

# Fundamentals of Research Methodology

## A Practical Approach



By

Dr. Purna Chandra Pradhan

Dr. Prosper Kwame Klu

Dr. Dipak Chaulagain

Dr. Harshitha Y S



ISBN: 978-81-988164-3-6

# **Fundamentals of Research Methodology**

## **A Practical Approach**

By

**Dr. Purna Chandra Pradhan**

**Dr. Prosper Kwame Klu**

**Dr. Dipak Chaulagain**

**Dr. Harshitha Y S**

ISBN 978-81-988164-3-6



9 788198 816436

Mode of publication: Print | Non-Print

Copyright © 2025: All Rights reserved by the Author/s

Month / Year: June, 2025

Published by:

A2Z EDULEARNINGHUB LLP

Door No. X/374, Kuthukallinkal

Udumbannoor, Cheenikuzhi P.O

Kerala, India - 685595

[www.edulearninghub.com](http://www.edulearninghub.com)

[editor@edulearninghub.com](mailto:editor@edulearninghub.com)

Total pages:134

Price: INR 475 / USD 7

ISBN: 978-81-988164-3-6

## **Preface**

Research stands as one of humanity's most profound endeavors, driving innovation, exploration, and the continuous expansion of knowledge across all disciplines. At its core, research is a systematic and organized investigation designed to deepen our understanding of specific topics, answer important questions, and solve pressing problems. Whether conducted in business, education, industry, medical sciences, military, or social sciences, research remains the cornerstone of advancement and intellectual growth. This book is dedicated to illuminating the multifaceted nature of research. It emphasizes that while research projects may differ in form and focus across disciplines, they are unified by a common goal: to contribute meaningfully to the broader landscape of human understanding. Research is not merely the collection of facts; it is a dynamic and iterative process that involves systematic inquiry, critical interpretation of data, formulation of hypotheses, and rigorous validation of ideas.

Furthermore, this work highlights the vital distinction and interdependence between research and research methodology. Research methodology provides the structured theoretical framework and procedural guidelines that inform the investigation, ensuring that every step, from formulating a problem to drawing conclusions, is carried out with scientific rigor and integrity. Together, research and research methodology create a foundation for meaningful inquiry that can lead to new inventions, improved procedures, and innovative solutions where conventional techniques fall short. As readers engage with this text, they will gain a deeper appreciation for the discipline, creativity, and intellectual curiosity required to conduct impactful research. It is our hope that this book serves as a valuable guide for scholars, practitioners, and students alike, inspiring meticulous inquiry and fostering a lifelong commitment to discovery and knowledge creation.

**Authors**

## **Fundamentals of Research Methodology : A Practical Approach**

### **Chapter 1: Introduction to Research.....07**

- Meaning, objectives, and importance of research
- Types of research: exploratory, descriptive, analytical, fundamental, applied
- Characteristics of good research
- The research process from idea to outcome

### **Chapter 2: Research Problem and Hypothesis.....16**

- Identifying and refining the research problem
- Formulating objectives, questions, and hypotheses
- Variables and operational definitions
- Understanding causality and correlation

### **Chapter 3: Review of Literature.....31**

- Purpose and strategies for reviewing literature
- Locating scholarly sources using Google Scholar, Scopus
- Organizing with reference managers such as Mendeley or Zotero
- Writing a structured review for a thesis or article

### **Chapter 4: Research Design and Sampling Techniques.....41**

- Research design types: cross-sectional, longitudinal, experimental, quasi-experimental
- Sampling methods: simple random, stratified, cluster, purposive, snowball
- Determining appropriate sample size for studies using formulas or software

### **Chapter 5: Data Collection Methods.....53**

- Primary vs secondary data
- Tools for surveys, interviews, and observations
- Online platforms like Google Forms, Microsoft Forms, and mobile-based apps
- Ensuring validity, reliability, and ethical data collection

**Chapter 6: Measurement and Scaling Techniques.....67**

- Levels of measurement: nominal, ordinal, interval, ratio
- Scaling techniques: Likert, Guttman, semantic differential
- Creating reliable and consistent research instruments

**Chapter 7: Data Analysis and Interpretation.....83**

- Organizing data using coding and tabulation
- Descriptive statistics: mean, mode, median, standard deviation
- Inferential techniques: t-test, chi-square, ANOVA, correlation
- Data analysis using SPSS, Excel, and R (basic commands and charts)
- Interpreting output in the context of research questions

**Chapter 8: Qualitative Research Methods.....97**

- Nature of qualitative research and its applications
- Methods: interviews, focus groups, case study, content analysis
- Manual coding and software like NVivo or Atlas.ti
- Ensuring trustworthiness: credibility, transferability, dependability

**Chapter 9: Report Writing and Research Ethics.....108**

- Structuring research reports, dissertations, and papers
- Plagiarism and citation standards (APA, MLA, Chicago)
- Ethical clearance, informed consent, and participant confidentiality
- Using Turnitin and Grammarly for ethical and polished writing

**Chapter 10: Publishing Research and Academic Visibility.....120**

- Journal types, indexing (UGC-CARE, Scopus, Web of Science)
- The peer-review process and avoiding predatory journals
- Creating and maintaining academic profiles: ORCID, ResearchGate, Google Scholar
- Tracking citations and metrics using Publish or Perish, Scopus tools

**“ The goal of research is not only to discover new knowledge but also to transform society with that knowledge.”**

— Dr. APJ Abdul Kalam

## CHAPTER 1

### INTRODUCTION TO RESEARCH

Research is both the process of collecting data -- and of using that data to develop generalizable knowledge. In this case, we mean the research that's conducted using scientific methods. Some researchers, and some kinds of research, focus on collecting data in very formal and structured ways. Other researchers use much looser and informal structures to gather information. The objectives of research are the enhancement of human knowledge and finding answers to our questions. Why do the planets circle around the Sun, rather than being scattered through the Universe? Why do children learn their mother language so quickly and easily, whereas it is so hard for adults to learn a second language? All these questions stimulate the human mind, and research attempts to provide answers to such questions. Human knowledge has been evolving incrementally ever since the dawn of conscious thinking. Research is important because it is the vehicle through which you explore and enrich your understanding of the world. Exploration is not limited to research in the natural and social sciences; the humanities have also contributed much to our understanding of the wonders of man and his environment. Even the applied sciences, while using a more applied approach, do rely on research in the natural and physical sciences.

### MEANING, OBJECTIVES, AND IMPORTANCE OF RESEARCH

Research is an essential aspect of academic and professional pursuits, defining activity that leads to mastery of knowledge. It is an element of formulation, testing, confirming, and disputing knowledge in the realm of theory or practice and leading to the development of new knowledge. The term research comes from the French word meaning to search for something again. It essentially means more than simply looking for something. Research, on the other hand, entails a careful search for facts, then analysis of the facts, and finally synthesizing them into a more coherent picture that fills the gaps and extends the limits of previous knowledge. Research is thus an inquiry that is both in-depth and comprehensive. In the formal academic sense, research is the pursuit of knowledge through a structured process. However, with recent technological advances leading to increasingly easy access to the internet and large amounts of keyword indexed data, there is a strong temptation to take shortcuts and provide research work that is superficial at best. The increased emphasis on using factors that reflect performance and cross-discipline influence of journals as a proxy

measure of research quality has led many researchers to take less than ethical shortcuts to present their research to propagate knowledge.



Figure 1.1 Research objective

### **TYPES OF RESEARCH: EXPLORATORY, DESCRIPTIVE, ANALYTICAL, FUNDAMENTAL, APPLIED**

The variety of disciplines studied will produce an equally varied list of research methods. Each research activity tends to fall into one of the following categories: exploratory, descriptive, analytical, fundamental, or applied. However, these types are not strict categories. Researchers will often use a combination of techniques in one study, or use a single technique across several studies. Exploratory research is a preliminary stage of research that helps the researcher gain as much background information as possible before filing a research question. At this stage, the researcher is trying to familiarize themselves with the problem area, while also developing a clearer understanding of how to formulate a viable research question. In this phase, the researcher may speak with an expert, or conduct in-depth interviews and focus groups with a small group of respondents. Case studies are another commonly used source of exploratory research. Due to the lack of large-scale information and the lack of control over a causal relationship, exploratory research cannot be conclusive. Descriptive research answers the how, what, when and where of a research question. Surveys are the most popular tool for collecting descriptive research data. There are several types of surveys that varied in the way results are collected. Surveys can be conducted over the phone, via paper, in person, or through web-based services. In using a survey, attention must be paid to the survey design, as a poorly designed survey can produce questions that result in unreliable responses. A survey

must also be given to a sample population that's large enough to be representative of a larger population, and the right analysis must be employed to evaluate the data from the survey.

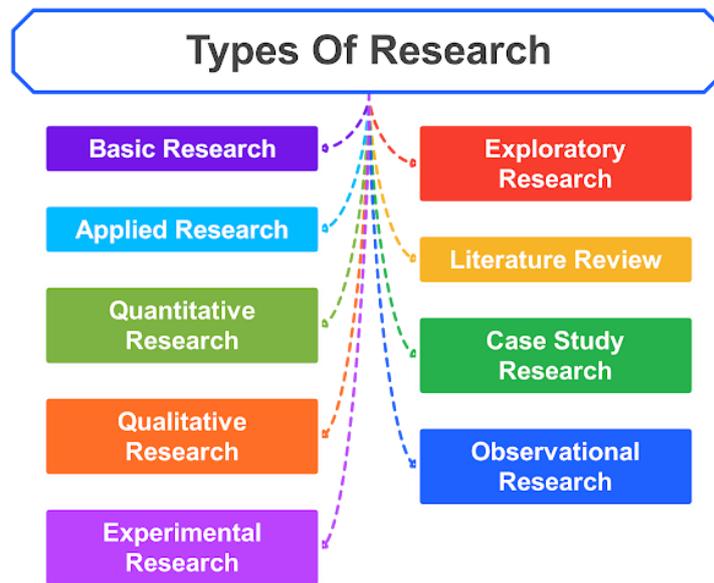


Figure 1.2 Types of research

### ➤ Exploratory Research

Richard H. Thaler quoted Einstein: "If I had one hour to solve a problem, I would spend 55 minutes thinking about the problem and five minutes thinking about solutions." What Einstein meant is that any researcher ought to spend sufficient time thinking about the problem before trying to solve it. "The Challenge" or "The Research Question," the essence of any research, comes at the end of exploratory research. Exploratory research has a special role in the research process. Exploratory research is undertaken to obtain insight into an issue gone through little or no prior work. Often the goal is to formulate the issue in a way that can be addressed through further evidence or modeling. The research might take the form of a detailed pilot study, or simply a gathering of evidence regarding in-field expert opinions. Exploratory research typically precedes both causal and descriptive research. It can also be the first step following a realization that the research described earlier either failed to address an inspiring hypothesis or perhaps flunked attempting to do so and/or examine in greater detail either a newly-formulated hypothesis, or a different aspect of an earlier analysis. A hallmark of exploratory research is that it has come before using scientific techniques, or its

unscientific status may to some extent remain burdened in even follow-on causal or descriptive research.

➤ Descriptive Research

Descriptive studies aim to accurately and systematically describe a population or phenomenon, answering the questions of who, what, when, where, and how. Descriptive research has historically been considered a subset of and a precursor to exploratory research, but is occasionally characterized as a distinct form of research. The area of descriptive research is vast and includes the use of both quantitative and qualitative methods. The focus of this entry is on quantitative methods. Describing a population often requires testing hypotheses, such as testing the effectiveness of a policy intervention. Researchers often create a causal model that describes how the mechanisms linking the independent variable to the dependent variable work, which may or may not invoke unobservable factors shaping behavior. The rationale for a descriptive study may include providing necessary groundwork for developing a causal model or making a recommendation on a policy intervention. Descriptive studies may also validate or calibrate a previously developed model, generalize behavioral relationships to a population whose characteristics differ from those of the reference population, or provide an indication of likely relationships among variables for use in decision-making. Regardless of the rationale, observational studies must be designed thoughtfully, using established statistical techniques to avoid pitfalls commonly encountered in conducting observational research. Key issues to consider in the study design include questions such as how well and how soon will you be able to identify the incident cases, what data will you need to collect, how long will you need to follow the subjects, how will your sample size affect your estimates, and how will your ethical requirements affect the design?

➤ Analytical Research

Analytical research, which is understood to be an explanatory study, establishes a causal relationship between two or more dimensions. The research questions associated with this type of research center on explaining why specific phenomena occur. These types of questions usually begin with the words which, who, when, how, and why. Analytical studies are typically performed when some relationship between the studied variables has been previously established or hypothesized and what the researcher attempts to do is confirm that hypothesis. When there is sufficient previous research on a topic and researchers have accumulated enough knowledge, it may be possible to develop a causal statement that allows

finer predictions. However, these statements are always formed about probabilistic properties of variables, that is, given conditions, it is possible to predict the behavior of some variable, but not the behavior of that variable precisely. It is useful to clarify that analytical studies include the generation of new knowledge; the difference between exploratory and analytical studies is in the prior knowledge of the topic and the nature of the questions posed. In the social sciences, the vast majority of variables are multicausal, and although this makes the analytical studies difficult, it does not prevent them from being carried out. Nevertheless, one limitation imposed by some social scientists to the possibility of developing analytical studies is the empirical restriction of “no perfect collinearity,” that is, the necessity of not having a linear relationship with a slope coefficient equal to zero between the variables that specify a multiple regression model. This limitation is imposed in order to avoid indeterminacy in the estimation of the parameters of the linear regression model. In this way, analytical studies typically seek macroexplanations of how some type of institutional arrangement influences the behavior of a relevant variable at the aggregate level, and the empirical validation of these relationships is made through the estimation of a linear regression model.

➤ Fundamental Research

Before we delve any further, it is prudent to define the different kinds of research we may encounter in the course of our work. One of the first distinctions made in Research Methodology is between Fundamental – or Basic – research and Applied Research. Fundamental Research is directed toward a more profound comprehension of the underlying facts and principles of nature, without regard to the practical benefits to be derived from such study. Basic research makes no demands on the part of a field of research; rather, the results speak for themselves. The terms of the final definition reflect the intrinsic value of such discoveries; as with contributions to the Theory of Everything and the pursuit of knowledge for the sake of knowledge. Others, however, see significance in the potential for expanding general knowledge that may be of use in the future. An example of such functionally basic research would be the work done to confirm that the mass of the neutrino would be less than the mass of the electron. An example of Functional Basic Research might be the use of high-energy physics to study the structure of atomic nuclei, prior to the modeling of nuclear power plants and the motivation for that study. Another example could be Basic Research on the nature of light; functions of purely intellectual inquiry that later led to Applied Research on the use of laser beams as guidance for the precision machining of semiconductor chips. Other examples include basic research on fermentation, which led to the application of this work to

the production of yeast and vinegar and the pasteurization of milk, and the application of a process developed by others.

➤ Applied Research

Applied research aims to solve specific, practical issues impacting a particular person or group of people. Most of the time, applied research is used to improve a situation, as opposed to understanding it only, making it a type of knowledge construction. Other than being problem-oriented, applied research usually relies on quantitative methods and takes their foundations from the already available theoretical explanation, typically shaped by fundamental research. Applied research may also be conducted for a particular group or company. In this context, it is carried out under an agreement between the parties involved. The leading research center of a multinational corporation may be in charge of defining the criteria for the development of a product, deciding to invest in such a project, and then hiring external researchers to carry it out. However, it must not be considered an exclusive prerogative of the private sector, with many public research centers assigned to it as well. In this case, though, the main reason for carrying out this project is to increase the overall competitiveness of an economic sector, the results becoming public. The latter might be defined as a pure contract or mixed contract research, with companies or economic institutions becoming increasingly aware that the success of certain applied research projects depends to a large extent on their relation to fundamental research and on the actual transfer of new knowledge. In other words, pure mixed contract research projects that cannot take advantage of certain scientific results have proven less successful. Research is considered applied research when it seeks to answer a question or solve a problem in a specific situation.

### **CHARACTERISTICS OF GOOD RESEARCH**

Research is a systematic investigation, such as a scientific or scholarly study, to discover facts or collect information on a subject. When someone is interested in research, he is interested in knowing the answer to a question or discovering something about a particular field. Research is important because it enhances and increases the body of knowledge that is available for people. Good research uses neutral and risky methods so that the findings are valid and not misleading. There are a number of characteristics that make research good or of higher quality, and they are as follows.

1. A Good Research Should be Goal-Oriented: The research ought to be guided by pre-specified objectives. It must exemplify not only the highest ethical standards but also promote the collective well-being of humanity. Whether academic or applied, a good research should be able to resolve human problems. Most scientists, while designing their research projects, will be heavily motivated by the desire to achieve an improvement in some aspect of the human condition, health, environmental, economic, or physical.

2. A Good Research Should be Systematic: A good research should be systematic. It does not simply involve curiosity and exercise of intelligence. The gathering of knowledge may be accidental and haphazard but not research. A good research is a systematic and organized investigation to discover the answers to questions. It is characterized by carefully thought out procedures and methodical data collection. It is not merely the use of accepted techniques but also the decision process to use those techniques and the chosen design that distinguishes research from other professional endeavors.

3. A Good Research Should be Empirical: A good research should be empirical. In scholarly research, most observations involve the collection of quantitative information as numerical data that can be analyzed to test hypotheses. A good research does not merely attempt to describe but explains. In general, research attempts to answer questions about observable phenomena. Such phenomena have been observed either through direct or surrogate observation using reliable indicators.

## **THE RESEARCH PROCESS FROM IDEA TO OUTCOME**

Much of the time these roles seem distinct from one another, which can lead to the impression that research is a sequence of separate and consecutive steps; asking a question, gathering data and so on. Phases can indeed be conceptualized this way, each accomplished one lapse or another by different practitioners. However, research is ordinarily a much messier process in practice, not always so linear but one in which each can heavily inform the others. We begin by describing a linear sequence and will therefore outline three phases through which researchers typically pass. Because it inevitably involves choice, topic or problem selection is the most important early stage in the research process. The areas, or vast expanse, of human activity that is possible to explore is virtually infinite in its breadth and depth: no person has the capacity to investigate all of it. Social scientists and everyday observers have commented on the rich diversity of societies and cultures throughout the world, suggesting highly varied avenues of exploration, inquiry, and investigation into the

nature of being human. However, what is to be investigated, and why that topic, is obviously a crucial question which underlies any serious academic inquiry. Most researchers would agree that the selected topic should be both interesting to the researcher and relevant to our knowledge of the particular area examined. What could be worse than spending a year or more studying a topic that is dull to the researcher herself? In traditional empirical research practice, selection is of the superficial, and supererogatory, kind and involves choosing the topic within what is sometimes called “the framework” or “the criteria” of the field previously thoroughly canvassed and established by prior research; those regarding the significance and capacity of such research to further increase insight and understanding of the area of interest. The topic has thus selected then informs the choice of the particular specific problem the research is meant to address.

Research remains, at least in the hard sciences, an arduous endeavor associated in the popular imagination with the lonely and obsessive figure of the scientist or an academic hermit driven by blind curiosity to venture into unexplored territories of knowledge. A more pragmatic approach informs intellectual development in the humanities and social sciences, and involves the methodical solution of specific research problems arising within the context of a scientific discipline and well identified by its community of practitioners. Although this contrast in the public imagination of research in the two realms is perhaps overstated, there is a difference in style that calls for comment. Teachers of undergraduate students will agree that both disciplines demand a lengthy period of preparation during which the student’s intellectual development is nudged in the right direction. The research initiative came all but spontaneously with the task of writing a dissertation: the choice of topic and the range of sources to consult were left in the hands of each student. Editors of scholarly publications will agree that while research in the social sciences and humanities is not lacking in originality or creativity, it is far more formulaic than in science. The research plans of scientists enter by the dozens every month in every recognized research laboratory in the world. So do the outlines for publications in the title pages of original papers in learned journals. The satisfaction resulting from finally discovering more than you had expected waiting for you at the specified coordinates is unique, and it outweighs the consideration that someone in the past has probably pieced together the same data with even less effort. The inverse is not necessarily true. There is no disappointment like that which besets you when, after a long and arduous effort, you discover that what you have turned out is little more than a compilation of data readily available.

## Chapter summary

The first chapter lays the foundation for understanding the concept and significance of research. It begins by defining research as a systematic and logical process aimed at discovering new knowledge, validating existing facts, or solving problems. The objectives of research are multifaceted—ranging from gaining familiarity with phenomena (exploratory), providing accurate descriptions (descriptive), testing hypotheses (analytical), generating theoretical knowledge (fundamental), to solving practical problems (applied). The chapter emphasizes that research is central to academic and professional inquiry, providing the means through which disciplines evolve and decision-making is refined. The chapter also discusses various types of research, each with its unique purpose and methodology. Exploratory research is used when a problem is not clearly defined, often serving as a precursor to more detailed study. Descriptive research provides a picture of phenomena as they exist, while analytical research tests hypotheses and explains relationships. Fundamental research is aimed at developing theories and expanding knowledge without immediate practical application, whereas applied research focuses on real-world problem-solving. By highlighting these categories, the chapter provides learners with a framework to select the appropriate approach based on their research goals.

A crucial component of this chapter is the characteristics of good research. Effective research is systematic, logical, empirical, replicable, and ethical. It should be based on observable experiences and data, allow for verification, and follow a clear, structured methodology. The importance of objectivity, clarity of purpose, and careful planning is emphasized to ensure the validity and reliability of the findings. Furthermore, ethical considerations such as honesty, integrity, and respect for participants are highlighted as essential to sound research practices. Finally, the chapter outlines the research process, illustrating the journey from idea to outcome. This process typically includes identifying a problem, reviewing the literature, formulating research questions or hypotheses, designing the study, collecting and analyzing data, and drawing conclusions. Emphasis is placed on the iterative nature of research, where refining and revising are part of achieving clarity and depth. This step-by-step approach serves as a practical guide for beginners, enabling them to visualize how abstract concepts are translated into concrete outcomes through methodical inquiry.

\*\*\*\*\*

## CHAPTER 2

### RESEARCH PROBLEM AND HYPOTHESIS

In looking at how research proceeds we are usually informed, especially if we are good listeners, by the research of others. We make discoveries from this prior work, informally or formally evaluate and reject or accept some research as flawed or only partially developed, and in so doing develop interest in a work that at some point has been sufficiently informative, exciting, and novel to lead the author to make an initial research contribution. With the most fortunate or involved among us, reviewing this research becomes our life's work, and it leads to the development of criteria and infrastructure to support the work of others. Eventually, the reviewers or the packagers of the work help others by promulgating a discipline to aid others in searching the literature for new discoveries, and of course, these new discoveries inform our own future research problem development decisions.

### IDENTIFYING AND REFINING THE RESEARCH PROBLEM

Research is a planned search that develops knowledge in a particular area. It considers previously proven guidelines but is goal-oriented, involving both logical and empirical methods. Before research can be performed or evaluated, a problem must first be clearly expressed. Although it may seem easy to identify and rely on an existing research problem or hypothesis, performing original research requires careful and focused efforts tailored to the needs of the individual laboratory. Most students and researchers are never explicitly taught the skills and approaches to authoring a research proposal, and very little information on the topic is available. As new researchers transition into independence as faculty, clinicians, and independent investigators, they must create a project using knowledge accumulated during training and apply it toward new avenues of inquiry. Either implicitly or explicitly, each research proposal derives its originality and focus from an underlying question or hypothesis. As it applies to research, the question or hypothesis poses a request for inclusion and accepts unknown results. However, not all questions are created equal. Some questions will undoubtedly lean toward the impossible, while other questions may lie along the lines of playing God. The focus or importance of some questions may indeed lie on the fringe of scientific inquiry.

In every case, the request for inclusion that is presented in a research proposal must be focused, eligible, and garner interest from the scientific community. The significance of a

question, as well as a balance of possible and probable results, is what drives a successful research proposal. The question or hypothesis is the proposal's most valuable and, at times, difficult part to create. In this chapter, we will focus on the basic techniques and considerations of question creation and aim to present a guiding framework for doing so. Research problems have evolved to become more and more complex. Distinctions between successful problems, such as those that have broad significance across multiple fields of study, are often lost on new investigators. The steps of idea conception, existence, and gestation are unfulfilled within a research group if senior group members and mentors do not communicate the process to less experienced group members.



Figure 2.1 Research Problem

### ➤ Importance of a Clear Research Problem

Formulating or identifying the right problem is often considered the most important aspect of conducting research. This is true not only in the business area but also in other academic research areas. Problems can often be so complex and multi-dimensional that no single specific problem can be identified. In such situations, it is beneficial for the researcher to more carefully and meticulously analyze the problem situation before formulating a general research problem. Following this general direction or guidance, this chapter first discusses some aspects related to the importance of clearly specifying the research problem area. Defining and formulating the right research problem is one of the most critical steps in the research process. It is widely stated and accepted that a clearly stated and well-structured research problem helps to focus the research effort. It can help to refine the scope of the

research, minimize research biases, and thus assist in the selection of an appropriate research methodology and design that is best suited to the research problem. Neatly specifying the designated problem area also assists the researcher in determining the meaning and importance of the research results, and their relationship to prior research work. Further, it helps to clarify the contribution that the research makes to an understanding of the research problem area. Indeed, the importance of formulating a clear research question is often considered a hallmark of high-quality and reputable social scientific research.

- Techniques for Problem Identification

This section describes some of the common techniques for identifying problems that researchers can employ to build up a list of potential research problems. The technical information-processing principles and a typology of European Marketing research projects are presented.

- Diary Method of Problem Discovery

The Diary method is the traditional and widely accepted method used to identify research problems. The researcher will read published research and practitioner papers over a period of time to identify topics that are being repeatedly raised. The discoveries are usually grouped under conceptual problem areas.

- Examination of Technical Documents

Technical documents in the form of trade publications, applied research reports, and technical journals often provide useful inputs to identify problem areas that warrant further attention. Such sources help the researcher group the problems.

- Asking Prominent People and Focus Group Technique

Seek evaluation from knowledgeable people in the form of face-to-face interviews, telephone interviews, or by asking opinions in a focus group setting. What is added to the problems identified by the researcher by personal opinions is the evaluative inputs on the importance of the problems. Such informal discussions can provide useful clues to organize problems that require further examination.

- Earlier Research Studies

Explaining earlier studies made in the same area can help identify stereotypes in the found problems and different markers commonly present in the results. This technique can be more useful when modifying the design from previous studies that require minor changes.

- Criteria for Refining Research Problems

Once the researcher has specified the areas and population of interest together with general questions of interest, concrete research problems need to be formulated for practical and analytical purposes. These problems essentially provide the narrowed vision required for the research to be manageable and focused. The research problem serves as the focus of such inquiry and acts as a reference point in the whole research process. A clear research problem statement increases the efficiency of the research process since it clarifies the scope and task of the project, identifies the core variables and identifies their relationship. Clarification of research questions can improve both the research proposal and the report. A statement reflecting the main research problem, its implications and relevance helps researchers formulate subsidiary questions directing the development of specific hypotheses. Numerous postulated criteria exist to assist researchers in achieving a well-articulated problem or research question that would improve the quality of the inquiries results. Such criteria can have been formulated by following the model of communication. First, inquiring what, then why, and lastly how, the three types of communication questions emerge. The what questions typically inquire after the nature of the relationship that the researchers want to investigate. The why questions are those that refer to the specification of potential supporting or explanatory variables for that relationship. The how questions regard the methodology of the inquiry, the investigation's design, the data sources and methods of data collection and preparation, the method of statistical or qualitative data analysis and the time horizon of the research. The criteria for formulating a good research problem can conceptually span informativeness, internal coherence, external coherence, relevance and feasibility dimensions.

## **FORMULATING OBJECTIVES, QUESTIONS, AND HYPOTHESES**

Formulating Objectives, Questions, and Hypotheses Research does not exist in a vacuum. Before a researcher embarks on any research, there are four questions they have to ponder: What is the problem? Why is it a problem? How will the researcher know whether the problem has been solved? What does the researcher hope to achieve amongst presenting these

answers? The answers to these questions will need to be operationalized in the form of objectives, questions, and hypotheses. They serve as the roadmap, guiding the researcher through the journey of the research project and ultimately answering the problem statement.

**Defining Research Objectives** The overall goals of the research are expressed in the form of objectives. The objectives narrow down the information in the problem statement, establishing specifically what it is the researcher hopes to achieve in conducting the research. It helps the researcher by providing a basis against which the researcher can compare their results, as well as outlining the specific steps to take to arrive at their final conclusion. These objectives can also help to highlight the limitations of the research, allowing the researcher to avoid paralysis by analysis, maintaining their focus on the issue at hand.

**Crafting Research Questions** Most research projects are answered with a set of research questions, breaking down the problem statement into a smaller set of focused questions the researcher needs to answer to elicit the results needed. Questions are important as they prove the researcher's curiosity about the problem statement and are used as the basis for engaging with their intended audience; conducting interviews with participants in qualitative research, or designing a questionnaire to solicit feedback in quantitative research.

**Developing Hypotheses** A hypothesis is a specific statement about the expected relationship between variables. It is informative in that it tells the researcher which variables they need to examine, and breaks the research questions down to the simplest form. It also serves as the basis for any tests the researcher conducts to determine whether the data they have supports or rejects their hypothesis, with the research being classified as hypothesis-driven research.

