (preface, table of contents, etc.) use small Roman numerals (i, ii, iii, etc.), placed one inch from the bottom.
- Page numbers should stand alone, without periods, hyphens, or dashes.

Proofreading

- Review the manuscript critically for inaccuracies, omissions, inconsistencies, and errors in quotations, footnotes, tables, figures, paragraphing, sentence structure, headings, spelling, style, and bibliography.
- Mark corrections clearly for the typist.

Layout of the Research Report

A comprehensive layout should include:

(a) Preliminary Pages

- Title and date
- Acknowledgements (preface or foreword)
- Table of contents
- List of tables and illustrations

(b) Main Text

- **Introduction:** State objectives, background, hypotheses, definitions, methodology, sample details, statistical analysis, scope, and limitations.
- **Statement of findings and recommendations:** Present findings and recommendations in non-technical language; summarize if extensive.
- **Results:** Detailed presentation of findings with supporting data, tables, charts, and validation. Present all relevant results in logical sequence.
- **Implications of the results:** Discuss inferences, limitations, unanswered questions, and suggestions for further research.
- **Summary:** Briefly restate the research problem, methodology, major findings, and conclusions.

(c) End Matter

- Appendices (technical data, questionnaires, sample info, mathematical derivations, etc.)
- Bibliography of sources consulted
- Index (alphabetical listing of names, places, topics with page references)

Lecture 25: Research Proposal

Research Proposal

A research proposal presents the author's plan for the research they intend to conduct. Its purpose is to demonstrate the relevance, necessity, and feasibility of the proposed research, often for approval by a supervisor, department, or funding body. In some cases, submitting a research proposal is a requirement for graduate school applications or for securing funding.

A well-written research proposal shows:

- How and why the research is relevant to the field.
- That the research fills a gap, supports, or adds new knowledge to existing literature.
- The author's capability and academic background to carry out the research.
- The academic merit and feasibility of the proposed ideas.
- The methodology, tools, and procedures for data collection, analysis, and interpretation.
- Consideration of budget and institutional or programmatic constraints.

The proposal must also include a literature review, which details the sources to be used, how they will be used, and their relevance to the research.

Length of Proposal

Research proposals for bachelor's and master's theses are typically a few pages long, while those for Ph.D. dissertations or funding requests are longer and more detailed. The focus should be on including all necessary elements, not on meeting a specific word or page count.

Research Proposal Structure

A standard research proposal generally includes the following sections:

Introduction

- Introduces the research topic.
- States the problem statement and research questions.

- Provides context for the research.

In some cases, an abstract and/or table of contents may precede the introduction.

Background Significance

- Explains why the research is necessary.
- Relates the research to existing studies.
- Clearly defines the problems addressed.
- Outlines the scope and any related questions not covered.

Literature Review

- Introduces and discusses all sources relevant to the research.
- Explains how each source will be used and its significance.
- Goes beyond listing sources by analyzing and contextualizing them.

Research Design, Methods, and Schedule

- Specifies the type of research (qualitative, quantitative, experimental, correlational, descriptive).
- Describes the data, population, and subject selection.
- Details data collection tools and methods (sampling frame, sampling method, statistics, experiments, surveys, observations).
- Justifies the chosen methods.
- Includes a research timeline, budget, and anticipated obstacles with plans to address them.

Suppositions and Implications

- Discusses how the research may challenge existing theories or assumptions.
- Explains the potential for future research based on the findings.
- Describes the practical value and possible impact of findings (for practitioners, educators, policymakers, etc.).
- States how findings could be implemented and their transformative potential.

Conclusion

- Summarizes the proposal and reinforces the research's purpose.

Bibliography

- Lists all sources referenced in the proposal, formatted according to the relevant academic style (APA, IEEE, Chicago, etc.).
- Sometimes a references list is sufficient for shorter proposals.

How to Write a Research Proposal

- Write in a formal, objective, and concise academic tone.
- Follow the standard structure so the reader can easily follow and evaluate your proposal.
- Ensure every section answers potential questions from the reader.

Common Mistakes to Avoid

- **Being too wordy:** Keep writing brief and to the point.
- **Failing to cite relevant sources:** Reference landmark studies and connect your work to the field.
- **Focusing too much on minor issues:** Address only the key questions and issues central to your research.
- **Failing to make a strong argument:** Clearly persuade the reader of your research's importance and feasibility.

Polishing Your Proposal

- Ensure the proposal is free from spelling and grammatical errors.
- Maintain an appropriate, consistent academic tone.
- Revise awkward phrasing to strengthen clarity and credibility.