#### ➤ Defining Research Objectives

Research objectives are specific objectives that the researcher wants to achieve by conducting research. They define what we want to achieve by doing our purpose. The research objectives are derived from the research purpose statement which is, in general, broader and more meaningful than research objectives. However, the research objectives must align and be consistent with the purpose of research. The research objectives are generally enumerated in number. The objectives can be classified as exploratory objectives and specific objectives. Exploratory objectives are preliminary objectives generally found in exploratory research. The exploratory objectives need to be further defined in specific objectives. The specific objectives are meant to help answer the research questions or the hypotheses. Generally, the specific objectives must relate to the various dependant and independent variables of research. The research objectives must be made clear to help measure and assess the research and its

results. Poorly defined objectives may lead to confusion in the research process, subsequently diluting research results. The research objectives must be described in detail in the research project proposal so that the objectives are understandable not only to the researcher but also to the sponsors, stakeholders of the research. This transparency promotes collaboration, advice, puts checks and balances to ensure the research stays on track, and assists in answering the research questions. Some amount of exploratory work is usually required for many research topics before appropriate objectives can be defined. Data collection and analysis procedures are then developed based on the objectives. Exploratory research can play an important role in both quantitative and qualitative approaches.

### ➤ Crafting Research Questions

The world is full of mysteries waiting to be uncovered — often the common practice of asking a question and reading the answer is sufficient to do the trick, but generally it is more complex. There are numerous questions we could ask, often countless possible answers. Most of those questions won't yield a satisfactory, insightful response. The careful selection of a careful research question is what separates professional researchers from amateurs. The question we select will greatly impact our research. Thus, it should be definable and appropriate with respect to time, deliverables, and our skills. The better our question is, the easier it will be to generate our insights and the more impact our findings will have. As explained so far in this section, the research problem evolves from more or less broad research objectives, and leads to more specific questions. The research problem in turn informs the goals put into an agreement of a research study or project, either personal or business-related. The questions themselves cannot be too broad — otherwise the research may take forever and deliver little value. But, they cannot be so tightly defined that there is no leverage for real discovery and innovation. It takes experience and several iterations until the question design is good enough to proceed with research. Moreover, often it becomes obvious during the work on the project that we actually want to know something else, and the questions need to be changed accordingly. It is important that all team members have sincerely aligned interests, understand what is expected to happen, and use the same language.

### ➤ Developing Hypotheses

Hypotheses form a key part of the ideas that form an integral part of whichever the research problem may be. The research hypothesis mentions the dependent and independent variables clearly and refers to some underlying patterned relationships among the varying variables

involved. It helps in research design as it mentions the relationship about the variables that can be tested and validated. Any hypothesis formed has to be either true or false and must be stated simply. The hypothesis of a study is a statement of the relationship which we expect to exist among the variables. Our clinical intuition, or theory should suggest possible relationships or groups of relationships, either general or specific. Empirical research aimed at explaining and predicting in the particular area requires the specifying of testable hypotheses. Suppose a research interest area led the researcher to a group of subjects who were especially alike with respect to several variables. The goal of research may lead the researcher to state a hypothesis predicting that the variables differ in some important manner. Whenever possible and when resources are available, the researcher should consider verifying this prediction before working with the larger group. Eager to identify significant explanation or prediction, the researcher starts with a broad area of interest. The researcher considers ideas that could be translated into a specific explanatory or predictive formulation, and then searching for methods that could be compiled into a formal empirical investigation. The problem definition focuses the research process, to state testable hypotheses conditions of possible generality, and expects to discover empirical violations to test the limits of applicability.

#### ➤ Types of Hypotheses

Hypotheses serve as specific expressions of predictions generated by theory, each of which is then tested for its validity. However, the specific type of prediction expressed by each hypothesis differs, depending on the type of hypothesis being examined. There are several types of hypothesis pertaining to some independent or predictor variable causing or predicting some dependent or criterion variable, including initial, directional, etc. An initial hypothesis expresses a prediction that has no particular direction and only states that a relationship exists between the independent and dependent variables under consideration. In contrast, a directional hypothesis makes a specific prediction about the nature of the relationship between two variables. Some researchers refer to directional hypotheses as more “powerful” hypotheses or predictions, because the hypotheses are more difficult to fray, requiring researchers to muster more evidence. Hypotheses can also be classified as positive versus negative. A positive hypothesis asserts that a change in the independent variable will result in a change or effect on the dependent variable. A negative hypothesis asserts that a change in the independent variable will not result in a change or effect on the dependent variable. There is also a classification scheme that divides hypotheses into simple versus

complex. A simple hypothesis states the relationship between two variables only, consisting of an independent and dependent variable. A complex hypothesis states the relationship between three or more variables, consisting of an independent variable and two or more dependent variables. Hypotheses can even be classified, according to the number of times they have been examined, as first-use versus multiple-use.

## **VARIABLES AND OPERATIONAL DEFINITIONS**

In the world of research, a fundamental yet complex phenomenon observed is the fact that we cannot study the entirety of any specific issue or possibility. This is caused by theoretical and practical limitations. Irrespective of the efforts made, it is always practically impossible, due to various constraints, to explore every single person in the public domain, thereby making research in which the empirical data is derived from a representative sample indisputable. Thus, it is of extreme importance to understand that in research, a multitude of variables (or random quantities) is present, which become clearer when discussing predictions. Hence, what becomes essential is understanding the variables within this domain.

**Understanding Variables** A variable can be defined as a characteristic that can take on different values for different individuals (persons or objects) at a single time. In simpler terms, anything that differs over individuals and can take on at one point of time at least two different values is known as a variable. A particular value of a variable for an individual is termed the "attribute." The value for an individual is often denoted as  $X_{ij}$ , where "i" stands for "i-th" individual, "j" for the "j-th" variable. If there are "I" individuals and "J" variables, then the data matrix consists of "I" rows and "J" columns. A specific attribute is itself an "i-th" sample of "j-th" random variable. It is also often "j-th" component of "I" samples taken from "J" random variables, where each random variable corresponds to an individual, and each sample is a vector of "J" components comprising the attributes corresponding to the "J" random variables.

**Types of Variables** Based on the type of information they convey, variables can further be classified into the following types: 1. Qualitative variables 2. Quantitative variables 3. Discrete variables (or count variables) 4. Continuous variables

**Operationalizing Variables** The concept of variables is so central to the process of scientific research that it is surprising how little is known about the way scientific variables acquire their meaning. A variable is a qualitative or quantitative property whose value may vary from one sample to another. All variables are generally classified into two major types: 1) Qualitative variables. 2) Quantitative variables. A quantitative variable is one whose possible values are numbers and that represents a measurable quantity. Qualitative variables are variables that can take on one

of a limited and usually fixed number of possible values, assigning each individual or other unit of observation to a particular group or nominal category on the basis of some qualitative property. Qualitative variables are also known as attribute variables.

### ➤ Understanding Variables

A variable is some characteristic of an object that can take on different values. It is a property, but one that changes or varies. A research variable must have at least two values or levels, although it may have common values throughout the study. The independent variable is manipulated or controlled by the researcher. The dependent variable's value is affected by the independent variable. For example, suppose a researcher wanted to study the effects of a training program on sales performance. The presence or absence of a training program is the independent variable; sales performance is the dependent variable. The researcher hypothesizes that by changing the independent variable, the dependent variable will also change. Look at another example: Suppose a researcher wanted to study the effects of cadmium on rats. Cadmium is the independent variable, and the rats' health and behavior are the dependent variables. The researcher can manipulate or control the amount of cadmium given to the rats, and hypothesizes that the dependent variables will be affected to some extent as the independent variable changes. In this chapter, we will concentrate on examining variables as they relate to the research problem and hypothesis. In order to conduct a study, researchers must identify certain variables that play a significant role and define those variables so that others can understand what they mean. Defining variables often takes the form of an operational definition, which describes the research procedures and techniques for measuring the actual values of the variable. The next section discusses operational definitions in more detail. But first, we must understand why researchers start by identifying variables.

### ➤ Types of Variables

Research, by its intentional nature, examines some features of the world whilst holding other features constant. Consequently, all research involves at least two extremes: a constant that is not being studied, and a variable of some type that is being studied. Various types of variables can be recognized depending upon how they relate to the hypothesis. First, there are the independent variables that effect the relationship being investigated. These are sometimes referred to as 'explanatory' or 'stimulus' variables. In social research, they tend to include social characteristics of individuals or groups such as class, sex, age, acting on surveys or experiments conducted with people in their role as social group members. Then there are the

dependent variables that effect the relationship being investigated. The dependent variable is what is being modeled or predicted and by its nature must be more varied. There are many other distinguishable links between the variables involved in the research design and execution which help to clarify and elaborate on the nature of the relationship with respect to the research problem. Correlated variables are simply those that demonstrate some relationship during the research, but without any actual cause and effect linkage being established. An intervening variable is a more complex type of correlation. Intervening variables are related to both independent and dependent variables in such a manner as to influence how contact between the two occurs. Thus, they are seen as occurring first in the chain of contact, or effecting the strength or ability of the independent variable to impact the dependent variable. Demographics characteristic, for example, could be acting like intervening variables in an effect investigation about voting preferences consequential to candidate's partisan. The demographics would also be intervals effecting sensitivity of all social groups to the effects of candidates' parties.

#### ➤ Operationalizing Variables

If we are to gather as information the answers to the questions of our research, the variables must be approached in an operational manner. Not all results are in the intention of observing the direct relation between two variables, some responses can be focused on the symptoms that these variables represent. This may be the case of our study when it comes to measuring the emotions that a person feels when interacting with a brand through their presence on social media, or even more subtle feelings and opinions in which they share the same feeling of belonging at a certain time of a country or group –that is, social movements– appearing as followers or supporters of it. One of the main variables in our study is that of the emotions aroused by a certain brand; and we question how its construction occurs through the brand's day-to-day posts. These emotions are difficult to measure –some of them are abstract– but have been previously studied. Subsequently, we will focus on some methodologies that allow us to measure these emotions in the best possible way. The fact that a variable has an arbitrary relation with its operational indicator should not be worrying. What researchers should worry about is that the relation between the variable and its indicator makes sense from the theoretical perspective. In the theoretical definition of the variable, the theoretical system implies relations with other terms that make its meaning clear. This theoretical consideration responds to the principle of pertinence: by establishing a theoretical meaning, we indicate which facts are relevant to explain the variable and which are not. If the selected

indicators are within the relevant set, it can be assumed that they indicate the theoretical variable, and its evolution will reflect that of the variable.

#### ➤ Importance of Clear Definitions

A common concern in the social sciences is that they do not have clear definitions of the terms they use, i.e. that their definitions do not correspond to the definitions used in the natural sciences. The argument is as follows. In the natural sciences, there is agreement regarding fundamental definitions and the researchers use the same definitions while in the social sciences definitions differ from researcher to researcher and the definitions proposed are not verified with the same care as in the natural sciences. Because of this lack of clearly defined fundamental terms, other problems are created. For example, it is possible to have causal exchanges between different levels of definitions. A researcher may explain at time  $t$  something that is measured at time  $t - 1$  using different definitions that entail different types of causal forces. Or a researcher may explain at one time and one space what he wants to explain using theories that employ different theoretical terms, and would therefore map their corresponding causal mechanisms in different ways. The problems highlighted in these examples are avoided by the definitions proposed in this chapter. The crucial observation is that the path specification does not involve an additional, independent causation relation; the path specification involves activities that are more concrete in character and that reflect the nature of the causal connection postulated by the theory employed. More specifically, the path specification consists of defining more precisely the action or property that is stated in the hypothesis to be in charge of explaining the *causa facienda*. Moreover, the highlighted examples of inter-level and inter-theoretical exchange problems involve definitions that relate to links that are not path specifications. For the reasons stated above, the definitions proposed here are therefore the ones that can be used to check path specification, and to clarify interlevel and intertheoretical exchanges.

## **UNDERSTANDING CAUSALITY AND CORRELATION**

Research often investigates the relationship between two or more variables. For example, a social scientist might look at social media use and self-reported loneliness scores. In this study, social media use is referred to as the independent variable or predictor variable while loneliness is considered the dependent variable or outcome variable. A researcher might test whether an increase in social media use is related to an increase in reported feelings of loneliness. If so, can we say that social media use causes increased loneliness? If not, what

can we conclude? These questions relate to the concepts of causality and correlation. Causality is the act of causing or producing, while a cause is an agency which brings about a certain effect or result. In the social sciences, we often want to make claims about how one thing affects another. A study examining the impact of a smoking ban on air quality is determining the effectiveness of a smoking ban. A study examining the impact of temperature on ice cream sales is asking how temperature influences ice cream sales. A study examining the impact of hours studying on a test score is investigating how study time affects a test score. Claims of causation usually make some sort of time-based prediction. If variable X is truly causing variable Y, then a change in variable Y cannot happen before a change in variable X. Causation indicates that an event is the result of the action of another event. For example, the increase in prices of gold and diamonds will cause a decrease in the prices of artificial jewels. Correlation indicates a relationship between events; that events appear to be related to each other without necessarily being the result of one another. Correlation does not imply causation. In the previous example, the fact that the prices of gold and diamonds are increasing does not mean necessarily that it is the effect of fallen prices of artificial jewels.

➤ Defining Causality

The ability to understand and express the relationship of causal and non-causal events is essential to making sense of the world; it is the most important task of science. We day-to-day infer the causes of natural and social events, and we learn the likely consequences of novel acts or of actions by systems of a known class. However, the action-predicting use of the word “cause” is subject to limitations, and scientists need to think carefully when they address the issue of causation. There is no simple answer to the question “What does it mean in a scientific sense to say that A causes B?” A formal definition of causal relations in terms of structural equational models with unobserved or latent variables has been proposed, while other contributors to the scientific literature assume that a causal relation is one for which intervention produces a change. A good definition must be sufficiently general to handle actual scientists’ usage, whether or not they have these or other definitions in mind. Philosophical models of causation pivot on the matter of understanding boundaries and meaning of counterfactual situations. For example, principles assert the importance of establishing that A is a constant conjunction of events that differ on dimension B, or that the A do but not the non-A produce changes on dimensions on which B varies. However, social scientists usually concerned with regularities rather than constants, are assumed by most

social science researchers to be cause events capable of producing change of pattern on dimension B. Causation is not just another correlation.

➤ Correlation vs. Causation

How do we know that an association between X and Y is a real causal relationship, rather than mere correlation? It is common in every day life to hear someone say that one thing causes another. For example, "Drinking alcohol causes drunkenness." Or, "Getting arthritis causes your knees to ache." As tempting as it is to accept these types of statements at face value, a cautious approach is to first check whether X and Y are closely correlated with each other, and then see if the correlation passes the tests discussed here. If your tests indicate that the answer linking your two variables is "not proven," then it is best to refrain from making what could be a misleading causal interpretation.

It is possible to carve up the term "correlation" in different ways. There is the statistical correlation that is conveyed by a common correlation coefficient. That coefficient quantitatively expresses the degree of linear association between your two variables. Then again, you may have a dynamic correlation in mind, which is the correlation between time changes of the two variables. When discussing correlation, you may also mean to use "correlation" as shorthand to refer to other types of undirected association. These different aspects of correlation share one key aspect: the variables are associated. However, markup is easy and may do more harm than good in the end. A cautious reader will keep this peculiarity in mind prior to finding conclusions derived from correlation studies. If your study is about a statistically understood set of micro-economic variables, and/or if you are studying very long time series, then you may get away with discussing restrictions.

➤ Methods for Establishing Causality

It is important to note that correlation does not mean that one variable causes another, and it is equally important to test relations of causality by additional methods. Several methods of establishing causation are available, but it needs to be kept in mind that they can help you to only strengthen your evidence for the proposed cause, and cannot in any way provide you with certainty that this is it. These methods are, (a) Time-order, (b) Covariation of X and Y, (c) Nonspuriousness, (d) Experimental Designs, (e) Manipulation of Suspected Causes, (f) Control of Covariates, (g) Statistical Testing, and (h) Specificity of Relations. More will be said about this in Part II of this chapter. You will rarely, if ever, find all of these methods at

the same time. If you do, the proposed relation is almost always causative. The time-order principle simply states that the cause must precede the outcome in time, or, in simpler terms, that the cause needs to come before the effect in order for the relation to be causal. Covariation means that when X changes, Y changes too. In the case that X is a cause of Y, when the value on the X-variable increases or decreases the value on the outcome also does so or does the opposite if it is negative. The third principle called non-spuriousness states that there is no other reason why both variables, X and Y, develop a relationship. A spurious correlation exists without any correlation between the two variables, X and Y, in negative or positive direction, but because of a third variable, controlling for this third variable the correlation between X and Y vanishes.

### **Chapter summary**

Chapter 2 delves into one of the most critical phases of the research process—identifying and refining the research problem. A well-defined research problem serves as the foundation upon which the entire study is built. This chapter guides the reader through the process of recognizing a broad area of interest, narrowing it down to a specific issue, and converting it into a clearly articulated research problem. The importance of choosing a problem that is researchable, significant, and feasible is emphasized, ensuring that the study remains focused and contributes meaningfully to the field. Once the research problem is established, the next step is to formulate precise research objectives, questions, and hypotheses. Objectives outline what the researcher intends to achieve, while research questions guide the direction of inquiry. Hypotheses, on the other hand, are tentative explanations or predictions that can be tested through empirical data. The chapter explains the different types of hypotheses—null and alternative—and the importance of clarity and testability in their formulation. This process helps bridge theory and practice, laying the groundwork for data collection and analysis.

The chapter also introduces the concepts of variables and operational definitions, which are fundamental in designing any empirical study. Variables are classified as independent, dependent, control, and extraneous, depending on their role in the research. Operational definitions provide specific meanings to the variables within the context of the study, making abstract concepts measurable and testable. This clarity ensures consistency in understanding, interpretation, and application of the terms used throughout the research. Finally, the chapter addresses the concepts of causality and correlation, which are central to analyzing relationships between variables. Causality refers to a cause-effect relationship where one

variable directly influences another, while correlation indicates the degree of association between two variables without implying causation. Understanding these distinctions is vital for drawing accurate conclusions from research findings. The chapter highlights that while correlation may exist between variables, it does not confirm causation without rigorous testing and control of influencing factors.

\*\*\*\*\*

## CHAPTER 3

### REVIEW OF LITERATURE

Any academic research work is better explained, judged, and appreciated in the context of its relationship with available or existing knowledge in the discipline. All predictions or expectations need to be compared with what has already been found. Hence every research work must take into account the previous works and it is important that work is cited appropriately in the text. Recognition of past works has an importance of its own course on academic progress. In depth knowledge of previous works in the disciplines becomes the basis from where the researcher must embark on the new discovery or body of knowledge. It is academia's way of keeping a careful and recorded track of the intellectual history of discovery. Therefore, there is no question about the importance of the literature review process in ensuring the sanctity of the reported findings. A literature review is by no means a mere summary of previous studies.

Figure 3.1 Review Of Literature



A literature review surveys academic articles, books, and other sources relevant to a particular area of research. It provides a description, summary, and critical evaluation of these works. It is a process of scanning the literature of a particular field of study and synthesizing it into some key ideas that need to be investigated. It also identifies research gaps in prior studies which warrant additional exploration. There are various types of literature review within academic domains. It may vary from a more traditional type, reporting on what is

already known about a work, to ones that are narrative in nature, describing the state of the art in the area, or critical analysis, emphasizing the shortcomings of certain types of work, or positioning it within an epistemological context.

## **PURPOSE AND STRATEGIES FOR REVIEWING LITERATURE**

Research in a discipline cannot be merely a repeat of what has already been done previously. To create new knowledge, researchers have to view their own research engagement as part of a larger discourse in the topic or area of inquiry. To establish that connection, researchers must write a summary or a review of the literature. In academic research, a review of the literature must serve a different purpose from those created in undergraduate study. Some studies treat literature reviews as an introductory or background component of the research report, with little inherent value within the report beyond furnishing a fixed platform from which the practical work proceeds. Not all literature reviews are designed primarily to propel readers into a new piece of work. Some written reports have literary merit in themselves. Sweeping literary reviews, for example, synthesize a substantive body of knowledge in a discipline, making knowledge more accessible and providing a solid foundation for subsequent research endeavors. Findings of a sound review may well spur other researchers to conduct similar review projects in other disciplines. They may even stimulate the development of comparison frameworks or models that span reviewing studies across various disciplines. They need not necessarily be limited to comparing models of specific topics. Pan-disciplinary reviews may also provide incentives for researchers in other disciplines to cooperate at inter-disciplinary levels.

### **Strategies for Conducting a Literature Review**

The first step in conducting a literature review is to ask the right questions. This probably should start with the three main questions posed in any research. What not much talked about is the adjustments that are necessary to make a literature review truly unique. Thus, adjustments are necessary at all levels. The first adjustment is based on what is absent. The function of any research is to add knowledge, not just review it. If you are reviewing what is already available but not adding to that, the review should just carry the label of being a paper. Therefore, explicit questions need to be raised about what is absent in the reviews already available. The questions about what are the redundancies and inconsistencies would be the next steps. Consider all the various research scholars would have reviewed, asking the same questions, but produced different answers. What could be the reasons for that? Then there are

the differences based on cultural, social, political, and environmental factors based on the location of the research scholars. Analyses along these variables should be seen carefully. What may not have been asked in conventional location specific reviews? These are factors beset by and large headline glamorization research questions. The next is the question faced by anyone studying the country's economic development – class analysis. Are there sections of the population excluded from the focus of the inquiries? What specifically are the questions that need to be asked? And to finally produce a paper that says nothing unique, is ridden with the so-called research methodology. The three questions discussed above lead to the next critical dimension of the research paper. What is the sample used in the research? Is it conventional size, selection and representation, and how does this enhance the validity of the paper? The next additional processes should contain a critical element. What are the adjudged, available methodologies? Which particular aspect is considered innovatively, and what gets added that others haven't? The final pointer to adjustment is consistency. What are the broad statements made about the research questions, and what are the broad pointers employed to ensure it?

### **LOCATING SCHOLARLY SOURCES USING GOOGLE SCHOLAR, SCOPUS**

The most critical challenge for most practicing academic researchers is locating appropriate scholarly sources that become the foundation for the literature review in their research. There are several ways to locate scholarly work that is relevant to your research project through the general internet and through specialized databases that generally include only scholarly work. However, these databases do require a strong institutional investment for access. The more commonly used search engine on the general internet is Google. However, Google has a specialized search engine for scholarly work. This option is free for everyone to use which makes it attractive, if not a bit limited compared to the institutional databases. If you need to access more scholarly documents than what this option can produce, you might try one of the specialized institutional databases. However, these databases are not free but require some level of institutional funding for access. However, both are very good at producing useful documents for academic researchers involving many fields of scholarship and the sophistication of the search features to generate valid hits is highly sophisticated. This option is the scholar option for the massive search engine that we all know. It's basically a scholarly filtering function for the general search engine. You can locate this option by going to the specific site and regular functions can also be accessed by going to the general search engine and clicking on the first entry for its own preference setting. You can search in this option in

the same manner as you would search on general search. The function is easy to use, but the algorithm is not as precise as other more formalized databases developed over decades. However, if you can successfully search in general search, you can acquire the same skills and search functions in this option. Though it does not give scholars as many filtering options as specific databases, it does operate in a manner that is better for those less familiar with more robust database functioning.

➤ Using Google Scholar

The search engine is widely used to search for scholarly literature. It offers the benefits typically associated with public web search engines, including free access to statement and citation search options, a large index that encompasses gray literature, the option to search across multiple academic disciplines simultaneously, and a user-friendly interface. However, using an academic search engine does also have drawbacks. For example, users may be overwhelmed by the sheer number of sources available by exclusive reliance on it, especially if the search is not refined to a particular area of research. It is sensitive only to exact-phrase searches; “matching-all-words” searches will retrieve every document containing the keyword terms anywhere in the text. As a result, the researcher may be inundated with irrelevant citations. Irrelevance can also become an issue when keyword terms are word forms that appear in an unusually high number of distinct titles, keywords, or abstracts. There is also the problem common to any search engine employing algorithms to rank citations according to relevance: that the order of returned citations is not indicative of real applicability. Given these issues, users should be cautious to cross-check relevant literature with search engines employing alternative querying approaches, such as subject-based indexing. Therefore, it is best viewed as a supplement to other academic research search engines. Despite these limitations, it offers indispensable features describing the status of the sources available through its search engine. Each source is linked with a maximum of five highlighted options beneath its description. The first two options indicate how to access the sources available through specific web pages, while the remaining three provide contextual research-related information.

➤ Utilizing Scopus

Alternatively, research content may also be searched at Scopus. Scopus is an abstract and citation database of peer-reviewed literature in the fields of science, technology, medicine, social sciences, and more. In addition, Scopus has content coverage from a wide range of

subjects, provides different types of documents, has better quality control procedures, and has a robust metric. Scopus can also generate a list of documents available in the database that contain specific keywords. Search in Scopus is free for users from academic institutions that are subscribing to Scopus. Access guidelines and subscription information may be found on the Scopus homepage. In Scopus, searches may be done on keywords in documents' titles, abstracts, and keywords; the authors' names; a functional affiliation of authors; years of publication; the types of documents; the source title; volume number, issue number, and page numbers; language and status of documents; journal publisher; and DOI number of documents. Keywords in the title of documents are the most reliable indexes of documents, as they are used directly by authors to describe their works concisely. Keywords in the title may return a higher-quality list of documents than using those in abstracts that are usually composed of multiple sentences. However, the search program does not compare keywords in a phrase form using quotation marks. So, authors should use the AND condition of keywords in title searches on words that must appear in documents. The OR condition may also be used for mutual exclusions of document keywords; however, this renders the return list of documents lengthy and fuzzy, and it may return irrelevant documents as well.

## **ORGANIZING WITH REFERENCE MANAGERS SUCH AS MENDELEY OR ZOTERO**

For literature reviews in academic research, sources must be well-organized to facilitate writing and integrate text cohesively. One method is to use keywords or text notes physically within document files stored in directories. Alternatively, a more direct and convenient solution is to use a reference manager. These tools allow users to create folders and subfolders to categorize paper pdf documents according to their review themes. Such notes and keywords can be managed within the reference manager for when papers are stored in folders for various ideas they cover. These add-ons also aid in creating bibliographies for papers and proposals. These reference managers simplify the task of properly co-citing multiple papers in-text. The following sections provide short descriptions of both reference managers to help researchers decide the appropriate reference manager for them. One of the first successful reference managers and academic social networks captures citation details from research databases automatically. The general downloading process involves creating an account, downloading the software, installing a browser add-on, and using the add-on to pull citations from research databases, after which pdf versions of the papers can be downloaded into the reference manager using another add-on. Papers can be stored in different folders

created for different topics or ideas. Keywords and comments can then be added to facilitate searching for literature. It also generates bi-weekly activities for people in your network and articles you did not find otherwise. Papers can be tagged under various categories and oftentimes, new keywords will be prompted, which explains techniques and areas termed in the paper. You can also see related articles, who have cited this work, self-citations, and co-citations. The plugin also generates ready-to-use citations with database links as shown for various formats.

### Which reference manager should I use for organizing research?

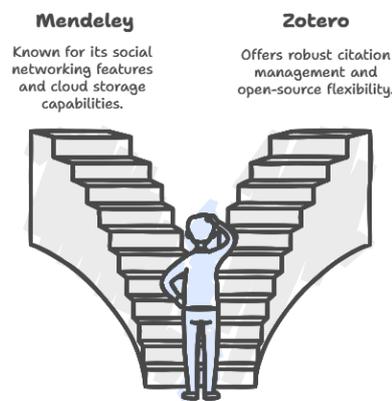


Figure 3.2 Organizing With Reference Managers

#### ➤ Overview of Mendeley

Mendeley is a reference manager, popular among researchers and students, that combines a desktop application and an online service for managing your literature library and PDFs. You can create your Mendeley database by importing references from your own library or public databases or manually entering references. Using the Mendeley desktop client, you can classify the documents in your library by keywords or author, mark the papers you want to read, and save notes and annotate the documents. The Mendeley desktop also allows you to create groups of researchers and share references and documents with them. Mendeley integrates well with Microsoft Word, Libre Office, and LaTeX users, helping you create citations and bibliographies in various styles. You can also integrate Mendeley with Wordperfect and startcite. Mendeley is available for different operating systems: Windows, Linux, and Mac OS X. Mendeley is free to use, but depending on the requirements, you can opt for the paid version of Mendeley if you require additional storage space. The free version

of Mendeley includes 2GB of web storage, allowing you to manage up to 100 references in your Mendeley library. You can create a private group and also create two groups of up to three users to collaborate on projects. You can opt for the different optional Mendeley Pro subscription plans based on your requirements. You can get the Pro version to unlock unlimited storage in your web library and unlimited private groups. If you have a specific requirement, then you can also watch the two-step pricing tailored for your requirements.

➤ Overview of Zotero

Zotero is a free, open-source reference manager designed to help researchers organize, cite, and share their research materials. Specific to the two approaches commonly used to manage references, Zotero supports a local-based solution as well as an online solution. To implement the local-based solution, researchers can download the Zotero standalone app. Further, it can also be used as a browser-based reference manager, where researchers can save and work with references by creating an account. This way, researchers can use Zotero on any device, be it Windows, macOS, iPadOS, Android, Chrome OS, or Linux. While Mendeley is primarily an academic social network, Zotero focuses on the adding and access features. The user interface is easier, more straightforward, and more intuitive. Instead of organizing references using keyword tags, Zotero uses a hierarchy of folders and subfolders. Unlike Mendeley, Zotero does not provide citation software directly within the word processor. For researchers looking for a standalone citation program, Zotero is a better choice. Even though Zotero was previously limited to referencing only the most common types of sources, they now have a massive collection of citation styles, much larger than Mendeley's. Furthermore, users can also design custom citation styles as needed. Like all free reference managers, Zotero also comes with a specific storage limit size. For those looking for more storage, Zotero provides paid plans too. This makes Zotero an ideal choice, especially for those in the humanities and social sciences fields.

## **WRITING A STRUCTURED REVIEW FOR A THESIS OR ARTICLE**

In the previous chapters we have discussed various roles of a literature review, and have also explored the structure of a literature review. Now we focus on the more practical dimensions of reviewing literature. The first point we would like to discuss is the advantages and challenges of writing a structured review. Both types of reviews have their respective merits. A traditional narrative review is often the most succinct way to convey what has been done before and what are the gaps. Some reviewers may argue, with merit, that structure in a

review could result in unnecessary complexity. In our own experience in reading, and writing both types of reviews, a structured assignment gives more clarity, is easier to follow, while at the same time helping the author avoid the usual pitfalls associated with writing a literature review. Apart from the inherent advantages a structured approach to writing has, there are also numerous pitfalls associated with writing a literature review. Different styles of literature reviews could make it difficult to write one in a precise, focused format. Careless introductions that do not explain the rationale for writing the review are all too common. Deficient reporting systems or guidelines, or the mixing of different methods and resultant problems with comparability and synthesis, can make it difficult to write a coherent results and discussion section. The results sections are usually patches of summarizing different studies, with little thought to a deeper analysis of the implications of the work, and often note the merits and limitations of the literature being reviewed.

In conclusion, taking a structured approach to writing is fraught with challenges but at the same time has clear advantages. In the next section we will describe the components of a structured literature review and the common pitfalls associated with writing. Hopefully this would make our readers enjoy writing literature reviews in the time to come.

#### ➤ Components of a Structured Review

Writing a literature review is rarely a simple matter of gathering papers together. Rather, it is an exercise in analysis, synthesis and evaluation, and the review article can vary greatly in style and approach. Nevertheless, a structured review contains many common elements. Generally, an abstract summarises the review's content, context and significance, and warrants further reading. There is usually a short introduction that sets the review into context, describes the aims of the work and gives a short description of the reviewed topic. This is followed by a description of the methods used to explore the literature; the operations that a reviewer undertakes to find and manage references. Clearly describing how you found the literature is very important; good reviewers easily wield a body of literature, but show little to no information about how they found it. This is particularly important in identifying the contributions of the works. The next sections contain the content of the review — the previously published works. These sections are shaped by the annotations — the notes that a reviewer writes on top of the papers — and the review should be carefully structured. A concise discussion follows, placing emphasis on conclusions and recommendations arising from the review, and any topics that were not covered that may be the subject of further

reviews. Finally, a list of works that were cited in the review is provided, always ensuring it is complete and comprehensive. Some journals will request specific formats for your review, requiring for example sections outlining content keywords, summaries of the presentation, or expected new contributions. These journal considerations must be adhered to, but even when none are requested, following what are conventions of other types of academic structure, or when they contain specific elements for reviews, is a good idea. This is because academic writing will usually have an intention: to guide its reader(s) through the review.

### ➤ Common Pitfalls in Writing

In research, writing the literature review is often perceived as a task done only to fulfill a requirement. Many researchers, thus, look for quick ways of preparing the literature review as they do not see it contributing to their research as much as the research itself. This is detrimental because not only is the process of preparing the literature review enriching but also a good literature review adds significant weight to the research by placing it in the right context. Moreover, it is also one of the earliest aspects of research that is exposed to the external audience, and hence should be written clearly and methodically. Writing a literature review does not entail going through all existing literature and summarizing its findings. If this was the case, we could easily duplicate research findings by merely reading the papers! A literature review requires a careful selection of literature, classification of the literature into related strands, and a synthesis of key findings, contributions, limitations, and techniques. Writing a literature review is not the same as writing an essay. In an essay, all one has to do is simply put forward arguments for or against the subject. In a literature review, connections between various research findings and arguments have to be established in addition to assessing how past work has built up over time so that the research question of the review can be answered. This is what makes writing a literature review methodologically different. Furthermore, using a structure can help in this task – it allows the researcher to focus on specific aspects of different sources of past work, which are in turn linked up through the structure chosen. When we transgress from the structure, we then focus on other aspects and thus are in a good position to connect together points from other parts of past research because of the variety of perspective.

### **Chapter summary**

The primary purpose of a literature review is to understand what has already been studied, identify gaps in existing knowledge, avoid duplication, and provide a theoretical basis for the

current research. By critically examining prior work, researchers can position their study within the broader academic discourse and justify the relevance of their research problem. The chapter outlines key strategies for conducting a literature review, including starting with broad searches and gradually narrowing the focus to specific themes related to the research topic. The chapter also introduces essential tools for locating reliable and scholarly sources. Researchers are guided on how to use academic databases such as Google Scholar and Scopus effectively. These platforms offer access to peer-reviewed journal articles, books, conference papers, and other academic materials. The chapter advises using relevant keywords, Boolean operators, and filters to refine search results, ensuring that the sources are credible and up-to-date. Mastery of these tools is vital for accessing the most relevant literature in an efficient manner.

To manage the large volume of references gathered during the review, the chapter recommends the use of reference management tools such as Mendeley and Zotero. These tools help in storing, organizing, annotating, and citing sources easily. With features like citation plugins for word processors, they streamline the referencing process and ensure consistency in citation styles like APA, MLA, or Chicago. By using such tools, researchers can focus more on analyzing content rather than manually managing bibliographic details. Finally, the chapter outlines the structure and format of writing a comprehensive literature review. A well-structured review should include an introduction, thematic or chronological categorization of literature, critical evaluation, and a conclusion that highlights the research gap. Researchers are encouraged to synthesize findings, compare and contrast perspectives, and critically appraise the strengths and limitations of previous studies. This organized approach not only demonstrates a deep understanding of the subject but also lays a solid groundwork for the rest of the research study.

\*\*\*\*\*

## CHAPTER 4

### RESEARCH DESIGN AND SAMPLING TECHNIQUES

The design of a research study is crucial to its successful conduct. A carefully drawn up research design is imperative for the successful outcome of the validation of the research hypothesis. This is emphatically true for medical research where any lapse in the study design can lead to dire consequences because of the direct effects on patients in the research trial. The strength of the study's conclusions is directly related to the design and methodology that have been used. Strikingly, both poor design and innovative design can yield equally interesting and entertaining publications. A bad design is one that fails to answer the questions posed. However, a good design is the one that does not leave bias at the end of the analysis. A research design is like a map indicating the important landmarks and the route to be taken to reach the destination, that is, answering the research question. More importantly, in order to eliminate any combination of chance, bias, or confounding from invalidating the results of the study, the research design should include the sampling methodology.

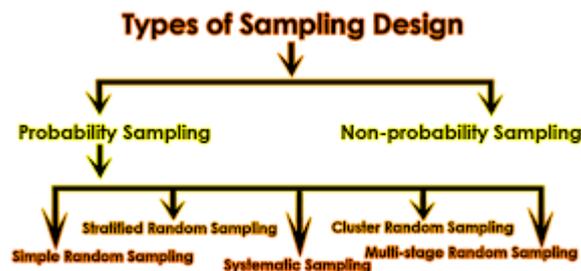


Figure 4.1 Types of sampling design

### RESEARCH DESIGN TYPES: CROSS-SECTIONAL, LONGITUDINAL, EXPERIMENTAL, QUASI-EXPERIMENTAL

The design of a given study typically is the most critical aspect of the research process. It is the framework or layout, the road map by which the research is undertaken. Simply stated, research design reflects the overall strategy or plan that the researcher adopts in order to be able to answer the research question and fulfil the research objectives and purpose. The research design enables the researcher to address the what, why, where, when, who, and how things in relation to the research problem under study. Basically, a research design is a plan of how you are going to conduct the research; it defines the basic structure of the study and

sets out the overall procedural approach for conducting the research. The research design may be defined as the plan that specifies the methods and procedures for collecting and analyzing the needed information. It is the decision-making process in which the researchers subject their research to a certain amount of control, in order to specify how, when, where, and from whom they will collect the data required to answer the research questions. The research design should at least incorporate the following decisions: the type of data needed, the type of sampling strategy to adopt, the procedures for collecting the data, the visit schedule for field work, the type of contact method to be used, and the methods for analyzing the data. This chapter discusses briefly the different types of research design in relation to the unique requirements of survey research. There are several types of research designs and for the purpose of clarity, we primarily focus on cross-sectional design, longitudinal design, experimental design, and quasi-experimental design.

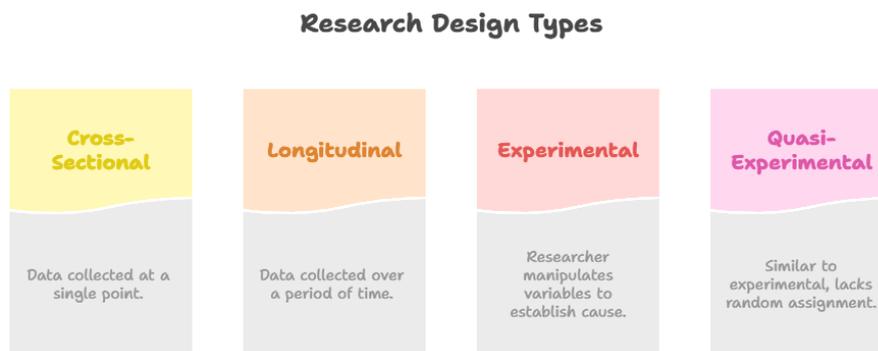


Figure 4.2 research design

### ➤ Cross-Sectional Design

Broadly speaking, there are two types of research designs: descriptive and experimental. The essential character of both types of designs is that they differ in the extent to which they can provide evidence for a causal explanation of the association under investigation. A descriptive design is needed to provide an initial assessment of the suggested association, but an experimental design must be used to establish the causal nature of the association, conceptualized at the outset as a hypothesis. In this brief overview, we will discuss four types of descriptive research designs: cross-sectional; longitudinal; experimental; and quasi-experimental. A commonly used observational study design is the cross-sectional study, referred to as a prevalence study. Cross-sectional studies are characterized by an examination

of the association between the exposure and the outcome of interest at the same time point. That is, the non-exposed and exposed groups are identified simultaneously, and information is collected on the presence or absence of the outcome of interest at the same time or time period. The presence or absence of exposure and disease for the subjects is ascertained at the same point in time or for the same time period. We provide a brief description of the strengths and limitations of cross-sectional studies compared to other designs. A cross-sectional study is considered a simple, fast, and economical approach to studying an association, using a short-term timescale. Both the outcome and the exposure are assessed simultaneously, without concern over temporal sequencing. Compared to cohort studies, which follow study subjects over time to determine incident disease by comparing those with and without exposure, cross-sectional studies provide ‘snapshots’ of disease prevalence in a population at a specific time. However, while cohort studies track an exposure or exposures over time, cross-sectional studies do not, and instead rely on the memory of study participants for prior exposure and disease information.

➤ Longitudinal Design

Longitudinal designs involve the collection of data at multiple time points. With the same subjects providing data over time, longitudinal designs can measure change and advance our understanding of dynamic processes. The increased risk of attrition compared to cross-sectional designs leads to a number of tradeoffs around sample selection, statistical analysis, and data collection method. Because they are suitable for studying how influences change over time, they represent a stronger test of causality than cross-sectional designs. Longitudinal studies are popular in social and behavioral sciences and other theoretical areas where researchers expect influences to change or become more or less salient over time. There are a variety of possible longitudinal designs, including time-series designs, where several measures of the same variable(s) are taken on the same subject over time; forecasting designs; rotated panel designs; and panel designs with or without prior conditioning. In longitudinal research the same subjects are repeatedly measured over time. This measurement schedule extends the time frame of the study while it collects predictor and outcome measures at several time points. Longitudinal studies often collect data on predictor and outcome variables at least 6 months apart, so they are less frequent than cross-sectional studies. Instead of being limited to data collection methods that can be used to assess predictors across a range of time points, longitudinal studies can utilize methods that cannot be repeated because of excessive subject burden. Lengthy interviews or focus groups can

probe specific issues of particular interest, and observation can be used to validate the reported data and to collect more sensitive variables such as peer support and conflict.

### ➤ Experimental Design

Experimental research is one of the three main types of research design. These three main types of research design are experimental, correlational, and descriptive. Experimental research is considered the gold standard because it has rigorous structure and process to rule out alternative assumptions, and increase confidence that change in the dependent variable is due to change in the independent variable. To better understand this simple explanation, let's take a look at the parts of an experimental design. This may help you understand the importance of having a rigorous structure and process in an experimental design. In the simplest case, experimental research has only two groups, called the experimental and control groups. When a new treatment is tested, the experimental group receives it and the control group does not receive it. Or, more accurately, the control group may receive a placebo treatment, which is designed to appear identical to the treatment being tested as possible, but is inactive. Unlike in correlational and descriptive designs, experiments have at least one independent variable and its levels are created by the researcher. All other possibly influencing independent variables, called extraneous variables, are held constant, or controlled. Experimental group assignment is done through random assignment, which is an essential component that differentiates experimental design from other types of designs. Once participants are assigned to the experimental or control group, the researcher can be more confident that the groups are comparable and have similar characteristics. Random assignment balances measured and unmeasured characteristics between the experimental and control groups.

### ➤ Quasi-Experimental Design

Quasi-experimental designs share some of the properties of true experiments, such as the manipulation of the independent variable or the treatment and the measurement of the dependent variable or the outcome, but, unlike true experiments, they lack randomization. Specifically, there is no random assignment of patients to comparison and treatment groups. Quasi-experiments allow researchers to conduct studies in natural settings, where it would be impractical to randomly assign subjects to treatment groups. Thus, many intervention studies and program evaluations are conducted in the manner of quasi-experiments. A major shortcoming of quasi-experimental designs is that their internal validity is very low. There are

many varieties of quasi-experimental designs. The classic two-group design without random assignment has many names, including the pre-post design, the event-history design, and time-dummy design. It involves the treatment of a single population, pre-post measurement of change, inference only about that specific population's relative treatment effect, threat from other secular trends, and usually threat from measurement and regression artifacts. The time-series design, which is often confused with the classic pre-post design, is more complex. Generally, there are five or more pre-post outcome measures in both the treatment and control groups, providing both temporal and cross-sectional data. Because of the many comparisons that can be made, said design is capable of tremendous statistical penetration. It has a more elaborate notation than the classic two-group design but is still close, with the contrast between Y and C being expanded to account for multiple observations. Although there are many studies that are referred to as experiments and that employ some type of quasi-experimental design, few researchers can agree on what criteria constitute a convincing quasi-experiment.

### **SAMPLING METHODS: SIMPLE RANDOM, STRATIFIED, CLUSTER, PURPOSIVE, SNOWBALL**

Most published research may be assumed to be a sample which the researchers have drawn from a larger and more important universe. The conclusions reached by the researchers are based on the information provided by the sample and while they may answer specific questions effectively and economically, the risk that they fail to provide an accurate account of what is going on in the wider universe must be acknowledged. Sampling is a natural solution for information-gathering problems when budgets limit the resources available, yet the numbers of addressable or available units in the universe are large. It is no coincidence that many of the examples given in this chapter concern market research which, by its nature, addresses large populations. When researchers use a sample, it is important to ensure that it is representative of the universe of interest. One of the goals of the research should be to understand the universe fully enough to make decisions about stratification as required. For example, surveys of consumer expenditure patterns take up to 12 months and use around 100 local interviewers, trained to follow a standard protocol, to collect the information from a sample of around 7500 households. A great deal of effort goes into ensuring that the sample is the best that can be achieved. The households or "units" are initially selected using a two-stage stratified sampling procedure, in which primary sampling units, known as segments, are selected with probability proportional to size, then second-stage sampling units, known as

housing units, are selected systematically within the selected segments. Households are selected systematically within the selected housing units.

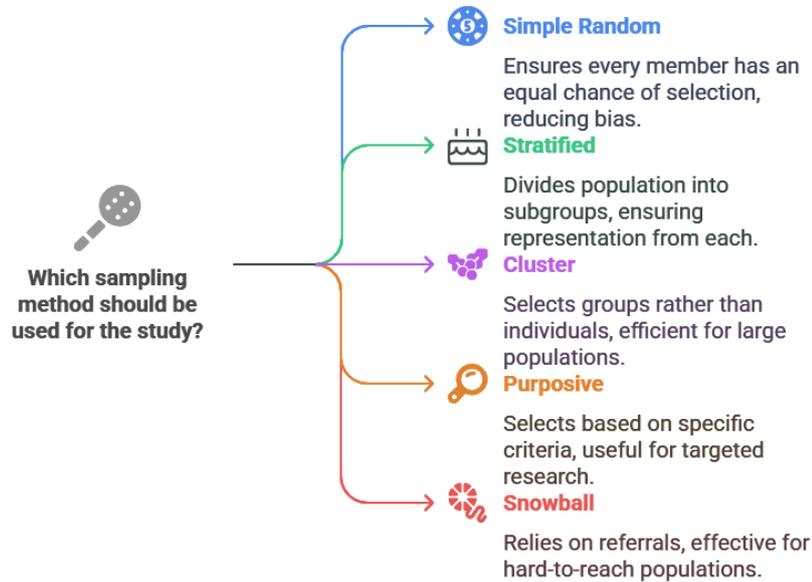


Figure 4.3 Sampling Methods

### Simple Random Sampling

In a random sample, every individual has an equal, non-zero chance of being selected. When we select researchers by simple random sampling, selection is accomplished in the following way: We list all members of the population. For our purpose here, a simple random sample is then obtained using a random number table. The selection of research participants can be replicated by other researchers who have access to the same population list at the same time. This exact replication is possible because listing and selection use explicit procedures: A complete listing of the population is required before sampling, and a neutral mechanism—namely the random number table—is used to obtain the sample. We use random numbers because researchers are human, and being human means we tend to introduce bias into things we select. Random numbers are not human, and this is good for us because it takes bias out of sampling. If we were to select research participants based solely on our judgment, we might be tempted to pick people who confirm our expectations or who just happen to be accessible to us. When we take shortcuts that offer convenience, we do not obtain a representative sample and lose the basis for making any inferences. This means that the data collected will likely contain biases that skew results, and our conclusions will be based on these unreliable findings. Simple random sampling attempts to minimize biases such as these.

## **Stratified Sampling**

Grouping the population into strata such as sex, age, education, or religion, and selecting samples from each stratum is referred to as stratified sampling. Stratification of the population is not only logical but it also usually reduces the variance of the analysis. Instead of treating our sample as a small version of the general population, thereby injecting diversity at the sample stage, we select elements by making sure each category is represented in the sample. A stratified sample might have more precision for a given total sample size than a simple random sample because the variance of estimates using stratified samples is generally less than the variance of estimates from simple random samples. Stratified sampling tends to be especially effective when there are large differences between strata relative to variation among elements in any one stratum. In stratified sampling, the variables used for stratification are not the variables being estimated; the estimates are for the variables not used for stratification. The variables used for stratification are sampled so that they are sufficient for representing the way in which samples are to be stratified. Sampling designs often use two or more heterogeneity variables for stratification. Sampling designs do not have to be probabilistic. Sampling designs for stratification can be non-probabilistic. Multiple-frame samples are stratified samples composed of the union of one or more or all of the frames from two or more sampling frames. In addition to controlling the precision of estimates, stratified sampling also can be used to control budgets for exploratory or pilot studies. Stratification is useful, as well, both practically and theoretically, in understanding how a population is constituted. Stratification can add to analysis but quantitative descriptions of households, individuals, and neighborhoods would be useful regardless of estimates. Stratified sampling provides a way to contribute to the evolving knowledge base.

## **Cluster Sampling**

Cluster sampling can save money and time, two extremely important factors in research. Thus, in small studies, individuals are randomly selected in a more traditional manner, but in larger projects, geographical areas (clusters) are randomly selected. All individuals in the selected clusters would then be included in the study. This method is effective when the costs of sampling in bulk are cheaper than doing the whole area. Much of cluster sampling derives from the work to improve the efficiency of survey sampling using spatially restricted designs, which tend to have a better non-response bias. Some of these designs also allow the budget optimization of the travel allocation. Geographical stratification is often complemented by a

cluster sampling of the clusters to deal with the rising budget. If the survey is geographically stratified, these are often the strata as a clustering of natural geographic units. It is also possible for practical reasons to partition the entire stratified sampling frame into non-overlapping clusters, especially in large industries and business surveys. A second constraint is to select the samples in the strata in a second phase. This allows the hierarchization of the different clusters by increasing any external auxiliary information. Since the pioneering work, the well-known model based on the direct estimator has allowed many developments in assessment. The specialization has described in detail the methods of estimation and variance, as well as the approximation and the calibration near establishment. This model also allows some applications in business surveys. These multi-wave surveys, whose aim is to maximize the reductions and the efficiency of the quality levels, are based on geo-clustering of the delimitation.

### **Purposive Sampling**

Purposive sampling is the sampling technique where the sample is collected from a population based on the knowledge of a population, its elements, and the purpose of the study. In purposive sampling, an individual or a group of individuals is selected to become a sample based on the purpose of research. In other words, researchers use this sampling technique to select a certain group of people in a population to study. With this sampling method, specific individuals are selected because they have specific characteristics necessary for the research; it is more qualitative and is often used in the field of qualitative research. Instead of random sampling, purposive sampling is used more commonly in qualitative research studies. Furthermore, it is also likely that the population available for research has already been involved in previous studies; therefore, these people can be selected based on their experience in a certain topic. Choosing a purposive sample is an important aspect of qualitative research; therefore, it is crucial that the right people are chosen to be included in the sample. The right people for the purposive sample must have experienced the phenomenon being studied and are very knowledgeable regarding it and can contribute useful thoughts or perceptions. Additionally, a small sample size is appropriate for this sampling technique. This technique has four steps; the first step is to identify the parameters of the population to be studied; the second step is to identify what group of investigation participants will contribute the most to the research; the third step is to define the strategy to recruit participants for the sample; and the final step is to recruit subjects. After recruiting the subjects, the data collection process will take place.

## **Snowball Sampling**

The procedure in snowball sampling (also known as chain referral sampling) begins with one or a few initial units (seeds) that belong to the target population of the research. After selecting the seeds, the researcher creates the sample by using these persons to locate the other members of the population. To locate members of the population, the researcher asks the seeds for the names of other people who meet the criteria specified for membership in the target population and who are willing to participate in the study. The researcher adds these names to the sample and uses them to locate still others. This procedure usually continues until the desired sample size is reached and/or the researcher determines that the sample is saturated or that the cost is greater than the use of the sample. Snowball sampling is often used to locate respondents in a hidden population – that is, a population in which the availability of respondents is severely restricted due to difficulties such as extreme physically visible sex traits, the stigma attached to homosexuality in certain cultures, or a lack of visible symbols of ethnicity. To illustrate these principles, suppose that a researcher is interested in studying the alcohol consumption patterns of gay men residing in urban areas. The proportion of gay men in the total male population is very small, reflecting the high degree of dissimulation associated with this sexual lifestyle. It is unlikely that a probability sample can be drawn of this population because currently available sampling frames contain no information about the sexual preferences of persons listed. Under such circumstances, a gay lifestyle cannot readily be identified from public sources of demographic data. Therefore, some social scientists resort to nonprobability sampling techniques to gain insight and expertise into certain classes of individuals, with the understanding that the goal of the research is to improve knowledge on certain demographic characteristics.

## **DETERMINING APPROPRIATE SAMPLE SIZE FOR STUDIES USING FORMULAS OR SOFTWARE**

When a researcher wants to obtain information about a population, the complete enumeration of all members in the population may not be feasible or practical. Therefore, a sample can be used to represent the population. The sample must be adequate to represent the population satisfactorily. To do this, the sample must be of a size that enables the researcher to make inferences that are more or less accurate. Many researchers, particularly in the field of social sciences, take sample sizes that are not scientifically derived. Determining sample size has been one of the most conceptualized procedures in the entire sampling process. The desired

degree of accuracy, the size of the population, the resources available for the study, and the research methodology have been recommended as the major criteria for the determination of sample size. Determination of sample size is particularly critical in quantitative research. Some qualitative researchers have argued that sample size in such research is a relatively unimportant factor. Sample sizes in qualitative research are generally smaller than quantitative research. Most qualitative researchers have recommendations for determining a sample size. The issue typically relates to when the sample size is large enough to satisfy the purpose of the study. The basic principle in determining sample size in qualitative research is saturation. This means that the researcher collects data until there is a lack of new information on the topic of interest. In this chapter, a few formulas and software programs that are substantially available to solve the sample size dilemma for the researcher based on suggestions made by many scholars are discussed. By reviewing the available tools and techniques, researchers can make informed choices because the methods selected will affect the accuracy and validity of study results.

#### ➤ Using Statistical Formulas

Using statistical formulas can be a relatively simple way to determine sample size, as long as the research team knows what variables to use to create the calculation. The most basic categories are the population variance, the margin of error, and the level of confidence. The formulas introduced below are for the easiest and most commonly used type of sample size calculation, which is for estimating a population proportion. In the following calculations,  $Z(\alpha/2)$  represents the value in the normal distribution with an area equal to the specified confidence level. For example, with a confidence level of 95%,  $Z(\alpha/2)$  is equal to 1.96. The simplest formula can be used when both  $p$  and  $(1-p)$  are known:  $n = (Z(\alpha/2)(\sqrt{p(1-p)}))^2 / E^2$ . This provides the sample size needed to estimate a population proportion with a specified margin of error. This formula can be used in two scenarios: when the sample size calculation is done before data collection, and when the sample size calculation is performed after the data collection is complete. For a given confidence level, a bounded interval for the estimate will be the true population proportion  $\pm E$ . The calculations will also include the finite population correction factor, which reduces the needed sample size when drawing from a small population. With this correction factor applied, the sample size to be computed will be  $n = n_0 / [(1 + n_0/N)]$ , where  $N$  is the population size. The needed sample size with a finite population is the larger of  $n_0$  and  $n$ . Note that the population size must be taken into account in the sample size calculation when  $N$  is smaller than about 10,000.

### ➤ Using Software Tools

Most researchers do not manually calculate the sample size using formulas. In most cases, they use software to perform the actual sample size calculations. These software programs may be standalone programs, or they may be modules available within a more general statistical package program. This section presents some of the more common statistical software programs. SAS is the most commonly used statistical software package in most research areas, particularly among pharmaceutical and medical researchers. In SAS, sample size calculation procedures are only available within specific modules. SAS macro code is available for performing calculations of many of the more commonly used sample size calculations for different research areas. STATA provides sample size analysis within the main program. Like other software, STATA does not require macro code as SAS does. Such calculations are available for the most commonly used sample size calculations for many different research design areas. However, the sample size functions do not provide full flexibility and options of the specific packages. PASS is the most comprehensive software for performing general sample size calculations; it is the most widely used package among researchers in different areas, including social sciences, marketing, and biomedical and pharmaceutical researchers. It is a very versatile and flexible package that can perform many specialized sample size calculations. It has a graphical interface where the users can select the options and input parameters required to perform the specified calculations. PASS is self-contained and very user-friendly, although not specifically designed for any particular area.

### **Chapter summary**

This chapter delves into the crucial aspects of research design and sampling techniques, which form the backbone of any empirical study. It begins by introducing different types of research designs—cross-sectional, longitudinal, experimental, and quasi-experimental—each suited for specific research objectives. Cross-sectional designs capture a snapshot of data at one point in time, while longitudinal studies track changes over time. Experimental designs involve manipulating variables under controlled conditions to determine cause-and-effect relationships, whereas quasi-experimental designs attempt similar analysis but lack random assignment. The choice of design significantly affects the validity, reliability, and generalizability of the research findings. The chapter then discusses various sampling methods used to select participants or data sources. It explains probability sampling techniques such as simple random, stratified, and cluster sampling, which ensure every unit in

the population has a known chance of being selected. These methods are essential for statistical inference and minimizing selection bias. On the other hand, non-probability methods like purposive and snowball sampling are useful in exploratory research, especially when dealing with hard-to-reach or specialized populations. The chapter emphasizes the importance of choosing a sampling strategy that aligns with the research design and objectives.

A key component addressed is the determination of an appropriate sample size. The chapter explains how the accuracy and reliability of a study depend heavily on selecting a sample size that balances feasibility with statistical power. Researchers are introduced to common formulas used for calculating sample size based on population parameters, confidence levels, and margin of error. Additionally, the use of statistical software such as G\*Power and online calculators is highlighted, offering practical guidance for estimating the required sample size for various types of studies, including those involving complex sampling structures. In conclusion, the chapter underscores the interconnection between research design and sampling techniques in producing meaningful and valid results. It encourages researchers to make informed choices by considering the nature of their research questions, available resources, and ethical considerations. A well-chosen design and sampling method not only enhance the scientific rigor of a study but also increase its applicability and impact. This chapter provides a comprehensive foundation for researchers to plan and implement methodologically sound investigations.

\*\*\*\*\*

## CHAPTER 5

### DATA COLLECTION METHODS

In research and analysis, data collection refers to a process of gathering and measuring information on targeted variables in an established and systematic way that enables one to answer relevant questions and evaluate outcomes. Such data collection enables researchers to obtain a clearer picture of something of interest. A research design is the framework or blueprint for a study, used as a guide in collecting and analyzing the data. Data collection can, thereby, be considered as a systematic approach to answering specific research questions. Data collection is an integral part of the larger research design and process. Collecting data can be done through different ways or methods. It is the first step in any research and assessment project. It allows you to define your goals, identify potential issues and ultimately explore solutions. Data collection is a process of acquiring information from a variety of sources and is the foundation of research in clinical laboratory. After the research question has been posed, the next stage is to determine the method of data collection. Data collection can be made in various ways. Some of these include investigation, observation, sampling, interview, and questionnaire. Each of these methods has its own strengths and weaknesses, and which is the most appropriate to use will, therefore, depend on the individual research project and question. Which data collection method(s) is/are used also depends on what type of study design is being carried out. For example, questionnaire methods may not be the most appropriate for qualitative studies, and measurement methods for randomized controlled trials.



Figure 5.1 Data collection

### PRIMARY VS SECONDARY DATA

The data that researchers use for analyses and to draw conclusions can either come from their own experiments, or be drawn from other sources. The use of data sources and the methodologies evolve naturally from the type and purpose of the study. There are two forms

of data that a researcher can choose from, and they differ on basis of the method of data collection. The two data types can be named as primary data and secondary data. Primary data refers to information that is collected firsthand and are specific to the study, such as in opinion surveys. The researcher creates the data, or 'constructs' them through experiments. Secondary data refers to the available information, and are explainable and quantifiable models that have been constructed by others. Secondary data can be obtained from periodicals, such as articles in newspapers and journals, books, etc. The data collection methods employed in business research vary according to whether primary data or secondary data is being used. Business research often requires primary data because existing data may not be adequate. However, many marketing research services hold extensive databases of secondary data and can make them available overnight. If the data needs to be collected or reported immediately, then a secondary data source has to be used. In some cases, such as annual sales reports, primary data can take years to complete. Primary Data Collection is most commonly associated with psychographics and opinion polling, which in turn involve the conducting of both qualitative and quantitative surveys. They however carry a larger cost envelope, and imply a veritable labor intensive exercise to reach the specified sample set. A significant percentage of marketing research activities make use of secondary data, either alone or in conjunction with investigative fieldwork.

## **TOOLS FOR SURVEYS, INTERVIEWS, AND OBSERVATIONS**

### **Quantitative Data Collection Methods**

Hypothesis testing, usually to prove or disprove a theory, is often the main method of using quantitative data in research. The primary goal of hypothesis testing is to detect true effects and rule out random sampling noise, and increasingly researchers argue that researchers should aim to collect large-scale data that enable them to make an unbiased inference – both sampling without replace and adjusting p values – aim for quality control, and not rely solely on statistical significance testing to make strong claims. The popular data collection techniques for quantitative methods include surveys and experiments. There are several advantages of a survey over doing an experiment. First, a survey allows data to be collected on many cases in a relatively short time span. This is especially important for studying rare groups. Second, it allows us to collect data on the current status of individuals. Third, survey methods can allow for a greater realism and can thus be more generalizable than experiments. However, surveys are not without their issues; conducting a survey on a web-based

environment means that not everyone in the population of interest will have access to the internet and thus a survey and it is difficult to establish timing and any temporal sequence. The quality of surveys is also affected by the sampling process, the structure of the questionnaire, the mode of data collection, and any response bias that arises, with many of these variables creating problems for constructing generalizable survey results. Experiments are another popular quantitative method of calculating a research question's answer. Experiments involve usually a randomized control trial design or something similar to it and test subjects – usually referred to as samples in experimental work.

#### ➤ Surveys

Surveys are a popular method of data collection among researchers for a variety of reasons. With the use of standardized questions, they can gather data from a variety of respondents, sometimes covering a quite large geographic area. Surveys can deliver participant responses relatively quickly, often in pre-determined time frames. Closed or fixed choice questions can be used to collect answers that are easy to convert to usable data for statistical analysis. Surveys can also be inexpensive, especially if the researcher incorporates electronic methods. However, surveys also have several limitations that should be considered. Questions must be clear and concise with well-defined response categories. This can sometimes be a difficult task when a researcher also hopes to gain in-depth insight into the study topic. Furthermore, surveys relying on a specific type of response option may lead to undesirable data. The questionnaire is also usually sent to a select group within a given population. If the entire population doesn't receive a questionnaire, the researcher is left unsure if the populated sample is representative or biased in some manner. A disadvantage of mail surveys is the returned response rate, which may demand that reminder postcards be sent to participants in order to increase the rate. Conversely, survey data may be difficult to analyze when people do not fill out questionnaires completely. Using a questionnaire to collect data can either precede or follow a qualitative data collection phase. Each phase can also be used to test results that emerged during the other phase, allowing for a better understanding of the research topic. Both approaches to data collection have advantages and disadvantages. By combining the methods, the researcher may be better able to understand the patterns that emerge.

#### ➤ Experiments

Experiments are a powerful method for collecting quantitative data, in which the researcher manipulates an independent variable (IV) and measures the impact on the dependent variable

(DV). The experimental design allows researchers to infer a causal relationship between an IV and a DV, typically referred to as an effect, while controlling for covariates or other factors that may confound the relationship. Other exploratory, correlational methods cannot be used to infer causation. Researchers use a variety of experimental designs in pursuit of causal inference. The gold standard randomized controlled trial (RCT) probabilistically assigns subjects or units to either a treatment or control for the purpose of creating balanced groups before collecting data to compare the treated and untreated. Many variations of RCTs exist, including using blocking or stratification to control for covariates, running experiments in the field or laboratory, and using cross-over designs. Beyond RCTs, many quasi-experimental designs are used to identify causal relationships. For example, researchers may exploit natural instruments, or matched samples with observable characteristics that are similar on both sides of a threshold. Propensity score matching, difference-in-difference designs, and synthetic controls are also examples of quasi-experimental methods that rely on control groups and covariate balancing, as well as potential outcomes. Each of these methods has assumptions that must be verified or satisfied for extracting causal identification.

### **Qualitative Data Collection Methods**

Qualitative research refers to a type of research, usually performed in social sciences, that focuses on descriptions, words, concepts, characteristics, and their meanings rather than numbers and the relationships between them. It is often used when the objective of the study is to better understand an issue or phenomena rather than to quantify a relationship, estimate a parameter, or test a hypothesis. Qualitative research can be used comparatively or experimentally but is most common with exploratory or interpretative research. The most common qualitative data collection methods are interviews, focus groups, and observations. Interviews are a dynamic information collection method and a key technique to capture the meanings. The interview is a conversation between two or more people where the interviewee is invited by the interviewer to talk about himself or herself and about a topic or topics that the researcher has defined from a previously specified theoretical background. The researcher already has some insight into questions that arise or issues that require clarification. Interviews can be conducted face-to-face, via telephone, or using electronic communications. They can be unstructured or semi-structured. In the unstructured interview, the researcher has no preconception about the issues; the interviewee reacts freely, and the direction and depth of the conversation are in the hands of the interviewee. In the semi-structured interview, the researcher has prepared a list of questions or topics to discuss but holds a conversation

around those subjects rather than conducting a Q&A session. In-depth interviews are usually done individually, loosely structured, with open-ended questions. The time frame is extended to allow for a more in-depth conversation.

Focus groups are group interview techniques used to consolidate and develop people's ideas about a topic, so focus groups can be seen as an attempt to generate discussion rather than individual responses. The discussion is centered on a topic chosen by the researchers. The topic is proposed by the researcher, and, with the collaboration of the participants, explored in depth. These discussions happen in small groups and are more controlled than normal group discussions. They last between one and three hours, are audio or video taped, and the moderator can take field notes. Focus groups are meant to stimulate interaction and amplify participants' perspectives. Observations also involve fieldwork and the collection of information or data through direct observation rather than interviews or other forms of self-reporting. Ethnographic and some case studies are characterized by long periods of observation in the field.



Figure 5.2 Qualitative Data Collection

### ➤ Interviews

The interview is an interactive and dyadic data collection method that relies on a series of questions and answers. Interviewing is a "Ritual of Uncertainty Reduction," broadly defined as "any question or information sequence through which one participant seeks to gain information about the experiences and thought of another." An interview is a two-person conversation initiated by the interviewer for the purpose of obtaining information from the interviewee. An interview differs from other conversations in that it has a structure and a

purpose. Interviews and interviewers come in various shapes and colors. Interviewing is a characteristic of human social existence. Interviews have an important, if modest, role in a variety of practical settings: advertising, academic research marketing, and patient-physician contact. An interview is an interaction of two people, whereby in response to the questions posed by the interviewer, the interviewee reveals information that relates to the researcher's specific area of inquiry. Surveys typically offer fixed responses, and although they can force respondents to choose among answers given, they also potentially restrict their response. In contrast, interviews are open-ended, exploratory, and flexible. Due to the dynamic nature and social interaction of a face-to-face interview, the interactivity and spontaneity of a conversation help minimize the distance. The fact that interviews elicit qualitative data about the interviewee's lived experiences can help researchers explore research questions that involve sensitive issues, social taboos, and non-visual phenomena.

#### ➤ Focus Groups

Focus groups are a qualitative research method that involves bringing people together in groups to discuss a specific topic of interest. Through a facilitator, focus groups enable in-depth discussions among participants, harvesting ideas and perceptions in an interactive group environment. Focus groups generally capture participant perspectives through their interactions with other participants and the moderator, using tools such as active listening and follow-up questions to guide discussions. Focus groups require trained professionals to carry out because the way the facilitator interacts with participants can easily influence results. Focus groups capture rich qualitative data based on participants' feelings and insights on research topics. The discussions provide a common framework for participants to express their thoughts and provide the researcher with an understanding of how and why particular ideas are relevant to those participants. Focus groups are helpful qualitative research tools for program development and initial consideration of health issues. They also provide useful input for the development of larger quantitative studies. Sometimes, researchers use them as the main data collection method in a study, although they are often used to augment information gathered from other methods. Focus groups are sensitive to the social context and interaction dynamics among participants, making them particularly beneficial for collecting data on topics that have a social or group-related quality. They are especially useful for providing collective insight into populations that would be difficult to study using more standardized methods. Given the interactive nature of focus groups, they are especially

suitable for topics that generate a wide variety of potential responses or involve complex community-based issues.

➤ Observations

Observations are probably the only data collection method for qualitative research which has their origins in the traditions of anthropology and sociology. The social anthropologist first does a field study, immersing him or herself in the field or the community under examination over a fairly long duration. During this period, he or she collects data by observing how the people in the community behave or what they do, and the various artifacts that the community has. This method is called participant observation because the researcher becomes part of the community and actually lives with the people. Observations can also be done from a distance, which obviously would not be as rich as participant observations. In sociology, observations have been an important method for collecting qualitative data in studying social problems and urban life. In qualitative research, observations are particularly suitable for use when the field setting is considered less artificial and the inquiry concerns behaviors that cannot be easily verbalized or when complex processes are involved or when the participants are unwilling to talk. Observations as a data collection method have the following distinctive features. To begin with, observations have no predetermined structure. The researcher goes into the field with an open mind, without the agendas set by a hypothesis testing study. Observations are, therefore, more exploratory, and the findings from an observational study are often richer than that come from an interview or survey study. Although qualitative researchers who conduct an observational study are uninterested in testing hypotheses, hypothesis generation is also a possible outcome. However, case study researchers sometimes use observations as a data collection method in conjunction with other methods to validate or compare information collected from other sources.

### **ONLINE PLATFORMS LIKE GOOGLE FORMS, MICROSOFT FORMS, AND MOBILE-BASED APPS**

The widespread availability of low-cost computing devices and mobile phones, along with the rising use of the Internet, has increased our reliance upon online data collection for research. Online data collection is a cost and time effective approach, enabling independent, trustworthy research, avoiding interviewers' effects. Oftentimes, filling in the responses online appears to be more comfortable for the respondents compared to filling a printed survey; they can be at the same time in their natural environment while trying to answer as

sincerely as possible. Innovative online data collection tools have greatly simplified the process of creating, managing, distributing, and analyzing data, making such tools appealing methods for novice survey researchers. The oldest and simplest method of conducting online surveys maximizes generalizability and turns to online data collection platforms. These platforms provide a very simple way to create and distribute surveys. You can create a survey, customize it with different themes, and share it in a variety of different ways. Responses are updated in real time in a spreadsheet, and you can limit them to prevent multiple submissions by the same person. Similar capabilities are available in other platforms. In user studies about online survey platforms, however, survey users often like survey applications even better. These mobile phone applications make it easy to collect and analyze data. They can prompt the user to answer periodically. Users can easily attach data, take notes, and include pictures as part of the answers. While there is often an expense associated with mobile survey applications and they may require more expertise to distribute and analyze, they often provide a much improved user experience and often remember choices in previous questions.

➤ Google Forms

Google Forms is a free application used to create forms and surveys for data collection. The data collected by the form is saved to a spreadsheet, making it instant and ready for analysis. Google Forms offers many sophisticated features such as the creation of customized questions in various formats with multiple conditional branching options to suit almost every research need. Google Forms allows researchers to publish their surveys and forms in multiple formats, including embedding them on websites, sharing them as pop-ups, or sending them via various social networking sites or emails. Google Forms is accessible from web browsers as well as from Android and iOS applications on mobile devices, making it a convenient platform for survey development. Researchers having an account can use Google Forms; no special software installation or training is required to use it. Obtaining data using Google Forms is free and takes only a few minutes to create a survey and share it for data collection. Data from the questionnaire results page can be directly exported from Google Forms to a spreadsheet for analysis. Google Forms also offers options for basic data analysis features such as graphing data from multiple-choice, checkbox, and dropdown questions. The forms created in Google Forms can be used to conduct online surveys for any kind of research data collection without needing any special coding language but providing multiple conditional branching and customized options. Google Forms has the option of creating customized templates for regularly conducted surveys, which can also be set for repeated uses

in future research. Google Forms is an accessible data collection tool for researchers and is used for various types of online research data collection, specifically in the field of health and education. The current report discusses the process of developing surveys and options available in Google Forms to enhance online research data collection.

➤ Microsoft Forms

Microsoft Forms is a tool that students and teachers can use to create surveys, quizzes, and polls. As part of Office 365 Education, it's easy to use, sharing and collaboration features allow results to be viewed in real time and analyzed in Excel. Microsoft Forms provides students with a simple way to consolidate thoughts, gather peer feedback, and conduct research. While much of the functionality is similar to other online data collection platforms, there are slight differences between the two in terms of capability. Although Microsoft once had a non-consumer version to support its easy online data collection tool, this has now been withdrawn. The current version of Microsoft Forms is associated with your Microsoft account, but will always be free to use, will show no adverts, and will allow your users to create as many forms as they want, as long as the forms are used on a small scale. Forms associated with personal accounts cannot currently be used to create quizzes, which can only be created with Forms that belong to an Office 365 account. The primary use case for Microsoft Forms is to create surveys and quizzes, either for a one-time or recurring purpose. Surveys can be simple for informal data collection, or complex for more serious investigations. Quizzes can be used for simple tests or as educational tools at the end of a unit of study. Questions can include open response, multiple choice, rating, or likert agreement scale. In addition to English, there are currently 107 supported languages available in Microsoft Forms.

➤ Mobile-Based Applications

Mobile-based data collection applications that allow researchers to build and deploy surveys and other data collection protocols have gained popularity over the past several years. Mobile-based applications allow mobile devices to be used for quantitative and qualitative data collection in a variety of subject areas, including health, epidemiology, economics, political science, marketing, anthropology, and many others. Researchers build surveys for mobile delivery using a web-based user interface, which then populates back-end database(s). Data can be collected and synchronized in real time, allowing centralized monitoring of incoming datasets. Most mobile-based applications are primarily designed for survey data

collection but can also accommodate mixed-method data collection with mobile-optimized text entry for qualitative data. Many applications allow geo-tagging of quantitative survey response data as well as images and videos for qualitative responses. Mobile applications typically function in mixed connectivity modes, allowing surveys to be collected with or without internet access. Although some applications have been designed to accommodate low bandwidth and low-memory settings by using drop-down menus in lieu of visual designs, and minimize the amount of data sent and stored, mobile applications require reasonably powerful modern devices to optimize battery life, maximize screen viewing for complex tasks, and minimize download and upload times. Most mobile-based applications have limited capabilities for more complex survey designs, such as conditional branching or ordering, audio, video, or interactive task-based elements. Most applications are not designed for complex qualitative data collection or question-answering protocols. Survey tools that offer user-friendly, accessible, and sophisticated designs are generally not designed to take advantage of mobile applications and have no offline data collection capabilities.

### **ENSURING VALIDITY, RELIABILITY, AND ETHICAL DATA COLLECTION**

While selecting data collection methods, the researcher must consider validity and reliability issues. The results of the research can only be generalized if the data collection methods used have such attributes. The research results lose scientific value if these two criteria are not passed. Before focusing on them, it is worth determining which of them comes first. However, many point out that validity is the superior attribute. With all this in mind, it is worth ensuring validity and reliability prior to data collection. Validity can be defined as the extent to which a test measures what it claims to measure. The simplest method to ensure validity is to rely on common logic or expert judgment. Using that explanation, we could consider validity a synonym for the word “truthful”. In contrast, some consider scientific validity to be socially constructed. Thus, validity translates into the degree to which social actors believe that a research accurately assesses the interpretations of reality that they present. Qualitative researchers would say that validity is higher when the perspectives of the subjects studied are pluralistically represented in the interpretation of the phenomena studied in the research report. Reliability is the second attribute proposed for the evaluation of scientific status. Reliability is a measure of the degree to which a research instrument consistently measures an unchanging attribute when the measurement is repeated. The degree to which conclusions and inferences can be trusted to be reliable. It can also be defined as the extent to which a test yields consistent results or a test can be repeated on different occasions, at different times

with similar results. However, subjective approaches regarding the subject of reliability have been proposed. For example, some qualitative research authors blame the traditional concept of reliability. A justification is that qualitative research focuses on description and understanding rather than on the prediction and quantification of events.

➤ Definition of Validity

In social research, as in psychology, it is most important that a scale measures what it is intended to measure. This aspect of measurement is called validity. Whether they be called by scale, test or index names, presumably at the core of one's concern is that an observed score is both an accurate and a valid one. An invalid scale cannot be interpreted correctly. But an invalid scale or test cannot be entirely discarded. Therefore, degrees of validity are assigned to measures – although certain standardized tests are purported to be 100 per cent valid. Generally, a distinction is made between three forms of validity: face validity, criterion validity, and construct validity.

1. Face Validity. Noting the word 'face', this flag may be considered as a layman's view of validity. The concern is, at face value, whether or not a scale or test measures what it is supposed to. Other things equal, an obvious attitudinal question is the item about, say, whether one strongly agrees that elderly people are a burden on the rest of society. Obviously, this item is a step towards achieving measuring something called 'ageism'. One furthermore hopes that there are others measuring different dimensions of this attitude. But, also obviously, this item cannot hope to be a 100 per cent indicator of ageism.

2. Criterion Validity. This form of validity is based on the verification of an actual score (based on the use of a scale) and the scale's 'gold standard' which is some external criterion. In the example above of measurement of 'ageism', you might want not only to compare the variance of a test item about undertaking volunteer work for the elderly with the variance of a theoretical model along these lines with your item, but that both distributions somehow agree with each other. If there is a 100 per cent correlation (of course with the same distribution) between both variables among seventy year-old males, there might be some justification for using the question as a measure of the concept 'ageism'.

➤ Definition of Reliability

Reliability is an important aspect of research design. It has been suggested that reliability is the degree to which data collection methods and measures produce stable and consistent

results. Data can be said to be reliable if, when a complete research study is repeated using the same methods, the same or similar results are obtained. For example, if a researcher repeatedly weighs himself on the same set of scales, he would hope to obtain the same result. Reliability often refers to the consistency of results across a number of measurements and the reliability of a research study is therefore assessed by examining the consistency of results over time and the degree of similarity between different measures of the same variable. It is possible to have reliability without validity; that is, people can become good at showing you the answer you want without it being a valid. Reliability is concerned with the accuracy of research findings. The reliability of qualitative data is frequently the focus of a debate between qualitative researchers and the proponents of the quantitative approach. It is possible to have reliability without validity; that is, people can become good at showing you the answer you want without it being a valid.

➤ **Methods to Ensure Validity**

Researchers can use several strategies to ensure the validity and accuracy of their research. It is wise to use more than one method to check validity. It is usually easy to achieve content validity; the most judgmental of the types of validity, it is a necessary but not sufficient condition for test validity. Other strategies might be necessary once the study is initiated. The strategies used often depend on the research design model and specific method being employed. For example, qualitative researchers tend to perform participant validation more than do quantitative researchers. On the other hand, more quantitative researchers can establish criterion or predictive validity by empirically comparing participant responses with the same results of other tests that assess the same concept. There are several techniques to confirm the study's content validity. During the study's conceptualization phase, a panel of authorities on the area of inquiry can identify the concepts covered; panel members might have to review the warehouse of interviews, field notes, participant and researcher checks, etc. They can use this technique only in a more qualitative study once data have been gathered, although it might not be as accurate, especially if the area of inquiry is less-known, which can compromise the codebook. The other technique requires an established test currently in use and comparing participants in the new test. The difficulty with this technique is that the criterion of the established test is sometimes impractical, difficult, or impossible to replicate.

### ➤ Methods to Ensure Reliability

Reliability ensues from a consistent outcome or application of the measurement tool. A tool is reliable if it returns similar results upon repeated measurements. The re-applied measurement must more closely resemble the first one than a measurement taken by a different tool. The need for a reliable research tool is vital since unreliable research will produce false conclusions and require repeat actions. In order to ensure reliability, quantitative researchers often run tests and re-tests using the same sample or different samples, applying the same measures. In effect, the participants must be measured more than once, within a time gap and with measurement stability ensured. In qualitative research, reliability determines how well the research documents the case and how well it describes the sample. Problems can arise not in the measurement tool but within the measurement process. In qualitative research, reliability is ensured by the transparency of the methods used, allowing them to be replicated in different contexts and times using different researchers without affecting reliability. The information must be systematically organized, rich in detail, represent different viewpoints and include thoughts and interpretations of the participants, and be well summarized in the interpretation and discussion. Another method to ensure qualitative reliability is through the recognition of participants' voices, allowing them to be openly quoted, substantiating the authenticity and credibility of the document. Reliability is further guaranteed by longitudinal studies, confirming if the sequence of events is indeed in the same order.

### **Chapter summary**

"Data Collection Methods," delves into the crucial aspects of gathering information for research. The chapter begins by establishing a fundamental distinction between primary and secondary data sources. Primary data represents original information collected directly by the researcher, while secondary data involves utilizing pre-existing information gathered by others. This distinction is critical in determining the appropriate methodologies and understanding the inherent strengths and weaknesses of each approach. Following this foundational understanding, the chapter explores various tools and techniques employed in data collection. This includes a detailed examination of survey methodologies, encompassing the design of questionnaires and the selection of appropriate sampling techniques. Interview methods, both structured and unstructured, are also discussed, along with the nuances of conducting effective and insightful interviews. Finally, the chapter covers observational

techniques, highlighting the importance of systematic observation and the potential biases inherent in this approach.

To further enhance the practicality of the chapter, various digital tools and platforms relevant to contemporary data collection are introduced. The chapter highlights the accessibility and user-friendliness of online survey platforms such as Google Forms and Microsoft Forms, and also discusses the growing importance of mobile-based data collection apps. These tools are presented not merely as convenient alternatives, but as integral components of efficient and effective modern research methodologies. Concluding the chapter, a vital focus is placed on ensuring the validity, reliability, and ethical conduct of data collection practices. The importance of using unbiased methods, ensuring data accuracy and consistency, and adhering to ethical guidelines for data privacy and informed consent is underscored. This section emphasizes the responsibility researchers have to maintain the integrity of their data and respect the rights and privacy of participants throughout the entire data collection process.

\*\*\*\*\*

## CHAPTER 6

### MEASUREMENT AND SCALING TECHNIQUES

Measurement is a powerful tool for collecting quantitative and qualitative data that is helpful to all disciplines. While quantitative data is typically made with tools and physical equipment, qualitative data is often created by careful observation and thoughtful assessment of the sample being measured. Regardless of the type of measurement used, the underlying processes are fundamentally identical: to take an abstract construct and derive an observable behavioral example. In instances of quantitative research, the focus is to conduct empirical association and causal research that will lead to generalizable findings. With qualitative research, the objective is to assess.... These tools and techniques for measurement and scaling are important to the University of Pittsburgh School of Information Sciences, Department of Library & Information Science, not only for the obvious fact that many of the researcher's databases are created through quantitative methodology nor that collection management, market analysis, and other job tasks are based on quantitative models. Measurement in general, as well as, measurement in both the qualitative and quantitative research paradigms will teach the individual valuable and applicable skills.

#### LEVELS OF MEASUREMENT: NOMINAL, ORDINAL, INTERVAL, RATIO

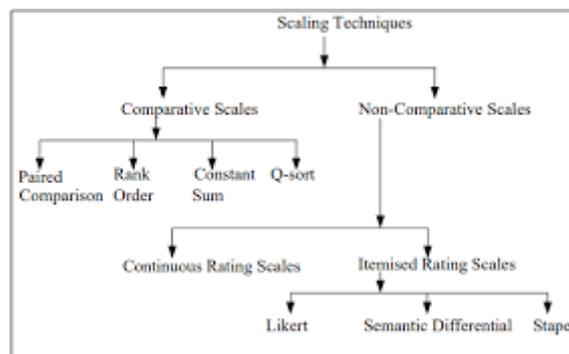


Figure 6.1 Levels of Measurement

A common question asked in measurement is, "How much of the attribute (e.g., attitudes, beliefs, intentions, and motivation) are we measuring?" The answer to that crucial question brings us to the topic of levels of measurements. It has been said that "the level of measurement" is the unit of account with which the scientist registers the result and that "different units of account are appropriate" at different stages of the research process. The developmental psychologist advises researchers that "it is very important to the whole

organization on the study and to its relation with those that follow to record in the right way" and that "in order to do justice, the right quantitative determinations must be made." Without that prior determination, the scientist falls into the risk of falling prey to a mistake made by "introducing the linear unit to measure the volume of a solid" and asks for it to measure "lengths of different orders," for which "the foot-metric contrivance must be called in." There are a number of different classification schemes, but the best known is that developed in the 1930s. A four-level classification scheme for measurement is based on the type of arithmetic performed on measurement data. The first level, nominal-level measurement, is just identification through labelling and allows for only qualitative comparisons. The second level, ordinal-level measurement, allows for a qualitative ordering of the data. The third level, interval-level measurement, adds the possible performance of addition and subtraction operations to the previous two levels and temporal comparisons; thus, both intervals between scale values can meaningfully be considered and equality or difference in value determined. The fourth level, ratio-level measurement, is the highest level of measurement, allowing not only for the performance of all arithmetic on the scale data but also proportion comparisons of the variant of the attribute being measured that is of interest to the researcher.

#### ➤ Nominal Measurement

Four levels of measurement exist — nominal, ordinal, interval, and ratio. The first level of measurement results from the mere categorization of data. Such measurement is called nominal because the only mathematical or statistical operation permitted on the data is to count the number of responses in each nominal category. Frequently, one will compute the percentage of total responses for each category. Some researchers also distinguish between nominal measurement that uses names or labels, and pronoun nominal measurement of data that is assigned a number that does not have any mathematical or statistical meaning except to act as a label (a student number would exemplify this second form). Nominal measurement only provides information that displays that one object is different from another. In social science research, most variables are assigned a nominal measurement. Examples include gender, which can be recorded as female or male; race, which can be coded as black, Asian, Latino, or white; religion, which might be assigned labels of Protestant, Catholic, or Jew; class status, which could be labeled working class, middle class, or upper class; or marital status, with categories of single, married, divorced, or widowed. For symbolic items like those mentioned above, it is important to note that the distance between any two items is meaningless in terms of how similar or dissimilar the two items actually might be. An Asian

responding to a question about race might be coded “2,” a Latino might be coded “3,” and a white person might be coded “4.” The distance between codes “2” and “3” or between codes “3” and “4” has no meaning in terms of the character of the response. An Asian is not exactly more similar to a Latino than to a white person.

#### ➤ Ordinal Measurement

Ordinal measurement is a step beyond the nominal classification system in that it has the ability to put things in rank order. For example, it usually makes sense to say that "good" is higher in value than "fair" if you are using a rating scale to describe the perceived quality of product advertising. We could add more categories such as superior, excellent, and poor, but this does not add much since there are a relatively small number of adjectives people use to describe their perceptions of quality. Certainly, it would make no sense to say that "good" was five times as good as "fair" which is theoretically required if the numbers are to reflect a difference in an abstract underlying dimension. The measurements, because they are ranks, can thus only indicate more (higher) or less (lower) of the attribute being evaluated. In this sense, ordinal measurement is more like interval measurement in that there is a common statistical processing assumption behind both types of data. The practical implication is that, if possible, you should use semantic differential or Likert scales with an odd number of values (including a neutral or no-opinion category) to denote degrees of differences between adjectives. You're also more likely to get honest responses since the higher the number of choices available, the more likely a respondent is to find a category that describes their feelings precisely, and they will avoid the trap of forced choice in a complicated judgment to themselves.

#### ➤ Interval Measurement

An interval scale has all the properties of an ordinal scale, but in addition is more informative, and the distances between scale points are equal. The common assertion that an interval scale conveys equal intervals is misleading. Metric scales convey ratios of distance, so it is misleading to assert that an interval scale conveys integer distances. Ratio relations for distances lie at the core of metrics. Therefore an interval scale conveys a weaker metric, or a metric of lesser generality, than a ratio scale. A more informative scale must use explicit labeling such as map convection or a report on the criterion. In practice the scale is an interval scale when the specified scale corresponds behaviorally to the phenomena of interest (namely when products of the phenomena taken at the point in the scale where the distance is

labeled one unit correspond conversationally to product mappings in the social psychological colloquial sense). Because an interval scale provides information on the distance between points on the scale, interval scales can be used with more complex analyses than nominal or ordinal scales. For instance, in the case of both ordinal and interval data, rank order is not only meaningful but is also quite informative about the measure of interest. In the case of interval data, however, we can also reason that a difference score—informed by an actual numeric distance on the scale—is meaningful, and we can construct further properties from examining participant's responses. In studying attitude theory more deeply one notices that a Likert-type item does in fact behave like a true interval scale in some aggregate settings. This is informative for the added assumptions used in these analyses, and these models attempt to justify the assumption of true interval data.

#### ➤ Ratio Measurement

Ratio measurement is the highest level of measurement in social science and thus enables the researcher to understand and examine the phenomena of interest in the fullest way possible. Ratio scales not only allow researchers to describe differences between and relationships among characteristics, but they also explicitly allow comparisons of absolute magnitude between the measured characteristics. Ratio scales are thus not only syntagmatic but also paradigmatic. Naturally, there are few characteristics that can be measured by ratio scales. However, such characteristics as height, weight, and speed can be and have been measured. All of these characteristics have an absolute zero that has meaning. For example, zero weight or zero height means that no weight or no height exists. In fact, meaning must always be attached to zero or ratio measurement is impossible. Since we can usually simulate absolute zero with other characteristics that are arranged in ordinal, interval, or perhaps nominal scales, it would seem that ratio measurement is seldom necessary. However, some variables in the various social science disciplines, such as rate of mobility in sociology or number of people in urban planning, lend themselves to ratio measurement. In such cases, the research results would surely obtain the information richness denied by any of the other scales.

### **SCALING TECHNIQUES: LIKERT, GUTTMAN, SEMANTIC DIFFERENTIAL**

It has been said that measurement is the language of research; scaling techniques are one way for researchers to create the ideas that make up that language. And just as a writer must consider the words he chooses to use, the researcher must also consider how the purpose of a relevant scale construct will lead them to select the appropriate scaling technique to use in

their survey. Based on the different goals of the researcher, there are different types of available instruments. If the researcher has well-defined scales, scales can be used to measure and as a result, classify data. If the researcher has structured scales, they can be used to measure an existing attitude, belief, or perception, and can therefore predict behavior. Or if the researcher has semi-structured or unstructured scales, they can be used to better explain some subjective aspect. We will now take a closer look at the types of techniques available for researchers to scale constructs and their measures. Scaling techniques are the most valuable tool available to marketing researchers, allowing researchers to measure concepts that are difficult to define operationally such as ideas, attributes, perceptions, and beliefs, among others. In marketing research, scaling techniques have been used to help determine a product's position in the marketplace, determine product and/or brand image, measure advertisement effectiveness, and help discover which attributes would be most effective in an ad. While different scaling techniques exist, only a few are as popular or widely used. Presented here are a brief description of the most popular scales: Likert scale, Guttman scale, and Semantic differential scale.

➤ Likert Scale

In 1932, the Likert technique evoked a concept which differentiated towards response scaling. It is a frequency approach based on item formation theory specifying the response to independently expressed statements. Its origin derives from people as "dichotomously weighted" positive-negative scales. It is installed as the extensive survey-metric classic instrument. Its utilization is not exclusive of opportunities of access to a population sample, which use other survey techniques such as the Questionnaire, Assisted Interview or Phone Interview. The longevity and basic acceptance of the Likert scale is mainly owed to its provisional nature. For research purposes, it rates the probabilities of the Likert conceptions. The "Likert" becomes characterized by the ordinal measurement that scales balance and events through its response situations. Its intransitive use is implemented in limited number of anchoring points, suited to answers testing with stimuli affected by the associating word. These may be: a silence word for neutral position goal, sometimes with the idea of agree-disagree only or other polar terms. Appearance determines subject subsistence for consideration, investigation and response goal. Discussion has been motivated recently towards the utilization of more points in the Likert scale. Hitherto it has limited the proposal to maximal point number of 7, sustaining the base hypothesis of each additional point the transfer of its utility in its utilitarian theory. This adds the usual statement focus through the

probed reaction proposing or denying ideas, feedbacks, affirmations and loose appreciations. The psychological reaction acts incorporating expanding distances to the associated event. The argument justifies an expert approach permitting to moderate conclusions by facilitating the spotlight of further consideration. Others remain reluctant to scale increasing points. They prefer the 5-points system, defending the standpoint option binary in demonstration of psychological opposition to stimuli presented with pole-motivated terms, exerting cognitive forces in its sectional anchoring.

➤ Guttman Scale

In the Guttman scale, also known as cumulative scale or scalogram analysis, the items represent intensities of the attribute being measured, ordered from weak to strong. As a respondent agrees with an item, he agrees with the one above it along the scale. These items can never be chosen in a random manner; ordering is essential which enables the Guttman scaling technique to measure the cumulative intensive part of the measuring attribute. Consequently, it is a non-random method since each item tested derives from the position of the item underneath it. The Guttman scale, which although in use well beyond its original intent, was invented by Louis Guttman in psychology. Due to its essential non-metric methodology, Guttman scaling is often misunderstood or overlooked in the fields of educational measurement and policy structuring. In psychological facets it may be helpful for both practitioner application and research. The Guttman scale synthesizes ordinal item data into multiple binary item-scale data, on which statistical operations can be performed.

A Guttman scale allows the question of interest, attribute or attitude being studied to be placed in the center of the research instead of being dispersed over different un-related questions. This increases reliability to a very high level graded higher than most attitude- or personality measures. Satisfaction, loyalty, importance, communication, needs are only a few attributes or attitudes for which Guttman scales have proven themselves to be reliable.

➤ Semantic Differential Scale

Semantic differential scale is an ordinal type of scale used to measure the connotative meaning of objects, events, and concepts, using a seven-point rating scale. It is a series of bipolar adjectives at each end of an odd-numbered scale where respondents indicate their position. This technique was developed during research on Affective Meaning in 1957. The scale consisted of as many pairs as there were categories and offered a range of five steps

between the opposing adjective pairs. The importance of determining the optimum number of items to be included was frequently discussed. Furthermore, since a multi-item scale was assumed to measure a single dimension, there was a need to check for the internal reliability of the results. Later on, a principle of a single-dimension checking procedure was dismissed by the use of large enough samples to carry out factor analysis. Semantic differentials can be used as single-item scales or multiple-item scales. In the single-item version, respondents check one item on its scale. This scale format has reportedly enjoyed more success in multi-item scales. Initially, in three required studies, data from nearly 300 adult respondents were analyzed and underwent factor analysis to define the items that should be included. An attempted validation of a five-dimensional solution based on a semantic differential grid format using individual items resulted in less than satisfactory results. These scales have recently been described and used in scientific papers across a multiplicity of discipline areas such as marketing, management, linguistics, and political science. The popularity of this tool may rely on both its flexibility and very simplicity, not to necessitate higher order cognitive processing in the technique.

### **Reliability in Measurement**

Reliability is the first of three properties any measurement should have. It refers to the extent to which a measure is free from random error and offers consistency across time and across various items used to measure the same construct. Measurement theory teaches us that to quantify a construct as abstract as an attitude using a set of manifest variables we have to accept the existence of random error in our item responses. Accordingly, a numerical score describing an individual's attitude on that latent construct is a reflection of that individual's true score plus an amount of random error associated with the act of measurement. There are four common sources of measurement error: the person being measured (e.g., error due to mistake or carelessness or random error in self-perception); the person measuring (e.g., error due to fatigue, bias); the instrument (e.g., the questionnaire designed, instruction used; error caused by failure to pre-test an instrument, failure to consider the clarity of specific items in a questionnaire; spontaneous changes in a respondent's cognitions or dispositions); and measurement conditions (the degree of precision we can obtain in objectively assessing individual characteristics; testing conditions, including physical interference such as temperature, lighting, noise, etc; reliability is reduced if measuring conditions fluctuate).

### ➤ Definition of Reliability

When we measure parts of a complex phenomenon like human behavior, we expect our measurement to reflect the characteristics we are trying to measure, without being distorted by unrelated influences. If we try to measure something four or five times and find the measurements differ widely each time, we are naturally skeptical about any inferences we might make from the measurements. We expect that measurement of a particular object, thing, person, or group under particular conditions should have relatively the same value—whether that object has been measured just once or many times. Certainly, the degree to which repeated measurements yield consistent results reflects on the quality of the measurement. This is the essence of the concept of reliability, which is defined in many ways. In simple terms, reliability is the degree to which measures are repeatable, reproducible, stable, and consistent with themselves across time and researchers, and researchers are increasingly concerned that measures need to be highly reliable in order to be considered valid, and every instantiation of the measure should yield comparable results. Measurement reliability refers to the extent to which a measure or indicator is a consistent measure of a concept, that is, the amount of error in it, but more precisely the proportion of measurement error variance relative to total variance. A reliable measure has reduced error relative to true variance. A measure that is not perfectly reliable does not mean it is not a good measure. A measure could have low reliability but still have good validity.

### ➤ Types of Reliability

Measurement error can result from various sources so that it is only logical, in fact, required, that we should assess the reliability of a scale from different points of view. Although there are different types of reliability, perhaps the three most frequently used are the following.

**Internal consistency: Split-half reliability.** Reliability, simply put, includes an assumption that a test or scale is homogeneous, or actually measures the same construct or factor throughout the length of the test. Internal consistency measures this homogeneity. Internal consistency can be assessed by calculating reliability coefficients, based on everyone's abridged scores on the different halves of the test. If there is perfect agreement between the two halves, and this is usually not the case, the reliability coefficient is equal to one. The Spearman-Brown Prophecy Formula corrects for the fact that when half the length is half the total score length you get a lower reliability than you would if any length permutation quotient of a test or scale were the same.

Test-Retest Reliability. Test-retest reliability improves dimensionality validity by providing additional evidence that the scale reliably measures what it was intended to measure. Test-retest reliability is simply the correlation among the same subjects who take the same survey at two different times. Time is an important element in test-retest reliability not only for collecting data but especially in considering the lapse duration between the two testings, and the time lapse does not create any additional issue in gathering sample data again.

#### ➤ Assessing Reliability

Given the key role of reliability in the measurement process, their first step following the development of a new measurement scale or the adaptation of an existing scale is to evaluate its reliability. Reliability assessment can be accomplished in a variety of ways. These methods are classified into two major approaches: computation of correlation coefficients and observed score variability estimation. Computation of correlation coefficients: A common way to assess the reliability of a measurement scale is to compute the correlation between two scores obtained from the same respondent on multiple occasions or from two different but similar scales or from two and only two items that two respondents selected. The reliability coefficient is represented by a correlation index. By convention, coefficient values below 0.60 are viewed as being low, those between 0.60-0.75 as being moderate, those between 0.75-0.90 as being strong, and those above .90 as being very strong indicators of reliability. Observed score variability estimation: Estimating the amount of observed score variability that is attributed to measurement error is an alternative way of assessing reliability, and the amount of observed score variability attributable to error can be calculated by employing the following general equation: measurement error = observed score variability - true score variability. The estimates of the true and error score variances are computed based on the test-retest and item-item approaches.

### **Validity in Measurement**

The purpose of measurement in research is to determine the amount of a property that an object possesses. Because popularly used physical and social measurements such as height, weight, temperature, and years of schooling have validities that are taken for granted by their users, the concept of validity tends to be overlooked. To psychologists and other researchers in the behavioral sciences, however, the concept of validity is of vital importance because the measurements on which their research rests are often either entirely or in large part self-constructed scales on broad and abstract concepts, which differ greatly in quality. The

magnitude of correlation between such a self-constructed scale and tangible behavior or between two self-constructed scales tapping the same concept is often touted as a measure of validity. While correlation is concordance in the domain of the operational definition of the scale, it is not validity. The validity of a measurement is the extent to which it actually measures, as opposed to, for example, relates to, the construct. The operational definition cannot fully represent the concept being measured, and different self-constructed scales tapping the same concept may correlate only moderately; yet both could have the same validity. Even so, such correlations can be used as (albeit imperfect) indices of validity.

### ➤ Types of Validity

Unlike the other qualities of measurements such as reliability and equidimensionality, validity cannot be quantified, but must be qualitatively established. While some researchers write on the three types of validity – content, criterion, and construct – others reduce it to two types by combining content and criterion validity. Content validity is the qualitative aspect of what the scale is actually measuring. It refers to whether the explicitly stated definition of the concept and the operational definition of the scale together encompass a sufficiently representative sample of the behaviors that constitute the concept being measured. After having established the definitions of measurement, measurement scales and research scales, it is now appropriate to explore their properties. A natural starting point is the quality of validity in measurement. Validity is the quality of truly measuring the concept one intends to measure. A scale is said to be valid if it actually measures what it is supposed to measure. For instance, if one is measuring retail price perceptions among consumers, and the measure is approvingly stating that "The prices of the items I buy are very high", then that is a clear case of a lack of validity. We are measuring something else, possibly self-esteem regarding the consumer, instead of price perceptions. Unfortunately, a more insidious threat is always operational. In social-sciences measurements, constructs are usually abstract, such as likelihood to recommend a hotel, future involvement with a brand, or mental imagery in the consumption of art. Furthermore, measurements are seldom playful and sometimes rather serious business, such as predicting a new product's success or predicting a company's corporate reputation among consumers. The insidious threat is that while researchers want to measure a construct, a different construct is being measured with the empirical procedure defined by the researchers. Sometimes researchers want to measure self-esteem, and the construct being actually measured is perhaps mentally imagined difference among questions. Sometimes researchers want to measure schema theory but what is really being measured is

something slightly different. If one is questing abstract, broad constructs, people identify behaviors to them and these behaviors are usually observed, sometimes through real life day-situations, sometimes through controlled settings in the lab. Other times the different definition is that the different behavior is being inferred a little less objectively through qualitative single case analyses on relevant behavior. When multiple concordant conclusions emerge, then that special behavior seen is indeed a construct.

#### ➤ Types of Validity

Validity is important to research because a poor measure prevents researchers from answering their research questions. Many measurement experts have described types of validity, including a multitrait-multimethod technique for testing the convergent and discriminant validity of measures. This work led to a large number of published validity taxonomy systems. From this work we extract the following types of validity: (1) content validity, (2) criterion-related validity, (3) construct-related validity, (4) internal validity, (5) external validity, (6) statistical conclusion validity, (7) face validity, and (8) consequential validity. Validity is a strength of the conclusions, inferences, or actions based on a measurement or other research activity. The concept of validity is not unequivocal. In the popular view, validity refers to whether a measure “actually” measures what it is purported to measure. In the psychometric theory of measurement, a validity coefficient is the proportion of variance in observed test scores that is due to variance in the true score. Construct validity is especially important in foundation measurement. In foundation measurement, a model for the construct being measured—such as intelligence as the “g” factor—describes how the specific indicators being measured relate to the construct. The foundation model provides the theoretical basis for the correlation between specific indicators and a measure of the construct—that is, the measure is an indicator of the specific value of the foundation model, which itself is related to the “true” but usually unmeasured construct.

#### ➤ Assessing Validity

Several techniques have been proposed to assess the validity of measures. These techniques depend upon the type of research involved. However, some techniques are common to both types of research. In predictive research, validity is evaluated by estimating correlations between the measure in question and its predetermined outcomes. Such validity estimates can only be based on correlational designs. In experimental research, the outcomes of interest are made to differ according to the experimental manipulation, and, therefore, the measure under

consideration should produce systematically different scores across situations if it has some predictive validity. Experimental designs provide more convincing evidence concerning the validity of measures because they eliminate the possibility of untested causes that influence both cause and effect. However, further experimental evidence is required to establish the alternative outcomes of interest, which can be predicted by the measure in question. In both cases, assessing predictive validity may necessitate disrupting participants by presenting them with stimuli designed to influence their predicted responses, using techniques such as a mood induction or a priming method. Such disruptions should be envisaged only as a last resort, after responding to the predicted outcomes through naturalistic channels has proven unsuccessful. The second technical method for assessing validity is to compare the interscale or interitem correlations of the measure in question with the interscale or interitem correlations predicted theoretically by integrative theories and paradigms. Such methods may be seen as an alternative to testing the predictive relationships of interest. From the researcher's standpoint, however, validity is about predicting and thus, when possible, is best approached in that way. The third technical means for assessing validity is factor analytic assessments of the validity of the measures. In exploratory factor analytic methods, sets of measurement scales are exposed to a factor analysis that yields a structure of intercorrelations among the measures. If measures are to be seen as valid assessments of specific constructs, it will be found that measures of the same construct load predominantly on the same factors.

### **CREATING RELIABLE AND CONSISTENT RESEARCH INSTRUMENTS**

Research instruments can take different forms, depending on the nature and depth of the information being sought, the types of investigation being undertaken, and the resources and time available for the project. What is important, however, is that the researcher does not venture forth and hammer together a crude instrument, when much is at stake, simply because a tool does not seem to be "to hand." Research instruments, which may also be called measurement tools or measures, are devised purposefully to gather information that will help the researcher resolve the questions or issues being examined. Also, designing instruments requires a set of specialized skills that draws heavily on the review projects undertaken by the research team, as well as any prior work available in the environment. Questionnaires are the instruments that researchers design most frequently. And it is not surprising. Questionnaires come in a number of forms, are easy to administer, and can usher vast amounts of information into a research project in short timeframes. Questionnaires are self-paced instruments that can be completed in one of three ways: unassisted and unsupervised; face-to-

face with an enumerator; or over the telephone. They can be completed privately or in a group setting. Questionnaires are structured, as they incorporate predetermined questions. As a result, this type of data collection is inexpensive, easy to analyze, and straightforward to code. Questionnaires require explicit respondent answers. However, they are limited to closed questions that respondents answer according to choice, scaling systems that have been devised. To elaborate, choice responses are structured according to possible answer categories - in other words, categorical answers - that respondents select from: yes/no or true/false, and one of a number of possible answers. Scaling response systems ask respondents to rank, in some manner, the responses they are choosing between.

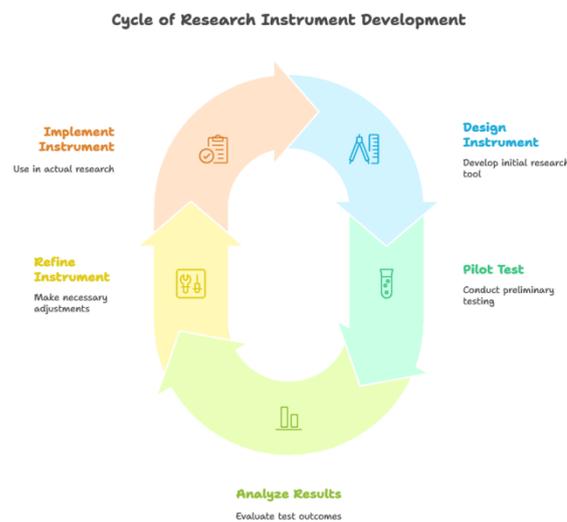


Figure 6.2 Research instruments

Prior to beginning the actual process of data collection, a researcher must develop a research instrument or procedure. A research instrument can take many forms in marketing management research. Traditionally, a research instrument was often a simple questionnaire containing open-ended and/or closed-form questions. With the advent of computer processing, some researchers also use computer applications to gather data—telephone and mail questionnaires are programmed and administered by using computer modems. The traditional paper-and-pencil questionnaire still remains a widely preferred method for gathering data in marketing research, primarily through personal and telephone interviews. In personal interviewing, interviewers personally administer the questionnaire, whereas in telephone interviewing, pre-programmed software and keypad responses facilitate data collection. This chapter provides some guidelines on how to create effective research instruments. The

principles discussed here are primarily for developing the traditional questionnaires used in personal, telephone, or mail interviews. The actual content of the questionnaire will depend on the information requirements of the study. Any computer-assisted interviewing application will, of course, follow the guidelines discussed herein. The principles discussed in this chapter apply mainly to questionnaire-based surveys; however, many of the principles apply to any type of data collection as well.

➤ Designing Questionnaires

While some authors devote entire volumes on the subject of questionnaire design, we outline some basic considerations in questionnaire design. The questionnaire should be designed to meet the research objectives and the information needs of the researcher. Hence a preliminary draught of the questionnaire should be drawn in order to link the individual variables to the specific questions. The wording of the questions should make it easy for respondents to understand the specific intent of the question, recall the relevant past experience accurately and formulate their response accurately. The visual layout of the questionnaire should be easy on the eyes and the respondents should be able to quickly grasp the intent of the question. While a question and the corresponding response should occupy no more than a single page of the questionnaire, the flow of the questionnaire should be logical and seamless.

➤ Pre-testing Instruments

When designing research instruments, researchers are keen to ensure that their questions work properly. By buttons press or letter sent, researchers can discover whether the instruments function correctly. Testing, like instrument designing, takes time and skill. To test successfully, researchers may consult specialists, recommendations, and texts. There are two basic approaches to study instruments: field testing and laboratory testing. Field testing requires availability of respondents and contacts, costs, and time. Laboratory testing does not require the research team to go to the respondents. The respondents are called to the testing company. There are other word-of-mouth pre-testing techniques. In-depth interviewing administrators may revise question wording, ask these pre-test respondents to itemize helpful and unhelpful questions, and ask what they think is about to be researched and how the study could be improved. Another type of pre-testing is cognitive interviewing, which gets at respondents' thinking. The interviewer asks respondents to think out loud or afterwards asks them questions about their thoughts. Additional behavioral pre-testing methods such as eye-tracking for web tests, focus groups, experience sampling, and think-aloud protocols should

be utilized. If the pre-test indicates that study questions need to be revised, there are various techniques researchers can use to improve the questions. Most of these question improvement techniques involve question wording or format changes. The most common class of question improvement methods involves changing question wording. This is the most common solution to a poor-performing question found during pre-testing. The validity and reliability of research instruments are key components for guaranteeing good design decisions.

➤ **Administering Instruments**

In light of all that has been discussed in the previous sections, actually administering the instrument is something that is left to the researcher. Among the things to be decided is whether to do a self-administration or not. A trained interviewer can add the greatest value to the answer by shaping it into the direction of aiding the decision-making process. A good interview can help clarify instructive questions, can probe into areas of deep interest for the study, and can help the researcher to listen beyond the words, and obtain the thoughts and feelings that shape the images portrayed in the interview. The interviewer can guide the process, helping the respondent feel at ease. The interviewer should be adept at asking questions that are open enough to get respondents thinking along the correct lines, yet not so vague that respondents start making assumptions. An important thing to do before the interview is set the purpose of the interview into perspective for the respondent so that the respondent feels at ease while being interviewed. If the respondent is not too happy to cooperate, the quality and quantity of information provided might not be very useful. If possible, the respondent should be familiarized with the contents of the interview and enabled to put forward his or her ideas before being asked direct questions. In a way, easing the respondent from general to specific questions will help the answers be more detailed and useful.

**Chapter summary**

Chapter 6, "Measurement and Scaling Techniques," focuses on the critical process of quantifying data for analysis. It begins by outlining the four fundamental levels of measurement: nominal (categorical), ordinal (ranking), interval (equal intervals between points, but no true zero), and ratio (equal intervals with a true zero point). Understanding these levels is crucial because it dictates the types of statistical analyses that can be appropriately applied to the data. The chapter emphasizes the importance of selecting the appropriate measurement level based on the nature of the variables being studied. The

chapter then delves into various scaling techniques used to create instruments for measuring constructs. It explores the Likert scale, a widely used method for measuring attitudes and opinions, along with the Guttman scale, which assesses cumulative scales of agreement. The semantic differential scale, which measures the meaning of concepts through bipolar adjectives, is also explained. The chapter provides guidance on choosing the most suitable scaling technique for specific research objectives, considering factors such as the complexity of the construct and the desired level of precision.

A significant portion of the chapter is dedicated to ensuring the reliability and consistency of research instruments. This involves discussing methods for assessing internal consistency (e.g., Cronbach's alpha) and test-retest reliability. The importance of pilot testing and refining instruments before full-scale data collection is highlighted. Strategies for minimizing measurement error and enhancing the validity of instruments are also explored, ensuring the accuracy and trustworthiness of the data collected. In conclusion, Chapter 6 emphasizes the meticulous process of developing and validating measurement instruments. It stresses the importance of selecting appropriate measurement levels and scaling techniques, and dedicates considerable attention to establishing the reliability and validity of these instruments to ensure the quality and integrity of research findings. The chapter equips researchers with the knowledge and tools necessary to create robust and reliable instruments for collecting meaningful and accurate data.

\*\*\*\*\*

## CHAPTER 7

### DATA ANALYSIS AND INTERPRETATION

Information is now being generated at an increasing rate through various systems. The computer systems perform many operations using multiplicative factors and there is a great requirement to analyze the results of these operations in a clear and coherent manner. There are various methods and techniques of data analysis and it is essential for a researcher to understand the need to apply the correct methods and techniques as required. In every research study, we collect a variety of information from a variety of sources. To put it precisely, we may have various answers to a particular question. If communication is the main part of research, that is, putting the results of the research and the information generated at the service of the public, it is essential that this is done with the aid of data. What happens when the data collected is too large and cannot be expressed with simple statistics like average or total percentage? Whether the data are large or small, it is evident that a given infinite set must be further subdivided into groups so that the data may be put to some proper use. Sorting the data is the next phase in the analysis. Data sorting refers to the arrangement of the data but should also reflect the aim of the data collection. Once we have the data sorted, we can find importance in one of the attributes of the data; that is, we can classify it according to some character. Further analysis of the data may be conducted depending on the attributes of importance, and data may be subclassified into additional attributes, as is necessary. Before we can determine the relevant inferences from the data and begin our hypothesis testing phase, it is important to understand the basic characteristics of the data. Finally, once the attributes of study have been established and sorted, one may examine the relationships and cause-and-effect that the different attributes reveal. This reason for such patterns be it normal or do we have outliers, can itself be the motivation generating further research questions and possibly leading to new hypotheses.

### ORGANIZING DATA USING CODING AND TABULATION

"Data organization is one of the primary objectives of any research project. Raw data lacks meaning until it is arranged and summarized. Once this has been accomplished, researchers can inspect the descriptions and see if they provide any patterns that suggest relationships. In addition, the data will be easier to analyze statistically when it has been transformed into a more meaningful and smaller set of numbers. Once the data reports consistent values, the next step is thinking of a coding scheme for the data. A code is a shorthand description of a

variable. In other words, information can be organized by placing into data files only the information needed to answer the research questions. Coding data is designed to limit the amount of data that is needed to be stored and analyzed, enhancing the interpretation of results. A variation of coding is applying weightings that assign values to discrete set levels of continuous variables or to categorical variables. Weighting can only be applied to variables that have ordinal properties. There are two traditional methods used in organizing data. Manual data organization techniques are labor intensive. Because of the technological advances of computer software programs, researchers today use statistical packages to organize data. Coding is a method of organizing and limiting data at the same time. For categorical data, coding is done by replacing the descriptive category title for a number. For discrete and continuous data, a coding scheme is developed to summarize the raw data into broader groups, thus eliminating repeated values."

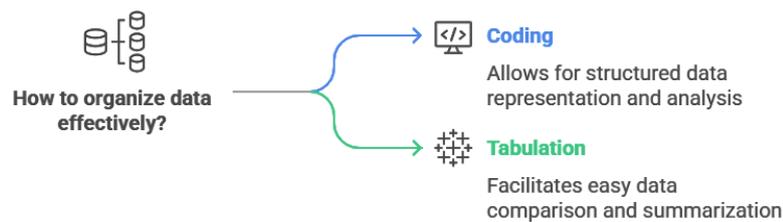


Figure 7.1 Data organization

### ➤ Coding Techniques

Coding is a mechanism for sorting data of similar types into similar categories, according to which the subsequent stages of analysis must be compatible. Coding means labeling units in a way that identifies them as being relevant to describe a characteristic of interest. It means also labeling all units in a way that makes it possible to compare them with one another because they belong to the same subcategory of that characteristic of interest. Coding creates a set of ordered categories for a particular variable. This representation must also be capable of being coded analytically with respect to the other variables present in the data file. Variables come in different forms. Some are simple, taking on only one specific value for a given case, and others are complex, such as sentences or paragraphs in a qualitative analysis where such variables may have a number of different interpretations for a given case. For simple variables, analysis can be performed after coding each of the values of the variable's limited number of categories. For complex variables, coding is done differently: either coding one

value for each group of words sharing the same interpretation, or identifying different codes for each identified interpretation. A max code assignment per interpretation is the easiest interpretation stage of a variable to analyze. The other coding process would involve number crunching on many more separate variables; hence analysis would require careful scrutiny of each subcategory separately. For a text variable, a group of words with the same field will usually follow the same semantic function, even expressing different syntactic functions, like the subject in the phrase “X is happy”. This means that their combined knowledge would allow creating a single code for that text variable, one per identified meaning.

#### ➤ Tabulation Methods

Tabulation is an essential method utilized in the field of research, capable of organizing large datasets into smaller, coherent, and readable sets for further analysis. It can increase accuracy, expedite the process, and reduce error risk in data analysis. Tabulation is the simplest method used in quantitative research, and it is typically combined with other statistical techniques for data explanation and interpretation. The simple tabulation may be described as an orderly arrangement of numbers, which shows the number of occurrences, usually by categories of the particular measurement for a given case. In simpler words, a simple table shows the frequency distribution of a single variable. Similarly, each group of the two categorized variable tables shows frequency distribution. The simple tabulation table is called a simple table, and two categorized frequency distribution is called a bivariate table. The frequency data may also be shown as a graphical representation like bar charts, histograms, etc. There are several types of tabulation based on complexity and consequences. A simple tabulation is the earliest stage of quantitative data analysis. If there is only one variable in the data, simple tabulation is directly organized to show the aggregated data frequency. If there are any two category variables into the data, bivariate bar charts, bivariate table charts, and bivariate frequency distribution tables can be prepared to present their frequency. Further, if there are any three categorical variables in the study, a three-way table is formed. There are further tabulations based on complexity and data readiness. Data may be in any form, like categorical, numerical, time-series, panel, etc.

### **DESCRIPTIVE STATISTICS: MEAN, MODE, MEDIAN, STANDARD DEVIATION**

Descriptive statistics are procedures that summarize or describe a set of data and provide simple numerical summaries about the sample and the measures. A combination of numerical and graphical techniques is sometimes used to describe a collection of quantitative data.

Because our main interest is in describing the participants involved in our study, only descriptive statistics relevant to the sample are presented. These include the mean, standard deviation, and frequency counts for categorical variables. Many of the research hypotheses test a relationship among variables. Descriptive statistics also summarize relationships among variables in a data set. Thus, they are used at the beginning of each chapter that tests a specific hypothesis throughout the text. Generally, mean and standard deviation summarize relationships among continuous variables, whereas frequency counts are used to summarize relationships among categorical variables and the mean and standard deviation among continuous variables. Graphical summaries often complement numerical summaries. In single-variable analyses, a histogram is useful for summarizing the distribution of the study variable. In multivariable analyses, scatterplots are useful for summarizing continuous-variable relationships.

The mean and standard deviation are the most widely used descriptive measures. The mean is a summary measure generally regarded as the best summary measure of the center of a distribution. It calculates the arithmetic average of a set of values. It is, however, less informative about the nature of the distribution than the mode and median. While the mode gives the most frequently measured value, the median is the value at the 50th percentile of the distribution. The mode is not commonly used to characterize continuous distributions, but it is useful to characterize categorical distributions when the categories have desirable properties distinguishing them essential to the test.

**Standard Deviation** The standard deviation is the most popular measure of variation. It provides a summary measure of the variation of the distribution. The same summary for sample proportion is the empirical rule, which states that approximately 68 percent of the data fall within one standard deviation of the mean, about 95 percent within two standard deviations, and about 99.7 percent within three standard deviations. The empirical rule applies only to distributions that are bell-shaped.

➤ Mean, Mode, and Median

Often, researchers cannot clearly interpret complex data or realize what it means. Descriptive statistics allow us to perform this properly. Collecting and analyzing the data taken as sample(s) can have so many possible outcomes. So, we need to summarize the results in such a way that reflects the sample's characteristics or provides insight into what it may have. Descriptive statistics involve methods for summarizing the data and interpreting it in simple

terms. For interpretation of the main fact of the study, it is better to describe it using graphics and tables. And, the description can be enhanced by estimation of statistics such as mean, standard deviation, mode, and median.

**Mean:** The most commonly used average is the arithmetic mean, which can be defined as the sum of a population divided by a number of individuals. The symbols  $n$  and  $N$  represent the number of values in a sample or the population. The mean is usually denoted by  $\bar{X}$  or  $\bar{X}$ . When it is for populations, we represent it as  $m$  or  $\mu$ . To calculate it: For its calculation, the values available were added. This sum would be divided by the number of terms. The mean has a quantitative effect. Each value contributes to its value. And unlike the other measures, even a small value at one end can pull the mean in that direction.

**Mode:** The mode is the most frequently occurring value in a population or sample. For example, if the values of a population are 1, 2, 2, 3, 3, 3, 4, 5, the mode is 3. The mode can be explained in another way: It is the number that appears most often in a data set. Unlike the median and mean, that can be also calculated for nominal data, it can be used only for the rest. The mode has many merits over the mean and median. When the data are normal, there are no modes. In a grouped distribution, usually is no unique mode. A multi-modal distribution can occur when the various modes exist in a variable.

#### ➤ Standard Deviation

Despite being a bit more complicated than the range and variance, the standard deviation is nonetheless the most useful measure of variability. The standard deviation is simply the square root of the variance. We'll denote the standard deviation of an entire population with the following Greek symbol:  $\sigma$  and the standard deviation of a sample with this symbol:  $s$ . With hypothesized parametric tests, in order to create sampling distributions that allow us to compute confidence intervals or test hypotheses, we must assume that the sampling distribution of the mean is either normal or approximately normal. Although it is common to assume that data are normally distributed, it is almost never the case that the data exactly meet that requirement. Yet the Central Limit Theorem guarantees that on average, for reasonably large samples, sampling distributions of the means are approximately normal. This means that if you took a random sample of scores from some normal distribution, whatever you choose the sample size to be, the mean of the sample will be normally distributed and thus have a known standard deviation. The standard deviation of that distribution is known as the standard error of the mean, and it is given by this formula:  $\sigma_{\bar{x}} =$

$\sigma/\sqrt{n}$ . If the population from which your sample is obtained is not normally distributed, the means of samples of 30 or more will, in the long run, have a normal distribution. The standard deviation of that distribution is obtained as follows:

### **INFERENCEAL TECHNIQUES: T-TEST, CHI-SQUARE, ANOVA, CORRELATION**

So far, we have discussed techniques used to describe and summarize the available data. What if we want to summarize our data findings and generalize them to a larger population? What if we observed some interesting differences between groups in our sample data, and we want to evaluate the statistical significance of those differences? In such cases, we can use inferential techniques. With inferential techniques, we can estimate and generalize the characteristics of a larger population from where our sample data was drawn. Inferential techniques are based on the probability theories of chance occurrence and on sampling. The t-test is a statistical test that helps us compare the means of two groups and determine if there is a statistical significance between the means of the two groups. The t-test allows us to determine whether two sample means are significantly different from each other. A t-test is a parametric test, and it is also commonly referred to as a type of t-test. Data must be continuous, the data should be randomly sampled from the population, and the data points are assumed to be independent. Chi-square tests are statistical tests that evaluate the difference between two or more groups of categorical data, and they help determine the significance of the difference between observed and expected frequencies in one or more categories.

#### ➤ t-Test

A t-test is a statistical test that is used to compare two groups or sample means and understand how far apart the groups are when taking into account the variability of the data and the sample sizes. It is popular for small sample sizes where the population variances are unknown but is also included in any sample size analysis and is widely used for any sample sizes. Like all the other statistical tests, there are many variations we can adjust according to the data we have and the questions we want to answer. All the t-tests work similarly, but they have different conditions of use and psychometric assumptions. The one-sample t-test lets you understand whether the mean of a single sample differs from a known or hypothesized mean. This difference could be statistically significant, for example, in a clinical trial where the novel treatment is tested against the usual one. The independent t-test calculates the difference between the means of two independent groups. It answers questions like whether two different means from different groups of people or animals are significantly different

from one another. This difference is usually an indicator of the treatments' different effects. The paired sample t-test works on paired data from two groups of people or animals. The interpretation of the t-test results is usually straightforward. In the case of independent or paired t-tests, a common output is a t-value and the associated p-value. The p-value indicates if the groups are different from each other at significance level 0.05. The t-value and p-value are typically presented in a table, along with the means and standard deviations of both groups. The accumulated information provides a great overview of the results and allows the user to determine whether any observed difference is meaningful for the study. In the case of t-tests on more than one sample index, the t-value and p-value are usually presented for one index at a time. Captions should provide insight into the comparisons performed within each plot.

#### ➤ Chi-Square Test

The goodness-of-fit Chi-square test is applied to assess whether the frequency distribution of a categorical variable is consistent with a previously expected distribution. In other words, it is used when you wish to determine if an observed frequency distribution differs from an expected frequency distribution. The degrees of freedom associated with the goodness-of-fit Chi-square test are equal to the number of categories minus one minus the number of estimated parameters.

For example, suppose that the 10 non-zero digits were to appear approximately uniformly in the last 10 digits of the Social Security numbers of persons in the United States. With respect to the goodness-of-fit Chi-square test computations, you could, for example, select a random sample of 50 Social Security numbers and tally the digits representing the last position in the passwords assigned to those persons into 10 categories: one for each digit value 0, 1, 2, ..., or 9. You would expect each digit to occur in the last position 5 times. The resulting observed and expected frequency distributions would then differ. The probability associated with the test of interest indicates that the observed distribution is detected to differ from uniformity at the 5% level and that the null hypothesis is rejected.

The Rao-Scott Chi-square test is similar to the Goodman-Kruskal Gamma statistic but applies to frequency tables having a greater number of categories. It decomposes the table into a number of simpler components—a separate Chi-square test for each diagonal cell summed to yield an overall Chi-square test. Care must be taken when interpreting the results. Each subtest assesses whether the diagonally opposite cells are in agreement. The consolidating

Rao-Scott version assesses whether the individual components differ from agreement. Applying this approach assumes that the agreement proportions differ while the disagreement proportions do not.

#### ➤ ANOVA

ANOVA (Analysis of Variance) is more powerful than t-tests when more than two means have to be compared: instead of making a series of pairwise comparisons as done by t-tests, ANOVA takes all groups and compares them at the same time. Using the logic of t-tests, if two comparisons are made, the chances of making an error is twice as great (if  $\alpha$  is the probability of making an error on one test, then the probability is  $2\alpha$  for making an error on either one of the tests). If “k” comparisons were made, then the chance of making an error is “ $k\alpha$ ”. Because a larger alpha means that the experiment is less sensitive to finding differences, instead of using the typical alpha of .05, a correction is made and use an alpha that is smaller by a factor of k. If k is large (i.e., if there are a lot of tests to determine whether means are different), then the chance of making an error is large; consequently, a small probability ( $\alpha$ ) for an individual test is necessary. This is the basis for the Bonferroni adjustment. However, this adjustment has been criticized since if an individual test’s p is small enough to make an error, it is not very interesting. Because of its inefficiencies, performing k tests is not a good strategy when k is large. Therefore, ANOVA was created to solve this problem of multiples comparisons: ANOVA compares all the means at the same time using an experimentwise alpha level of  $\alpha$ , while the extension of the Bonferroni adjustment for all the t-tests that would be performed is not trying out, only the smallest mean difference is used. If t-tests were performed, adjustments would need to be made to correct for the multiple tests being performed.

#### ➤ Correlation Analysis

Correlation analysis is concerned with measuring the strength of the relationship between two or more variables. The word 'correlation' is derived from the Latin word 'correlare', meaning "to relate". In research methodology, correlation is defined as the degree of relationship between two variables. Thus, correlation is a measure of the intensity of the association between two or more variables. Correlation analysis is a key area of multivariate analysis. Various statistical techniques are used for calculating correlation. The most common and widely used technique for assessing correlation is Pearson's correlation coefficient, though some use other measures like the Spearman rank correlation coefficient or other measures

designed for a specific purpose or to overcome the limitations of Pearson's correlation coefficient. Usually, correlation analysis is used with cross-sectional data, but it can also be used with time series data.

The most fundamental property of correlation analysis is that it can only confirm a relationship; it cannot prove or disprove a relationship. A correlation coefficient of 0.6 would indicate a fairly solid relationship, while a correlation coefficient of 0 would indicate that no relationship exists. Correlation theory is based on the premise that with a large sample size, the significances of correlation coefficients may be attained. For the most part, we agree that 25% of total variance explained is negligible, 25-49% is a small contribution, 50-69% is a medium contribution, and 70%+ is a large contribution. Correlation analysis requires that the variables be multidimensional and continuous. If the correlation analysis is severe, there are measures to control for it. Also, factor analysis and structural equation modeling also accomplish this objective.

## **DATA ANALYSIS USING SPSS, EXCEL, AND R (BASIC COMMANDS AND CHARTS)**

This chapter presents some of the most commonly used tools for data analysis. There are many proprietary or open source tools for data analysis, and quite a few are available from within popular programs such as word processors or spreadsheets. These tools form only a part of the analysis options that exist: some researchers prefer to set up their own analysis tools based on the structure of their data and the kinds of outputs they want to create. Programmatic solutions using programming languages with data analysis or statistical functions may also be appropriate, and most surveys have their own company-provided reporting formats available. This chapter gives basic syntax and/or basic capabilities for four popular analysis tools: SPSS, Excel, R, and Power BI. These are presented in that order because they are, in general, increasingly less powerful as analysis programs. SPSS is an analysis program, while Excel is a spreadsheet program which has some capabilities for data analysis. R is a tool for programming and data structure manipulation which has robust community-built packages for data analysis, and again, Power BI is primarily for business analytics, offering limited analysis capability. Because these programs tend to be covered very well in their help systems, this chapter covers only basic pieces of functionality for data analysis. For new users, exploring optimization questions using the help functions is the recommended approach, or else finding someone to teach you the basics.

➤ SPSS: Basic Commands and Charts

The Statistical Package for the Social Sciences (SPSS) software was developed as a stand-alone package to handle large scale social science data. Today, it has evolved as the most popular scientific data analysis software, with extensive capabilities not only for statistical analysis but also for data management and graphics. The salient features of SPSS include a wide range of established test and data manipulation modules; the ability to manage data files (including combining and merging separate data files); an ability to create or modify new variables; backward and forward compatibility with other major statistical packages; the ability to automate repetitive commands by creating macros; implemented graphical options; a wide range of graphics for graphical presentation and exploratory analysis, including building innovative graphics; and facilities to export and publish output in a variety of formats. The basic SPSS window has what is called a 'Data View', similar to a data sheet in Excel offer grid-like entry of data, with variable names in columns and cases in rows. In addition, SPSS file types are '.sav' files, which can be opened by SPSS. Once you install and open SPSS, you have the option to save your work at any point. A new session can be opened by selecting the file option 'New'. You can create one or more '.sav' files. In addition to creating new files, SPSS can read Excel '.xls' or '.xlsx' files. This is useful for running analyses on Excel files that were created by other users. You do not have to create a new file or input your data into SPSS from scratch.

➤ Excel: Data Analysis Features

Excel provides numerous tools for data analysis. The Data Analysis Tools option in the Tools menu includes such techniques as ANOVA, Correlation, Covariance, Exponential Smoothing, Fourier Analysis, F-test, Histogram, Moving Average, Random Number Generation, Regression, t-test, and Variance. To access these options, one must first install several Excel add-ins, specifically Analysis ToolPak and Analysis ToolPak-VBA add-ins. Most of the above commands are available through the Data, Data Analysis menu option. Users have to select one of the option commands available. When the desired technique is selected, a prompt screen will lead the user through the process of analysis by requesting required input information such as measurement level, whether labels are included, output range, etc. While the commands described are certainly valuable, many of Excel's extensive features enable researchers to develop their own analysis of data. Creating new commands or customizing existing ones developed by Microsoft or other individuals using a Visual Basic module can

also be done. Selecting Help from the menu bar provides an overview of these features, along with step-by-step instructions for using them to develop custom-prepared statistical features. Statistical analysis that can be accomplished through Excel includes measures of central tendency, variation, linear regression analysis, and probability distributions. One can also compute binomial cumulative distribution polynomial, exponential fade-point, negative binomial cumulative distribution polynomial, Poisson cumulative distribution polynomial, Student's t cumulative distribution polynomial, and Z-score. Other probabilities that can be computed include probability of standard normal deviate falling between two specified values and probability values associated with the binominal, negative binomial, Poisson, Student's t, and Z distributions. Excel also allows users to develop two and three-dimensional data visualization charts.

➤ R: Introduction to Commands

R is a powerful free software package used widely for data analysis. Unlike other similar packages, however, it is not interactive. This means that commands directing R to perform certain actions must be typed in directly and the results will be displayed usually on the screen and the files directed to a specified location or into a workspace. Although the learning curve is a bit steep, R has become the standard tool in applied statistical research, in subtle or complicated manipulations of data, or when researchers need to develop new statistical methods or tools. The need for budgetary economies has also led to the growing popularity of R. R, unlike the commercial packages, is specifically a scripting language. It is usually run in a window with no graphical menus or buttons. Graphics can be output in separate windows, or redirected to files in various formats. Typing commands into a console window one line at a time is generally the simplest approach for quick exploratory analyses and generating graphics. These commands can also be collected and run in batches from a script file. Although adding R commands to an existing script to be read and run in batch mode can be inefficient, it is often quicker and more convenient to test commands interactively as you work, and then collect them later into a script. R is based on the S language, originally developed at Bell Laboratories. S has a long and proud history in statistical computing. S is responsible for popularizing a number of important ideas. It was one of the first languages to make it easy to create high-quality graphics. It was the first statistical computing system to allow users to extend the language. S was one of the first languages to provide complex, multi-module, object-oriented paradigms.

## INTERPRETING OUTPUT IN THE CONTEXT OF RESEARCH QUESTIONS

In this chapter, we will discuss ways to interpret the output from MoDIS and other analytical methods of research. The more focused the interpretive criteria, the more likely it is that conclusions drawn from data will be relevant and meaningful. For this reason, prior to describing our interpretive process, we will illustrate the way that interpretive criteria can vary according to the nature of the research questions being studied. Research questions can differ dramatically in terms of scope, the manner in which they describe the subject of study and the neighbor roles within the model. If research interest at time is ‘what are these data indicative of vis-a-vis the research questions posed?’ we have undertaken associative research. If interest is ‘what are these data telling us about how to interpret other variables?’, an interpretive inquiry is being undertaken. If the intent is to compare to other groups or sets of variables, an explanatory inquiry is underway, with data generally being used to determine the role being considered. Simple as this classification may be, it makes a huge difference to the interpretive criteria. Associative research relates to the confirming, exploring or revealing of relationships; explanatory research aims to compare or determine direct relationships; and interpretive research deals with the interpreting of roles. This chapter's interpretive process is specifically geared to qualitative inquiry and it is assumed that any model will by need be simple, with no more than a handful of neighbor roles.

### ➤ Contextualizing Research Questions

In the beginning, the goal of the research question is to reduce uncertainty and allow for prediction of results. This, however, is not how data analysis generally works. Instead, data produces results which are raw and unexplained. Their un-contextualized brute probability is where we start. Applying sampling techniques to calculate the distribution of the output values allows for a base interpretation. The reduction of uncertainty can now be framed as a simple question – are our predicted distribution of results different from the un-contextualized brute data results? This is done using statistical methods. If they are different, that is something interesting. Something has explained the variation. However, the problem is that using only perceived different distributions can raise false flags. All data is noisy and can trigger things that are not real. That being said, sometimes answers do come without coercion when using big data methods. For example, one big data study on text found that the number of nouns in the text would generally rise as the predicted difficulty of the text rose. In the case of perception competency measure, it explained a percentage of the big data pattern.

This result was however indirect and remained unexplained in the study. Explaining variation implies some other measure of predictability use. At this point, we can put the plain data aside, if we want to. In the ideal case, we can proudly understand why our data produces the kind of effect that occurs seen at the start of analysis. However, this direct data analysis is dependent on usually Small N effect measure. Here, the focus goes back on the brunt data. The job of numerosity again is to be predictive.

➤ Drawing Conclusions from Data

The data output does not answer the research question of a study. It simply relates to it. The research questions were first drawn up by the investigators that drew up the study. The output data never answers in a structured way; it is against what science is. Science needs to be built. Only when researchers contextualize the data output and interpret it, does it answer the study's research question. Then investigators should try to explain the data output in a manner that was not expressed with data. Researchers generally contextualize the data output within what is known about the subject. The literature review is critical in order for scientists to understand how to interpret the data output. Only after comprehensively understanding the known information on the subject may investigators draw different conclusions in wealthier and richer hypotheses from that presented by the known literature. To generalize new hypotheses and predict the future, we must correlate the study data output with those from the literature. Data input means those data acquired with the method utilized that lead to the results. Data output is always relative. We humans were not blown out by this universe without a comparative process. Data output must be as understandable as possible by the common sense of the majority of people affected and intimate to those studied, or at least seem to be. If the population from which the study objects are taken distrust the argument and do not relate to the final output, they will generally refuse it and consider it invalid. To put less technically, researchers must try to convince the involved population that new information is useful; otherwise, research will not help.

### **Chapter summary**

It begins by outlining essential data organization techniques, including coding (assigning numerical values to categorical data) and tabulation (summarizing data in tables). This structured approach lays the foundation for subsequent analysis and interpretation. The chapter then introduces descriptive statistics, focusing on measures of central tendency (mean, median, mode) and dispersion (standard deviation). These techniques provide a concise

summary of the data's characteristics, offering an initial understanding of its distribution and key features. This section emphasizes the importance of selecting appropriate descriptive statistics based on the data's measurement level and distribution.

Building on descriptive statistics, the chapter delves into inferential statistical techniques. It covers commonly used methods such as the t-test (comparing means of two groups), chi-square test (analyzing categorical data), ANOVA (comparing means of three or more groups), and correlation analysis (measuring the relationship between two variables). The chapter explains the underlying principles of each test and their appropriate applications. Finally, Chapter 7 provides a practical overview of data analysis software. It offers a basic introduction to using SPSS, Excel, and R, showcasing common commands and chart creation techniques for data visualization and analysis. The chapter concludes by emphasizing the crucial step of interpreting the statistical output within the broader context of the original research questions, ensuring that the findings accurately reflect the study's aims and objectives.

\*\*\*\*\*

## CHAPTER 8

### QUALITATIVE RESEARCH METHODS

Qualitative research methods are of growing interest in business research as business becomes increasingly customer-focused. Many writers, however, have a critical approach to qualitative research. They ask how qualitative insights can be integrated into the quantitative models by which the managerial mind navigates. Others ask to what extent qualitative research is systematic, rigorous, and controlled, and with clear methodological guidelines. These are questions directed at business qualitative research in general. However, the treatment of qualitative research methods in business research textbooks is predictably uncritical. Such textbooks convey the impression that qualitative research can be conducted on a shoestring, and may provide surface observations of questionable depth, and yet provide equally valid answers as do the scientifically rigorous formal hypothesis-testing methods generally regarded as quantitative.

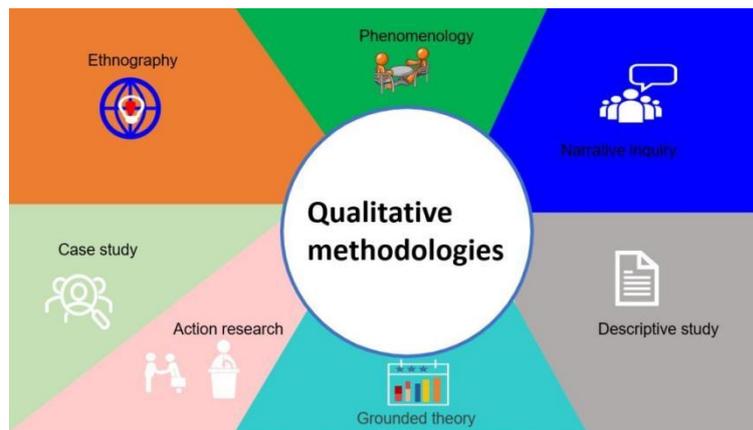


Figure 8.1 Qualitative research

### NATURE OF QUALITATIVE RESEARCH AND ITS APPLICATIONS

Qualitative research is the term given to research which produces findings not arrived at by statistical procedures or other means of quantification. Qualitative research, and the qualitative methods that go into it, have been around since researchers wanted to study what people do, study it in detail, and give implications based on that detailed study. A qualitative researcher begins research with an idea of how to follow a lead about a social phenomenon that also engages his or her interest. The researcher does not begin with a theory, and does not test a hypothesis, but does formulate a set of specific hypotheses as the research progresses, because a good qualitative study usually does have to do with making progress on

social theories. When qualitative research is able to put together a messy reality involving human beings that had previously escaped the notice of others, people pay attention. Qualitative research has a history of calling the attention of the social scientists to the small things that can make big differences. Qualitative research asks what is happening in a situation. What are the meanings people attach to what they do? What is the interplay of the different social influences? It is to qualitative research that colleagues turn for accounts of the variety of contexts in which human beings operate. All of social science has a qualitative side. The research methodology of choice is usually whichever one is most appropriate for the research question. In the best traditions of social science, qualitative research is problem-driven; it draws on methods and designs on offer from diverse disciplines. A wide range of methods can be used and adapted for qualitative research: interviews, focus groups, ethnography, open-ended surveys, naturally-occurring group talk, action research as well as designed and implemented social experiments. Likewise, qualitative research is frequently mixed with quantitative research. Unlike the processes of hypothesis testing favoured by quantitative researchers, qualitative research aims for a more holistic or narrative account of the complexity of behaviour in context.

### **Applications of Qualitative Research**

Qualitative research methods provide researchers ability to study complex phenomena in their natural settings. In many different disciplines, qualitative research has been employed in the following ways. One application of qualitative research is to support and enrich quantitative research. Using qualitative methods in conjunction with quantitative methods can reduce the chances of erroneously interpreting previous quantitative findings. Another approach is to use qualitative research to explore topics that quantitative research has failed to fully address. In this case, qualitative research can help ensure that the selected quantitative method options will be relevant and the selected quantitative measures will be appropriately utilized. A third application of qualitative research is to use it as a precursor to quantitative research, using qualitative interviews or focus groups to identify item response processes or capture the breadth of topic areas to be measured. The fourth application is to use qualitative research to inform the design of a new measure. This includes constructing question or item wording, determining the length of the measure, and determining how the measure will be implemented. Researchers may also choose to conduct a qualitative study for the primary purpose of generating hypotheses or questions to be tested in subsequent quantitative research. In addition to serving as a precursor to quantitative research, qualitative research may also

serve as a follow-up to quantitative research, attempting to understand the reasons for results that have been uncovered in quantitative studies. Another approach is to use qualitative methods to conduct validation research, providing evidence supporting the validity of a specific measure or assessing its relevance to the construct or population in question. It is also possible that the researcher may not be conducting multi-method research, but may instead be conducting a qualitative study that has as its sole purpose to develop an understanding of the participant's experience. For these studies the development of a new product or longer term, are not involved.

### **METHODS: INTERVIEWS, FOCUS GROUPS, CASE STUDY, CONTENT ANALYSIS**

Qualitative research is exploratory research that aims to study the various phenomena encountered in social life through in-depth analysis to reveal the nature and essential properties of these phenomena. It is carried out using non-statistical data collection and analysis techniques and methods. The goal is to understand how people think and feel, explore their understandings, share their experiences, and understand the meaning of or interpret the events in their lives. To understand people's behavior, researchers use qualitative methods to ask open-ended questions, or they conduct interviews, focus groups, case studies, or textual analysis. They analyze the transcripts to identify important concepts and frequent patterns, resulting in a rich description. The purpose of qualitative research is to uncover factors that affect how people feel about a particular situation. The specific result of qualitative research is the identification of factors that influence the attitudes and behavior of a population or an exploration of ways to find those factors. Qualitative research can also help to improve the process of collecting quantitative data and increase the response rates by easing the accompanying burden. This can be done by contributing to a better formulation of questions, thus helping to ensure that the questions are understood and the response options are appropriate. An additional qualitative step can be included in the area of cognitive assessment to provide cognitive interview techniques and focus groups.

#### ➤ Interviews

Qualitative interviews, akin to informal conversations, delve into participants' perspectives, experiences, and feelings, advancing their understanding rather than merely testing preconceived theories. Conversations typically take more time than mere discussions – anywhere from a quarter hour to many hours, even days. The interviewer usually asks many

more questions than the interlocutor answers but in response to the interviewee's own utterances rather than a fixed plan. Of course, interviewers often have a skeleton structure interspersing open questions with more structured – though not mechanistic – prompts to maintain focus on the broader social or cultural issues that interest the interviewer. Importantly, interviewers are willing to explore asides – unplanned turns in a conversation – if they seem especially revealing of the person's individual or social identity. Instead of one or a handful of researchers using an instrument to harvest data from hundreds or thousands of participants, qualitative interviewing tends to collect in-depth responses from no more than a few dozen. Interviewers make special use of the deep knowledge they develop with conversationalists to facilitate interactions that elicit multiple layers of understanding about the subjects' personal worlds and the meanings of their words and actions. Interview data generally present multiple layers of interpretation. In addition to quoting informants verbatim, qualitative researchers often draw on their own experiences, both meta-level and interpretive. Interview material may also be used to supplement more traditional quantitative content analysis of large numbers of survey or administrative data coded into defined categories. Indeed, qualitative interviewing is often performed repeatedly over time on the same individuals so that researchers can track changes over their lives as well as critically interrogate how and why they perceive events or conditions.

#### ➤ Focus Groups

In a standard focus group, a group of 6–10 participants who have been selected for their common characteristics sit together (usually around a table) and discuss a particular topic of interest for around 1–2 hours. The session is guided by a trained moderator, who is also present. Qualitative data are generated from both the interactions among the participants and the additional comments provided by the moderator. Since discussions focus on the opinions and experiences of the respondents, and since group discussions typically lead participants to expand and elaborate on their responses (and even challenge each other's views in a polite and respectful manner), the qualitative data available from focus group sessions is usually extensive. Focus groups are predominantly used to gather exploratory data on a topic in the public domain and private lives of the respondents. Focus groups originated in market research as a way to evaluate potential new products, concepts, or services. However, it was soon recognized that their interactive format made them a particularly useful qualitative method for understanding a variety of social phenomena. The social interaction of the respondents, or their group dynamics, can shed light on how the respondent perceives

something. Moreover, by tapping into the shared experiences and understandings between the group members, a focus group can produce discussion about actions, motivations, and feelings that may not arise from interviews. Because of these advantages, focus groups have increasingly been embraced by researchers from a wide range of disciplines inside and outside the social sciences including education, politics, health, and marketing. Focus groups are frequently carried out before interviews, which are considered to be the primary method. They are also used first to identify themes or concepts for subsequent individual interviews.

➤ Case Study

The case study design in qualitative research has been receiving increasing attention. More recent developments such as new family approaches to field territories, new strategies for hierarchical ordering of data, and the new turn towards other forms of research allow for new perspectives on the usually unstructured and unduly holistic nature of a case study approach. In particular, case studies that follow a holistic approach deliver little value that could not be collected through probability-based large sample surveys or ethnographies. Despite this, the case study approach to inquiry is being classified as either qualitative, quantitative or as both. The quality of research findings is often determined by the complexity of the underlying research process. A one-off examination is a frequent shortcoming in case study inquiries. A related issue is that existing theoretical perspectives are not used as a starting point for developing meaning to the individual case study. This leads to several specific techniques, principles, or overarching rules for developing case study inquiries. Despite its stated advantages and often promising possibilities, the case study approach remains shrouded by far greater levels of intellectual division in its appreciation than the field's main competing strategy. Nevertheless, there seems little doubt that case studies provide a unique opportunity for many international business scholars to take qualitative research seriously. The open-ended nature of many case studies resembles that of ethnographies. However, unlike ethnographies, case studies normally focus on specific research questions.

➤ Content Analysis

Content analysis entails collecting pieces of text, audio, or visual materials, to code and analyze their content. This method is well suited for addressing research questions such as “what are research papers on electric vehicles’ charging stations discussing?” and “how are various existential risks addressed by the media?”. Content analysis is a simple method for answering such research questions by discussing publicly available visual, written, or audio

materials. Content analysis is about documents. What you see is what you get. One is quite simply dealing with captured words, numbers, or symbols, the representations, constitutive elements of human activity. These representations and the interactions that produced them have been captured on tangible, permanent documents. This permanence means that the researcher is not present at the interactions that produced the data. This absence requires that the researcher infer from the captured documents the underlying intentions and, more generally, the answers to the research questions posed. Hence, content analysis can address most research questions and has many applications, especially when using new tools for networked visualization of text, or script processing, but at the expense of the deep connection that qualitative research via interviews, case study, or focus groups, create between researchers and the subjects.

### **MANUAL CODING AND SOFTWARE LIKE NVIVO OR ATLAS.TI**

Qualitative data analysis methods are organized within two categories: qualitative data manual coding and qualitative data software coding. Manual Coding, what we might call “qualitative data coding,” entails printing out all of our qualitative data (or hanging it on the walls) and analyzing it “by hand”: underlining or highlighting it, jotting notes on it, making margin comments, etc. We are producing a rough form of qualitative content analysis through this process: essentially tagging, track-marking, and categorizing the data that comprise our study. Software coding, what we might call qualitative data computer-assistant coding, takes on the mechanical work of qualitative data manual coding for us. Rather than the researcher engaging in every aspect of the qualitative data coding process—reading through and marking up the text, transcripts, program logs, and conversations over and over again—software does the “retrieval” tasks for us. Manual coding cannot be done on large segments of qualitative data; it is better suited to smaller qualitative datasets or for certain stages of larger qualitative data analysis processes. When looking to code one set of data or one set of coding categories several times in order to analyze different aspects of the data or to draw out different types of themes, codes, or units of analysis, manual coding is preferable. In small dataset analysis, a researcher might be able to analyze and comment upon each and every line of the qualitative text. The extreme level of nuance captures the data analysis within manual coding. It is very possible to write up rich qualitative texts as a result. Software programs take on the mechanical tasks involved in qualitative data analysis: storing the dataset, providing easy retrieval functionality using coding categories, generating frequency lists, generating and analyzing word and phrase distribution lists, and querying other aspects of the data. For

qualitative researchers coding by computer, qualitative analysis software is used as an expensive filing cabinet or elaborate index card catalog. Such programs do not take away the challenges and knowledge associated with analyzing qualitative data using qualitative content analysis methods. Programs such as these are primarily tools to store the qualitative data, facilitate data searching and retrieval, and do basic qualitative data coding work.

➤ Manual Coding

Choosing a method of coding depends largely on the research objectives. Generally speaking, we can code for the following reasons: 1. To index transcripts. Like indexing a book to provide a logical way back to the text, coding can help researchers build a framework for the analysis. This is helpful when there are multiple researchers working on the analysis. This can also help guide further analysis, e.g., finding negative cases. 2. To provide commented line-by-line analysis. Coding "at the line" level can provide a more thorough text analysis. 3. To guide analysis and reporting. Using a hierarchical system of codes can help structure a data analysis report or guide the writing of a theory-building paper. 4. To prompt questions about the text. Coding can also allow researchers to see repeated key ideas in their data, but prompt questions about those key ideas. 5. To facilitate flow analysis or capture structure in the data. The same types of flow structure that one would consider coding for in quantitative work can also guide qualitative coding decisions. 6. To extract quotes for reporting. Codes can aid in quickly scanning for a short passage to confirm other analysis conclusions. Codes can also help locate a key passage that can be quoted in an analysis report.

➤ Software Tools (NVivo, Atlas.ti)

Qualitative analysis software tools are widely used to support the qualitative analysis of data, such as interviews, articles, or online reviews, particularly for large datasets. They are designed to help with the tasks of data management, annotation, coding, and retrieval, but they do not support the analytical work, which is essential to qualitative research. In addition, these software tools offer rich functions to produce various visual representations. They can be used effectively for qualitative research, if researchers are aware of their strengths and weaknesses, and approach them with the appropriate theoretical and methodological backgrounds. These tools are designed to aid the researcher in the analysis of qualitative data, including open or moderated interviews, blog posts, articles, or online reviews. Qualitative software tools can be useful for qualitative research when the analysis process is well thought out. A fundamental aspect of qualitative analysis that software tools do not aid with is the

way the analytical insights are created. In qualitative research, the work of finding categories and establishing hierarchies is part of the analysis. Specialized software tools are not needed for this – familiar tools such as Post-it notes, scissors, and paper for categorizing are sufficient, and the use of colours facilitates the visual work at this stage in print form. Categorization and hierarchization, typically for printing multiple copies of the documents, may be followed by the gradual transition to a summary by word or paragraph, grouped according to the codes/concepts/categories chosen by the researchers: tables, with a combination of concepts and cases providing the guide, ordered lists of documents with the highlights for each chosen concept, and word and character clouds, these designed to facilitate visualization of the content resulting from analysis. Who uses which approach depends mainly on habit, but this offers no real advantages. In contrast, the later connecting of the findings to the words used by participants typically requires the tools' relative ease of access to the transcripts and their use for comparative purposes, tasks for which paper and scissors could be sufficient.

### **ENSURING TRUSTWORTHINESS: CREDIBILITY, TRANSFERABILITY, DEPENDABILITY**

Trustworthiness is the overarching effect of the different criteria that judges the quality of qualitative research. In qualitative research, the terms validation and verification are not usually used. When trustworthiness is being considered, qualitative researchers look to the concepts of credibility, transferability, dependability and confirmability. Credibility is the degree to which the researcher has established confidence that the findings, whether described as themes or categories, reflect the meaning participants ascribe to the phenomenon under study. In other words, credible qualitative research describes the multiple realities of social life and suggests that qualitative researcher has done adequate work to understand those realities and ensure that they are reflected in their findings. In qualitative research, external validity is more commonly referred to as transferability. Transferability is not the responsibility of the researcher; rather, it is the responsibility of the consumer of the research to determine the applicability of findings. Because small qualitative research samples are specific in nature, care should be taken in judging the transferability of research findings. Some researchers argue that semi-structured interviews can give us thick descriptions that facilitate transferability. Thick description is when we take the time and care to provide enough detailed description of the context of the research and the elements of the participants' culture that the reader can make an informed decision about transferability. Dependability

also has similarities to reliability in quantitative research; however, when it comes to confirmability, the qualitative researcher should be able to demonstrate the degree to which the findings of the research are a function solely of the participants' experiences and are not a reflection of the biases, motivations or interests of the researcher. Dependability is not simply a matter of checking for findings' consistency.

➤ **Credibility**

The dependence on inquirer skills and disposal of adequate resources as regards time, money and research tools does not allow extensive qualitative inquiries on large samples. Moreover, quantitative measures do not usually express the meanings of qualitative events. Since then, such events have been described employing qualitative methods that are aimed at exploring the perspectives of the groups being studied. It is not enough to note that the study cases are limited in quantity and quality. The selected cases need to be representative of unusual condition in qualitatively varied but intellectually comparable cases that allow deep analysis. Although the aim of the qualitative study is not generalizability, it is assured in terms of the global understanding of similar problems extending to comparable contexts and populations. By observing different perspectives of different cases, contrastive analysis can incorporate and enrich the imaginative accounting of the subjective experience of individuals. The researcher cannot be useful in accounting these subjective experiences if it does not audit the formulation of theoretical beliefs that have semantically and emotionally constructed these lived phenomena. It is about establishing the credibility of the study in the same way as the researcher needs to give account to itself of the evaluation of the subjective experience to locate the intersubjective reality. However, the complexities of individuals' constructions cannot limit, but can detail the problematic hypotheses that determined the conception of the research project.

➤ **Transferability**

The second criterion contributing to trustworthiness is transferability. This refers to whether the findings are valid in other contexts. Since qualitative research usually occurs on a small scale, findings that emerged from just a few participants do not automatically apply to the wider population. In fact, idiographic approaches aim to present a detailed analysis of particular cases, and it is not their intention to generalize the findings. Qualitative researchers, therefore, should provide a thorough description of their data sets in terms of context, as well as inform the reader about factors such as the number of participant samples with

demographic and other information in detail. This is critical for transferability, as the reader can then make informed decisions about similarities or discrepancies with their own context. If there was a clear sense of context and culture in the examples presented earlier, it would help the reader. Since qualitative research is generally context-bound, readers familiar with that context can make the connection and see whether or why the findings could hold true elsewhere. In addition, the researcher can provide a thick description of participant experiences in a different culture. This gives both the reader and researcher a general understanding of the potential transferability of the findings. Nonetheless, qualitative researchers should accept the burden of trying to demonstrate or argue that the data would make sense in different contexts. This can be done through illustrating how the results would be useful outside the original context.

➤ Dependability

Dependability is considered the third criterion of qualitative quality requirements. "Dependability" concerns the stability of the research's findings over time. It has to be guaranteed that if a second researcher would follow the same procedures as defined in the original study, he or she would also reach the same conclusion concerning the results. The actual way to ensure the aspect of dependability has risen to growing criticism. Many authors see the tools and constructs for enhancing the dependability of qualitative research difficult to respond to this request for more robust and sophisticated qualitative research designs. These conditions of subjective interpretation are inherent to the method, however, and therefore beyond the control of the researcher, regardless of how talented he or she is at executing the investigation. Critics demand an accounting of the constant subjective interaction between the researcher and the research participant that characterizes qualitative inquiry and influences the generation of results. For this reason, recently other methods have been proposed to ensure the trustworthiness of a qualitative research study since its design query construction. An explanation of one's theoretical position and also the aim of the study seems to be important to clarify the reader about some inherent biases that influence the interpretation and the reach of the research project. Several measures have been proposed to enhance the dependability and confirmability of qualitative research. Similar to the discussion for credibility, dependability can be achieved if dense, rich and contextualized data are generated, stored and analyzed. Respondent validation and using multiple sources of evidence are also methods to ensure dependability of qualitative research.

## Chapter summary

"Qualitative Research Methods," explores the distinctive nature of qualitative research, contrasting it with quantitative approaches and highlighting its suitability for in-depth understanding of complex social phenomena. It emphasizes the importance of exploring rich, nuanced data to gain insights into meanings, experiences, and perspectives, rather than solely focusing on numerical data. The chapter details the varied applications of qualitative research across disciplines, from sociology and anthropology to business and healthcare, showcasing its versatility in addressing a wide range of research questions. The core of the chapter focuses on various qualitative data collection methods. It provides a detailed overview of common techniques, including individual interviews (structured, semi-structured, and unstructured), focus group discussions, case studies, and content analysis. Each method is discussed with its strengths and limitations, guiding readers in selecting the most appropriate approach based on their research objectives and context. Practical considerations, such as developing interview guides and facilitating effective focus group discussions, are also addressed.

Data analysis in qualitative research is examined, covering both manual coding techniques— involving the systematic organization and categorization of textual data—and the use of specialized qualitative data analysis software such as NVivo and Atlas.ti. The chapter explains how these tools can assist in managing large datasets, identifying patterns, and building coherent narratives from qualitative data. It highlights the importance of choosing appropriate software based on project size and complexity, and emphasizes the iterative nature of qualitative data analysis. Finally, the chapter addresses the critical issue of ensuring trustworthiness in qualitative research. It explores key criteria for evaluating the quality and rigor of qualitative studies, including credibility (internal validity), transferability (external validity or generalizability), dependability (reliability), and confirmability (objectivity). The chapter emphasizes the importance of employing strategies to enhance trustworthiness throughout the research process, from careful data collection and rigorous analysis to transparent reporting and reflexivity on the researcher's role and potential biases.

\*\*\*\*\*

## CHAPTER 9

### REPORT WRITING AND RESEARCH ETHICS

Research is defined as systematic inquiry to develop knowledge. Knowledge is primarily reported through writing. A written report, termed a 'research report,' describes the research that the author has completed, includes every facet of the research, and analyses the results. The report is said to be the formal disseminator of knowledge – it converts ideas into information by making available to others analyses of original data. The act of research implies a relationship with the wider community that goes beyond the tasks of collection and analysis. Reporting research, like publishing, is a crucial act that ensures the flow of ideas in a research community. The imperative to report research is reinforced by funding bodies who demand accountability and expect public good in return for their financial support. A research report thus serves the dual purpose of giving and being accountable.

**Which aspect of research is more critical for ensuring reliable and responsible research?**



Figure 9.1 Report Writing

### STRUCTURING RESEARCH REPORTS, DISSERTATIONS, AND PAPERS

A research report presents your research study in the way that you judged was appropriate at the time, given your particular readership, supervisor or examiners, their requirements, the framework of rules interpreting the particular area of study or genre, and the academic customs of your field. Structuring a report is like preparing a jigsaw puzzle: as pieces can be hard to shift once joined together, the report is easier and quicker to write if you plan ahead a little. Provided you have taken into account the readership perspective and studied some of the models of reports specific to your field, structuring follows a predictable logic and the report is likely to be well understood and to make a clear contribution to the understanding of

the audience. Like most aspects of research study, the structure of a report is recurring, even cyclic. The academic style is recursive, which means that projects and reports are structured into subsections or recurrent components, each of which can be addressed quickly in turn. The general structure report closing the recursive model is: Introduction section, Specifics, General Discussion and Conclusion. The introduction asks “what is the problem the research intends to address?” and says what is “known” and what “is unknown.” The specific sections address the general area of study; then describe the problem in an articulate way, logically structured into smaller clauses; explain current possible solutions to the problem in a summary of previous research; present your research question and your contribution; and locate your work within the field. The General Discussion section or sections answers the meta-research questions “how did you do it?” and “what did you find?” for your specific research method and for each of your research questions. Finally, the Conclusion answers the question “why do the results matter?” and discusses the implications of your work visually identified as consequences of the results for that wide area of study.

#### ➤ Dissertations

A thesis or dissertation may be defined as a monograph describing original research undertaken in partial fulfilment of a degree. It is likely to be larger than a journal paper, but smaller than a book, generally between 60 and 350 pages long. It also may have some of the same structural components as a book, but they will be differently constituted, and likely not in the same order. There often is a longer introduction that does not just orient the reader to the work accomplished, and reveal the contribution, but explains why that research needs to be done in the first place, including a review of the relevant field that justifies the research. Dissertations are often unpublished in any other form, and thus may be available accessibly only through research archives or repositories of the institutions conferring the degrees. This creates a rarity of access, actively characterises the dissertation as not being publicly available for academic scrutiny, and establishes the author and their institution as custodians of the original research. Students looking to understand what a similar work might look like would be better advised searching on research archives for completed dissertations that are available to be read publicly, or reading Master's theses completed at the institution by senior students. These would provide models closer to the one that would be expected, and display a variety of styles and structures. There also may be institutional templates that students are encouraged to use. To aid accessibility, many institutional repositories now require that students upload their final PDF versions to the archive after graduation to enable findability.

### ➤ Research Papers

Research papers are an important vehicle for disseminating research findings. Research papers can vary from very short notes to full-fledged articles of 60 or more pages in length. Their length and complexity aside, a research paper has all the hallmarks of a typical research report: it investigates an area of inquiry and attempts to add to our knowledge about it. Research papers differ from research reports in one important respect: research papers often entail the analysis of a dataset and report the findings derived from this analysis. More specifically, they may develop and apply an algorithm that tackles a problem in the area of inquiry, or they may apply and report the findings from an established algorithm to advance a theory in the area of inquiry. One may argue that both categories are equally valid ways of conducting research, but it is by far the objective of a research paper to contribute new knowledge or insight, and lesser the objective of a research report. As with research reports, the most popular format for a research paper is based on the IMRaD structure: Introduction, Methods, Results, and Discussion. But in research papers, the Introduction is often extended to also offer information about related work in the area of inquiry, which is usually called the Literature Review section in research reports. In addition, the Discussion section may also discuss implications of the paper's findings for related work in the area of inquiry rather than just the paper's contribution to the area of inquiry. In particular, we have observed Comparative Discussion sections in research papers where these are dedicated to contrasting and comparing each individual contribution from related work to the paper's findings.

### **PLAGIARISM AND CITATION STANDARDS (APA, MLA, CHICAGO)**

Plagiarism occurs when an author includes another author's ideas, processes, results, or words without giving appropriate credit. Plagiarism is a serious ethical offense, and the consequences associated with it vary by institution and context. Most research ethics committees and institutions require applicants to be aware of and to comply with copyright laws. This includes using citation standards that are available from many resources or consult your instructor or school for more information and assistance with these laws and citation standards. Commonly accepted formats include APA format for the social sciences and sciences, MLA format for the humanities, and Chicago Manual of Style for the general field. The following summaries provide helpful information on each of the above formats: Research and writing can be complex, time-consuming, and confusing. Even experienced researchers and writers sometimes are uncertain about the "rules" surrounding research and

writing. It can become even more complicated when someone asks you to complete a project in a specific format for your course assignment. The following summarizes some of the widely utilized writing styles in academic research and writing: APA, MLA, and Chicago. Each writing style has specific formats for citation, but the method is the same regardless of the format. The information is presented in this paper to provide guidance only. Students should consult their instructor for their preferences (or requirements) for research papers submitted in their courses. The American Psychological Association (APA) style is composed of citation rules including word usage, grammar, and paper formatting developed by the American Psychological Association. The standard emphasizes clarity and conciseness (for example: using an active voice), which makes it a common choice in the social sciences and sciences. The APA format uses a simple in-text citation system. The author's name and the year of publication appear in parenthesis directly after the summary or quoted material (for example: Bandura, 1992). The detailed reference giving the full publication information, including the primary source, will be found in the reference list at the end of the document.

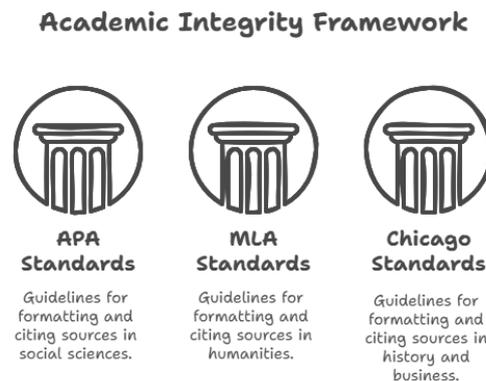


Figure 9.2 Plagiarism And Citation Standards

### ➤ APA Style

There are many citation standards in use today, but those most often required in social work and the related disciplines are the American Psychological Association and the Modern Language Association styles. The APA format offers a clear template for getting most of the basic elements into a reference list entry: Who wrote the work, when was it published, what is the title of the work, and where can the work be found? Other than including a DOI whenever possible, it does not worry about including the particulars of retrieval, such as the date you accessed an article or that it was in an archive. Once you understand the basic structure, creating citation entries and in-text citations is relatively easy, and it can be done

quickly. APA has become the consensus standard for social work writing, but its use in social work research writing stretches back decades. It can take time to become familiar with the variety of types of items that you will need to cite. A map through its formats is crucial – it becomes easy and routine to create entries when you begin to understand it, and it helps save someone's editor time. Look to news collections if you need a quick refresher on how to format a specific item. Don't be afraid to try formatting a couple of entries yourself before checking the guide.

### ➤ MLA Style

Modern Languages Association format is widely used in nearly all publications, as well as most college courses. General Guidelines – The major MLA style requirements include the following. A standard, easy-to-read font, appropriately sized. Most of your document should be double-spaced, with no extra spacing before or after paragraphs. Your citations and works cited pages should be formatted using hanging indent set to half an inch. If necessary in your document, notebook's default 1-inch margins on all sides. If necessary in your document, page numbers in the upper right corner of every page. An easily readable title, centered on the page, which does not stand out in any unusual way from the rest of your document. In-text Citations – The bulk of in-text citations in MLA format are parenthetical citations at the end of a quote, paraphrase, or summary. Most in-text citations are entered at the end of the borrowing. The citation consists of the author's last name and page number, separated by a space. The match between the in-text citation and an entry in the works cited page allows your instructor to locate the relevant source those words in your paper drew from. Works Cited – The works cited page is a list of the sources cited in your paper, located after your text and any appendices. It is generally on its own page, but in the case of very short documents may straight-follow the last page of the text. The sources in the works cited list are organized in alphabetical order based on the first item in each entry. Your list entries provide the information needed to locate your sources, including what type of source it is.

### ➤ Chicago Style

There are two types of Chicago citation styles: "Notes and Bibliography" (NB) and "Author-Date" (AD) (or "Reference List"). The first one is frequently used in the field of humanities, whereas the AD style is used mostly by sciences. Both are similar to one another but differ significantly. Also, it is quite different from other citation styles. There are modes for both in-text and end-text citation. For citation within the text, the author can either follow mode NB

or AD. For citing at the end of the work, either Bibliography (NB) or Reference List (AD) is to be made. For citing within the text, the preferred way is to use footnotes. But if footnotes are avoided, the author can choose the following way to cite sources in-text. In such cases, besides the provided item, the author must additionally provide the readers the source for the first citation or other citation modes for other authors. A page should provide full information at the first citation, and then they could just give short items for any subsequent citations. Chicago style very strictly prohibits paraphrasing or directly copying words or phrasing from the sources without proper citation and hence should be used only by a small number of people probably by people writing research able otherwise other people can be avoided by their references list. Chicago style uses footnotes to cite the works or provide further information; the author can use parenthetical notes as found in other styles but within texts, but within-citation notes are extremely discouraged. There are however multiple other points that the essay that have to be strictly kept in mind while using this style. When paraphrasing, the writer should frame the entire text in his/her own words and also should not copy not too many words. Quoting words from the sources is completely fine, hence mentioning other authors are ought to be cited completely appropriately via footnotes.

### **ETHICAL CLEARANCE, INFORMED CONSENT, AND PARTICIPANT CONFIDENTIALITY**

Ethics should be addressed at the beginning of a research project. Ethics is a moral code used to distinguish between right and wrong behavior and to provide guidelines on how people should behave. It is essential that every consideration is given to any possible harm that may come to your participants as a consequence of taking part in the research. Aside from setting the standards for researchers, ethics help maintain a discipline's reputation. It is encouraged to always be highly aware and utilize ethics in research. When ethical breaches are made, and researchers are disciplined or the research is retracted, it is the discipline's reputation that suffers. Ethics approval follows a process requiring you to submit your proposal to a review committee who will assess whether your research meets the relevant ethical principles before it is given the green light. It is possible for approval to be granted with conditions or for a proposal to be rejected altogether, meaning you must then make amendments to your proposal according to the outlined recommendations and resubmit it. Approval is granted on a case-by-case basis, taking into consideration the population being studied and the nature of the research. For example, if your proposal involves research with children or vulnerable communities you are required to demonstrate how you will minimize participant risk if the

research were to proceed. Furthermore, you will need to outline any possible benefits the participants might receive and why these outweigh the risks.

➤ Ethical Clearance

Research ethics should be appropriately addressed when conducting research on a field. Addressing ethics is necessary to show our respect for the fidelity of the researched persons as well as the discipline. Any inappropriate and unprincipled conduct in research would not be acceptable for the studied group, community, or territory involved. The purpose of ethical compliance is to protect participants from any potential risks and/or consequences associated with participating in the research. Ethical considerations should be addressed from the start to the end of the research, as well as during the dissemination and publication of the findings. Obtaining ethical clearance is important but not compulsory for a number of research studies. However, most research fields lay laws governing how ethical clearance must be obtained, and it is often mandated for research on human and animal subjects. While ethical clearance protects the interests of research subjects and participants, it also protects researchers, as well as confidential or proprietary information within their institutions or governments. Ethical clearance is often required for a number of research studies involving human and animal subjects. There are guidelines defining how and when an ethical board must be consulted in order to be granted appropriate ethical clearance. Ethical research on social groups is governed by these principles to assure the safety, certainty, well-being, confidentiality, and respect due to research participants. Institutions often have committees on human or animal research ethics that assess the research problems and provide the researcher with ethical clearance. This process provides researchers with feedback on the study design, and makes sure that all aspects of the project are ethical. Note that applying for clearance typically entails associating with an ethical approval procedure form and a consent form that the study participants may sign.

➤ Informed Consent

One of the most essential ethical concerns any researcher must address is whether or not to seek informed consent. Ethical adult research hinges on the premise that the decision to become a research participant should result from consideration and deliberation. Informed consent implies that researchers provide prospective participants with the relevant information needed for decision making regarding participation in the study. The principle of informed consent asserts that it is wrong for researchers to expose research participants to

risks without adequate justification. Ethics review committees scrutinize proposals to ensure that adequate measures are in place to ensure valid consent processes that comply with cultural conventions, legal requirements, and policies of protection. The most essential and universal provision is that participants receive adequate information specific to their situation to enable them to make deliberative, well-reasoned decisions about their options. The informed consent process has two main components: the description of the systematic investigation and its purpose and the decision about whether or not to participate in it. Failure to provide information to participants about the study makes it impossible to claim an ethical decision about whether or not to participate has been made. Furthermore, the quality of the information that a researcher provides for the participant's decision is important. For prospective adult participants, complete and detailed information allows individuals who are normally competent to weigh the risks and the potential benefits and to decide whether or not to enroll.

➤ Participant Confidentiality

Researchers must ensure that participants' confidentiality is protected when conducting interviews so that the participant's personal identifiers cannot be linked back to their responses, either directly or indirectly. In the context of qualitative research, consideration of participant identification and confidentiality is particularly important, as participant responses are often more revealing and sensitive than in quantitative studies. Strategies to protect participant confidentiality must be clearly outlined and explained to participants prior to starting the research to enable informed consent. There are a number of strategies for ensuring participant confidentiality. These strategies can be grouped into two broad categories: ensuring anonymity before the research occurs and ensuring anonymity after the research takes place. When research participants are anonymous, they cannot be linked to their responses in any way. This may be accomplished by using a focus group moderator who is not connected to the study sponsor, using a computerized assistant to facilitate data collection, or providing the research respondents with a means such as a unique identifier to allow them to identify themselves rather than the researcher. It may be possible to provide an extra incentive for participants to return for subsequent research without personally identifying them. When audio and video recordings are kept completely secure, removed immediately from view, and destroyed after a secured period, their temporary retention does not necessarily breach confidentiality. Token identifiers could be used to identify audio and video recordings, provided the recordings are secure and destroyed within a bounded period.

Similarly, unlinked, scrambled subject identifiers, which are paired with the subject's identifiers on a separate secured file, can be used during the period before the subject identifiers are destroyed.

## **USING TURNITIN AND GRAMMARLY FOR ETHICAL AND POLISHED WRITING**

The tools that comprise the professionalization of writing environment carry the trace of the relations and economy of the writing field. While formal guidelines, style manuals, and the act of proofreading itself are part of a currently taken-for-granted social consensus about the ethics of writing, there are others more recent tools that need to be described and discussed: academic writing usage checking tools. We mean by this expression the use of tools meant to facilitate the detection of these uses and, in so doing, to hinder their enthusiastic expansion. The two that have appeared as the major ones are plagiarism checking bots and grammar checkers. They represent not exactly opposite ends of an ethical control domain but more points along a continuum where the goal is to be as original as possible while remaining within the boundaries of the writing consensus. Our argument is, of course, that in fact, both types of checking point toward the same writing ethos, namely the necessity to follow concealed field destination rules as rigorously as it is possible. At the scholarship level, the fact that the majority of subscribers are aware that having their uses detected is a serious issue for them means that broadcasting last-minute addenda becomes a major service within the scholarship editing business. It also suggests that the agency available to authors is of a very low degree. This creates radical scriptural constraints and the possible loss of autonomy could be at the expense of credibility and reputation security. However, the reality of these services as far as copy-editors are concerned means that those who do not or cannot afford them remain susceptible to being caught in a scriptural loss of credibility.

### ➤ Using Turnitin

Turnitin is commonly known for its similarity detection function. This means that, if a student submits a paper, Turnitin will search its database for papers with similar phrases, sentences, or paragraphs, and alert the instructor. However, instructors can search students' papers for similarity independently of their role in Turnitin. This feature allows instructors to check for similarity before a paper is submitted by a student. We think this feature is particularly useful for papers where unintentional plagiarism could accidentally occur, such as when the source is in a foreign language. Using a similar paper as a template for

translation can result in sections that closely resemble the original without the proper attribution. Turnitin also has an unusual system for detecting similarity. Typically, similarity detection systems search for identical characters in a given order and location. Turnitin has the capacity to disregard common substitutions made to prevent detection while keeping the paper's meaning intact. One method is disregarding common substitutions entirely; by doing this, it can find all copies of a paper in which some letters or phrases were changed from Greek characters to Roman characters. Another method is substituting other characters with longer series of characters common to certain languages. This is done to detect mirror copies of papers, in which characters are substituted to avoid detection. In fact, a mirror copy of a well-known literary work has been detected.

➤ Using Grammarly

Grammarly is a writing assistant that helps users improve their writing. It has several features that distinguish it from basic writing programs. First, it provides advanced suggestions that take into account the context of the writing window; advanced suggestions are recognized by a blue circle. Second, the service will make suggestions in real time; basic error checking only activates when you complete a sentence. Third, the tool makes suggestions for a wider range of writing issues; for example, it will make suggestions to improve your sentence's clarity, engagement, delivery, and style. Fourth, it checks for several creative writing issues like wordiness, long sentences, passive voice, and more.

Another thing that it does differently from other services is that it provides what are known as "hard" and "soft" rules. Hard rules are rules that apply to every piece of writing the service checks. For example, hard rules include errors in standard grammar, spelling, punctuation, and connoisseur phrasing. On the other hand, soft rules are suggestions for errors that may be admitted depending on the style of the writing. For example, when writing formally, it might make sense to avoid contractions, use the Oxford comma, or use the active voice. These soft suggestions apply only to specific pieces of writing at the writer's discretion. Moreover, in the free version, the soft suggestions relate only to conciseness. The fully featured pro version provides specific soft suggestions related to a variety of categories; however, the free version is still useful on all counts. Because it uses advanced software and has a wide range of suggestions, it is worth loading a piece of writing into the program to get back feedback. Explore the suggestions with an eye to grammar but also to clarity, style, and conciseness.

## **Best Practices for Research Writing**

It is important to keep in mind that academic writing is different from other forms of writing. The most significant point of difference is that the intended audience is a scholarly audience. Academic writing has readers who are experts in their fields; therefore, the writing must be precise and must meet the expectations of research presentations. Rigor, clarity, and precision are key aspects of all forms of academic writing. Because of the scholarly nature of the audience, there is generally no need for definitions of basic concepts; specialists will already know the basis on which you are defining your concepts. The writing should be formal, so the casual, conversational style of creative writing is not appropriate, nor is the use of the first-person form of address. The writing should be tight; meaning unnecessary words and expressions should be trimmed if possible. Scholarly writing relies on fact rather than opinion. Furthermore, factual statements must be supported with evidence, and every part of the writing should establish the credibility to the audience. The use of various types of citations is an important part of establishing credibility, but citations should also be used judiciously, bearing in mind their intended purpose. Additionally, the use of headings is a key part of all academic writing, allowing for a clearer structure, which is helpful not only to the writer but also to the audience. The headings can also work as signposts for the reader, who may not need to read every word because they can skip to the specific points of interest. There is a logical flow to academic research writing. After what you intend to focus on is outlined in the introduction, a literature review provides a context for the study. This context is then followed by the methods used to gather and analyze data, and the results are in the subsequent section. The results provide answers to your research questions and the inferences that can be made from them. Following this is the discussion section, where the focus is on the wider significance of the results. Finally, the conclusion wraps up your paper.

### **Chapter summary**

The chapter begins by outlining the standard structure for various research outputs, including research reports, dissertations, and academic papers, emphasizing the importance of a clear and logical flow of information from introduction to conclusion. This section covers essential components like abstracts, literature reviews, methodology descriptions, results presentations, discussions, and references. A significant portion of the chapter is dedicated to addressing plagiarism and proper citation practices. It explains the importance of academic integrity and provides detailed guidance on using established citation styles such as APA, MLA, and

Chicago. The chapter emphasizes the consequences of plagiarism and offers practical advice on how to avoid it through proper paraphrasing, quoting, and referencing techniques. This section also likely includes examples of correct and incorrect citation formats.

Beyond the mechanics of writing, the chapter underscores the ethical responsibilities inherent in research. It covers the process of obtaining ethical clearance from Institutional Review Boards (IRBs) and emphasizes the critical need for informed consent from participants. Protecting participant confidentiality and anonymity is also a key focus, outlining strategies for ensuring data security and responsible data management. The chapter highlights the ethical implications of research design and data collection. Finally, the chapter introduces practical tools that can help researchers produce high-quality and ethical work. It discusses the use of Turnitin for plagiarism detection and Grammarly for enhancing the clarity and precision of writing. By incorporating these tools, researchers can improve the quality of their work and ensure compliance with ethical standards, contributing to the overall rigor and integrity of the research process.

\*\*\*\*\*

---

## CHAPTER 10

### PUBLISHING RESEARCH AND ACADEMIC VISIBILITY

The advent of the digital age through the World Wide Web has irrevocably transformed many facets of the human experience, for academia and scholarly communication is no exception. During the past three decades, we have witnessed astonishing developments in information sharing, digital scholarship, academic visibility and scholarly communications in general, from the last wave of diskettes used to share work-in-progress to the current preprint archives, from the personal web pages of researchers hosted on disused department servers to the research profile services, from the skeletal Web of Science to the increasingly diverse dimensions of the Semantic Scholar, especially in the wake of recent releases of AI-supported bots. Publishing research findings is still preferred by most — if not all — academic institutions around the globe as a primary way of proving that scholars contribute to the advancement of accumulated human knowledge while their ability to do so is being compensated by stakeholders.

#### JOURNAL TYPES, INDEXING (UGC-CARE, SCOPUS, WEB OF SCIENCE)

There are different types of academic journals available for publishing research. Without trying to categorize the journals available in the market, the following are the major types of journals that researchers often consider for publishing new knowledge. In India, UGC-CARE journals are the most important journals that provide academic recognition. All the journals listed by UGC-CARE are indexing journals. Journals listed by UGC-CARE are often considered for recruitment and promotion of teachers and researchers in universities and colleges. UGC-CARE journals have advantages and disadvantages. The advantages are: 1. Free to publish, 2. Widely used for recruitment and promotion in India, 3. Various subjects are covered in many languages, 4. Gives recognition for the work done. The disadvantages are: 1. Many articles are published questionably or inappropriately, 2. Heavy article processing time with no editorial control or unknown peer-review quality, 3. Research quality in many articles is questionable, 4. Service quality provided by the journals varies, 5. Some journals are considered predatory. Scopus journals are available. Both indexed and non-indexed journals are available in Scopus. Only indexed journals are considered for citation and other analysis. Journals become unavailable and/or are removed on a quarterly basis without prior notice. Journals are removed only after due process and consideration. The advantages of publishing in Scopus journals are: 1. Research content is made available for

global visibility, 2. Used for hiring, promotion and tenure, 3. Indexed journals make it easier to evaluate research quality. The disadvantages of submitting in Scopus journals are: 1. Publishing in low-quality journals may adversely affect research reputation, 2. Non-English journals may not be accessible, 3. Highly cited papers may be too few to measure impact, 4. Citations may not be the best measure of research quality and impact.

### Journal Types and Indexing Process

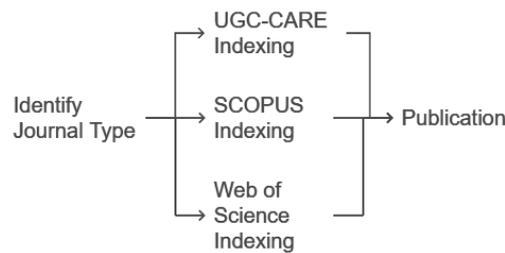


Figure 10.1 Journal Types

#### ➤ UGC-CARE Journals

In response to the lack of quality publications in social sciences and humanities that promote research visibility and contribute to impact factor frameworks, the University Grants Commission established the UGC-CARE databases by proactively collaborating with numerous disciplines. Within the UGC-CARE format, it supersedes its list of journals with pre-specified criteria that decide the quality of the journals one wishes to publish-in. To qualify for the archived UGC-CARE list, a journal should achieve five years of indexed publication, reporters with print and online circulation, peer-review policy which strictly adheres to ethical reporting, editorial board who contribute significantly to its journal policy, and timely publication with quality manuscripts that at least correlate with the index entry. Journals featured in the UGC-CARE list are further divided into two groups: Group I comprises journals indexed in globally reputed indexes and Group II features journals that pass the minimum quality-prevalence criteria as laid down by the UGC committee for the respective disciplines. The database is further sub-divided according to the respective disciplines. Although the UGC-CARE list is useful in helping researchers from non-STEM disciplines to enhance research visibility, the difference in UGC-CARE group parameters and information for submission differ greatly from journal to journal. As such, it mandates the decision of the researcher in choosing quality periodicals for publication and becomes

cumbersome. Moreover, the quality and visibility of a journal largely depend on how it manages its editorial position.

➤ Scopus Journals

Scopus is a bibliographic database containing abstracts and citations for academic journal articles. The database was launched in 2004. In addition to being a citation database, Scopus is also a publisher of primary literature in some fields through its various publishing imprints. Scopus provides researchers with a clear and comprehensive overview of the peer-reviewed research literature in the natural sciences, social sciences, medical sciences, physical sciences, and life sciences. Scopus also tracks citation statistics for articles and allows for analysis of citation trends. Scopus is the largest abstract and citation database for academic literature. It covers about 22,800 titles from approximately 5,000 publishers around the world. There are over 72 million records in Scopus, with 1.4 billion cited references back to 1970. Scopus also tracks sources published before 1970, with some sources traced back to the 1500s. The Scopus database contains articles from several types of content documents, including: over 22,800 titles published in peer-reviewed journals from all over the world; approximately 38,000 conference papers published in proceedings from about 3,500 conferences; a small number of trade articles published in trade publications; and patent records. Scopus has extensive international coverage of sources and large-scale citation tracking, offering researchers a powerful tool for comprehensive resource tracking.

➤ Web of Science Journals

The knowledge that only one fifth (20 percent) of the indexed journals actually on the list of journals that are indexed gets to occupy the first quartile (Q1) position in the impact factor list. Ties in impact factor collapse a few journals into each quartile based on the lower number of total top ranking journals. Depending on the size, sufficient numbers of those articles in those mostly high research output journals, actually published by authors based at funding sponsoring institutions, there is also speculation and concern by many Indonesian researchers on why high quality articles are hard to get published in those elite Q1 journals and hope that Indonesian universities publishing for the sake of being acknowledged by the university would change their mindset to produce more and better high quality research outputs. Yes very few Q1 journals do publish relatively smaller amounts of journals for countries such as Indonesia. In short even if the writing is accepted or published in those, the increased number of citation is not always guaranteed. The notion academic visibility is

synonymous to indexed mostly listed research output. It is the aim of researchers as these lists can be looked across continents to see if we are on par with other international renowned journals. These journals do sometimes prescribe to an extremely key policy strict on the aspects of wording technicality. Academic visibility in the digital repository department has to remain balanced as the 'key aspect' of higher ranking in journals is that of speed.

## **THE PEER-REVIEW PROCESS AND AVOIDING PREDATORY JOURNALS**

Peer review is an important part of academic publishing, which establishes trust that work being made freely available is of a certain level of quality. It involves the submission of work, usually in an article format, by a researcher to a journal editor. The editor accepts the work for submission to the journal, sends it to experts in that particular field, and makes a decision based on the feedback from those experts. If the expert feedback is that the work is not ready for publication but has merit, the editor will usually pass the suggestions on to the author for revision. Peer review emphasizes unbiased assessment of the research, guarding every field against the inclusion of work that is flawed, refuted, or misrepresented. A high-quality journal will typically have various checks and quality measures in place. While the peer-review process may appear tedious, it ultimately leads to the refinement of work, and the quality assurance means that published articles are more likely to be cited compared to those not formally reviewed. There are various types of peer review and equally, various issues that come with peer review. While most popular are double-blind peer review where both author and reviewer remain anonymous, and single-blind where the author remains anonymous but the reviewer does not, there are others where neither the reviewer nor the author is anonymous, or review is done by a panel of experts. Furthermore, many researchers are pushing back against peer review, having it removed entirely from the publishing processes, instead opting to upload their work to freely available preprint archives to introduce the work into the discourse.

### **Identifying and Avoiding Predatory Journals**

The fundamental element of the scholarly communication system is the publication of journal articles. The accountability of scholarship depends upon the integrity of the journals used in the publication output of any field or discipline of inquiry. But this accountability is being threatened by predatory journals that may have contributed an estimated one million or more publications in just a few years. Further, recognizing predatory journals requires access to expertise or tools developed for that purpose. And for authors, particularly those in

developing countries, properly written guidelines to assist them would be useful. This paper provides examples of guidelines to identify and avoid predatory journals. Scholarly journals serve the academic community in ways beyond facilitating and disseminating research and scholarship within an academic discipline, field, or special interest. They provide aspects of the peer-reviewed publishing bottleneck central to the system of scholarly communications. The peer-reviewed journal process provides the academic community with evaluations of ethics and accuracy. Such journals have become a means for scholarly assessment. A common tool in this assessment is the journal impact factor, a quantitative publication ratio used to evaluate the change in visibility of journals within specific categories over the years. However, the value of any impact factor is dependent upon the accuracy of methodologies that are specific to the discipline or field. Predatory journals have appeared in this landscape, taking advantage of this urge for researchers to as quickly as possible share their research, without the assurance of quality or ethics as would normally be supplied through traditional publishing services and procedures. And while this landscape is crowded with opportunists, some are worse than others, with varying levels of predatory behavior.

### **CREATING AND MAINTAINING ACADEMIC PROFILES: ORCID, RESEARCHGATE, GOOGLE SCHOLAR**

Research visibility is the by-product of participating in both the publishing process and its dissemination circuit. When research content is openly shared and pledged to a wider audience, the likelihood of others engaging with the work increases. During the writing process, authors should clearly signal their contributions to the knowledge universe and elaborate external signposts to where their work can be accessed. In this digital age, both the collective and the individual are responsible for that visibility. Open research and open access, open educational resources, social media promotion, and content sharing are initiatives that head in that direction. Once the research is published, however, authors should create academic profiles to build their own information infrastructure and rapporteur academic affiliation resources where others can discover their work. These digital footprints are not only about the need for academic visibility and identity; they also promote the structural organization and interconnection of academic content. It is through those academic profiles that academics position themselves in the academic community and endorse their visibility. In what follows, we outline three services that assist researchers in their academic profiling. Individually or collectively, they can already serve to channel the dissemination of academic content. The aim is not to disclose their merits or weaknesses or provide in-depth technical

instructions on how to choose and set them up, but rather to introduce them as working tools for visibility and exchange that should be mastered and conceived during doctoral training. We outline the distinct functions of an ORCID iD, a ResearchGate profile and a Google Scholar Profile, how each is structured, which elements can enhance their performance, how they relate to each other, and how non-academic readers can engage with them as resources to learn more about our capabilities and expertise. By no means do we intend to provide an exhaustive review of existing services. Academic profiling services encompass a wide range of solutions with distinct particularities and purposes. Yet, although existing, being heterogeneous according to disciplines or models of production, wrapper services are relatively limited.

### ➤ ORCID

The creation and maintenance of academic profiles enables a researcher to increase the visibility of their research outputs. Most repositories have an integration with ORCID, through which researchers may upload their previously published works to their profile page when the works are made available through the repository. In addition, most publishers require authors to create an ORCID ID and link it with their academic profile when submitting academic works for publication. ORCID provides a non-discriminative registration policy, supporting the list of identifiers of external users and allowing users to create their CV, accessible through a unique URL. The page gives general information such as name, e-mail, website, picture, and address and is a flexible tool in terms of presenting the researcher information. In fact, ORCID is much more than a unique identifier system. Actually, ORCID is a non-for-profit organization that aims to give a unique and persistent identifier for every researcher. The main features that make ORCID a first-rate unique identifier for researchers are that it provides: ORCID is thus a legitimate authentication tool. Thanks to its partnerships with numerous other publishers and institutions, it enables the effective management of content created by researchers throughout their academic careers, as well as the linking and sharing of this data between many academic platforms. In addition, it allows researchers to manage diverse identifiers through a single registration. When searching for information and academic content uploaded to other platforms, users' attention may be drawn to the user's diverse identifiers, linking their academic activities. Furthermore, some academic information systems may ask for an author's ORCID in order to validate their identity.

➤ ResearchGate

ResearchGate is a free social network aimed primarily at researchers. Founded in 2008, it has become the most comprehensive database of research publications, with more than 200 million users, 145 million publications, and 30 million full-text articles. Peer-reviewed articles, conference papers, preprint articles, posters, book chapters, patents, research proposals, and data can all be found on ResearchGate. Registered users can request full-texts uploaded by other authors, and ResearchGate will usually send the request to authors to facilitate obtaining the document. By making it possible to download documents published by peer researchers across the world, ResearchGate fills the information gap often caused by expensive paywalls at traditional publishers. ResearchGate removes this obstacle and allows researchers to expand their academic profile by sharing all kinds of research products, from their published papers to their grey literature. ResearchGate profiles provide information about research interests, as well as career milestones, such as education, previous work experience, current work experience, awards, and organization memberships. ResearchGate is a website that facilitates networking, and it has features that resemble other social media, such as a news feed, notifications, groups, and a system of votes and comments for uploaded content. However, unlike other services, ResearchGate does not need to monetize its platform through selling advertising, data collection and selling, or boosting featured content, as it is funded through institutional subscriptions. Such subscriptions allow institutions to access ResearchGate's Special Services features. ResearchGate also maintains a database of research metrics and a statistics/analytics dashboard that tracks numerous indicators related to specific authors. ResearchGate recently introduced Verified Accounts for select researchers.

➤ Google Scholar

Google Scholar, the academic feature of Google, was launched in November 2004 as a freely accessible web search engine that indexes the full text or metadata of scholarly literature across an array of publishing formats and disciplines. It uses Google's patented hyperlink analysis technology to enhance the precision of its search results, so researchers can quickly locate the top three results of a user-entered query string. Google Scholar's indexed content is freely visible but it is not necessarily freely available in full text. While content from many self-archived research studies are available in their entirety for free online, Google Scholar also indexes materials solely available in full text behind paywalls, such as papers published in subscription-access journals, eBooks published by commercial publishers, and

monographs in digital libraries. Google Scholar is predominantly used by students, researchers, and research related workers for academic purposes. It hosts a variety of features, such as millions of indexed datasets, and ability to search, browse, and analyze citation patterns, which have made it particularly attractive to scholars worldwide. Factors such as Google's market share, the presence of a search box in the homepage of Google Scholar, its feature of a single search at a single point, and its no-cost service provided to search and access digital data account for its ranking as the most popular search engine by Scholarly Electronic Environment Monitors. In the area of scholarly services, it was cited as the most established, desired, and developed academic search service for scientific research.

### **Maintaining Academic Profiles**

The swift technological advancements have rapidly changed how scholars and institutions promote their research; this has consequently shifted how science communication and engagement are achieved. Publishing research in repositories or online journals is no longer sufficient for amplifying research findings, as researchers also need to maintain a constant online presence by posting research updates, interacting with target audiences, and distributing dynamic content across multiple platforms and marketplaces. Online metrics, analytics, and visibility tools also enhance researchers' abilities to elevate engagement and visibility. Research engagement extends to academic profiles and institutional repositories, which are the first data source many journalists and content writers use as search engines. Though academic profiles are crucial channels to make research visibility much more feasible, they still remain underused by researchers. Important peer-to-peer platforms increase the chances of having colleagues or experts from other fields/studies contact them to exchange ideas, fund projects, co-author papers, or simply provide guidance on career paths. Academic profiles remain crucial steps to start bridging the visibility gap. Researchers have set up their profiles, mostly missing the crucial step of updating them with relevant information such as values for the links and section fields. Managing public profiles and institutional pages does not only enhance the strength of an academic's visibility and career but also encourages a greater community open science dialogue around research and data.

### **TRACKING CITATIONS AND METRICS USING PUBLISH OR PERISH, SCOPUS TOOLS**

Tracking your citations and metrics is a way to see what types of impact your work and the work around you is having. It can point to specific audiences or other fields of research you

may not expect. While many consider these net numbers just vanity metrics, they reveal networks that influence the wider community. Here are some of the things that citations and metrics can tell you: 1. Are you being read or cited in your own field? 2. Are you being read in a practical way, or are people implementing your research? 3. What regions or industries find your work important? 4. How do you compare with others? Scholar counts citations and shows resulting metrics with relative simplicity. There are many tools with good analysis and tracking methods. You can create multiple metrics, compare between authors, and dive deeper into your authorship. These services allow us to not only narrow down our results based on selection but also utilize other tools to get ultimately better results. Our plus-point including factoring in lots more types of metrics.

➤ Using Publish or Perish

There are many options for tracking your citations and metrics. However, we prefer the free software which is available for Windows, Mac, and Linux. It is a powerful, user-friendly application that retrieves and analyzes academic citations. This software allows you to calculate a range of citation metrics, including total number of citations, h-index, g-index, and more. The main interface is easy and intuitive, with one tab dedicated to data queries and a second tab with the analysis options. The query screen asks you for your name (or the name of another researcher), your institution, and your field of study. It also requests an exact search option, so that only relevant records will show. Once you query the data, the program will review the records you have generated and suggest an author review page. If the data seem correct, you can click to run the query and extract the raw data. The software will display the information needed and you can save it as a report; alternatively, you can extract the citations as a BibTeX file or to Zotero. This software can also analyze co-authorship and co-citation data, similar to those provided by other tools. There is also a support page with a user manual and a help forum where you can obtain further information on how to set the software for the best results. The data appears accurate and is helpful for locating researchers in a particular field.

➤ Utilizing Scopus Tools

Ever since many thousands of institutions have acquired access to this tool. Being among the world's largest abstract and citation databases for academic content, it currently tracks metrics for approximately 76 million items that include over 26 million journal articles published in around 41,000 active titles, and over 20 million book chapters and 25 million conference

papers published in over 4 million active publications. Unlike freely accessible databases, access is paid by an institution to have its users benefit from this tool. This subscription-based approach enables the employment of data curators whose job it is to verify that the journal titles included in the journal list are all reputable journals and to exclude predatory or disreputable publishers. Accessing this tool through an institution allows not only to conduct general searches such as understanding the volume of contributions made by a particular author, but also to explore metrics. Most of the metrics available to users are based on the publications included in the database. Authors are served by the Author Identifier which lists all indexed publications attributed to the profile or scholar in multiple forms and throughout the years, it presents both source normalized and unnormalized citation and h-indexes, as well as citation counts for the last three, five, and ten years. Funding institutions can explore the Affiliation ID based on affiliation names and their variations to assess the performance of a department or institution throughout the years. Journals can be tracked through the Source ID presented for each published article. In that case, metrics are calculated.

### **Chapter summary**

Chapter 10, "Publishing Research and Academic Visibility," guides researchers through the process of disseminating their findings and building their academic reputation. It begins by outlining the diverse landscape of academic journals, differentiating between various types and emphasizing the importance of understanding journal impact factors and indexing databases such as UGC-CARE, Scopus, and Web of Science. The chapter stresses the significance of selecting reputable journals aligned with the research's scope and audience, thus maximizing its potential impact. A key section focuses on navigating the peer-review process, a cornerstone of academic publishing. The chapter explains the stages involved, from submission to acceptance or rejection, and offers advice on addressing reviewers' comments effectively. Crucially, it provides strategies for identifying and avoiding predatory journals, which compromise academic integrity and may exploit researchers. This includes understanding red flags like exorbitant publication fees and lack of transparency in the peer-review process.

Beyond publication, the chapter explores strategies for enhancing academic visibility. It details the importance of creating and maintaining comprehensive academic profiles on platforms like ORCID, ResearchGate, and Google Scholar, emphasizing the value of these profiles for networking, showcasing research output, and increasing discoverability. The

chapter also highlights the role of these platforms in improving researchers' online presence and facilitating collaboration. Finally, the chapter addresses the measurement of research impact, explaining how to track citations and other metrics using tools such as Publish or Perish and Scopus. This section emphasizes the importance of understanding these metrics, acknowledging their limitations, and using them judiciously as indicators of research influence, rather than as the sole measure of academic success. Ultimately, the chapter equips researchers with the knowledge and tools to effectively publish their work, enhance their visibility, and contribute meaningfully to their respective fields.

## References

1. Bhattacharjee, A. (2012). *Social science research: Principles, methods, and practices*. Open Access Textbooks, 3.
2. Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research*. Harcourt College Publishers.
3. Boote, D. N., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Educational Researcher*, 34(6), 3–15.
4. Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage Publications.
5. Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), 1–4.
6. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
7. Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. *British Journal of Applied Science & Technology*, 7(4), 396–403.
8. Kim, H. Y. (2017). Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. *Restorative Dentistry & Endodontics*, 42(2), 152–155.
9. Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook*. Sage Publications.

10. Bazeley, P., & Jackson, K. (2013). *Qualitative data analysis with NVivo* (2nd ed.). Sage Publications.
11. McMillan, J. H., & Schumacher, S. (2010). *Research in education: Evidence-based inquiry* (7th ed.). Pearson.
12. Resnik, D. B. (2015). *What is ethics in research & why is it important*. National Institute of Environmental Health Sciences (NIEHS).
13. Wager, E., & Kleinert, S. (2011). *Responsible research publication: International standards for authors. A position statement developed at the 2nd World Conference on Research Integrity*.
14. Beall, J. (2012). *Predatory publishers are corrupting open access*. *Nature*, 489(7415), 179.
15. Harzing, A. W. (2010). *The Publish or Perish Book: Your guide to effective and responsible citation analysis*. Tarma Software Research.

\*\*\*\*\*

### Author's Biography

**DR. PURNA CHANDRA PRADHAN** is a seasoned professional with over two decades of diverse industry experience across sectors such as baby foods, retail finance, automotive batteries, investment banking, and e-commerce. He has held leadership roles in reputed organizations including Raptakos, Brett & Co., Citi Finance, Amara Raja, India Infoline, DBS Cholamandalam, and UTI, where he served as Chief Manager.

Academically, Dr. Pradhan holds an M.Phil, Ph.D., and MBA, and is a Gold Medalist in Public Policy. He serves as a visiting faculty member at leading business schools, sharing his knowledge and mentoring future leaders. As an entrepreneur, he has successfully launched two e-commerce ventures and currently heads a software consulting firm as Director. His consulting interests focus on e-Governance, especially in the fields of Agriculture, Health, and Education. A certified Black Belt Consultant, he also holds project management and development certifications from the USA. His academic contributions include authoring book chapters, co-authoring two books, and serving as editor of a published academic volume.

---

**DR. PROSPER KWAME KLU** is a passionate academic and researcher with over a decade of experience in teaching and scientific research. He currently serves as a Lecturer at the Department of Agricultural Engineering, Ho Technical University, Ghana, and is also affiliated with the School of Environmental and Biological Engineering at Nanjing University of Science and Technology, China. He holds a Ph.D. in Environmental Science and Engineering, an M.Phil in Nuclear and Environmental Protection, and a B.Ed. in Science.

His research focuses on environmental remediation, nano-catalysis, water treatment, and science education. He has published extensively in reputed journals such as *Chemosphere* and *Environmental Research*, contributing to the advancement of sustainable water treatment technologies. Dr. Klu is widely recognized for his ability to combine theoretical knowledge with practical insights, and he is deeply committed to mentoring students in scientific inquiry.

---

**DR. DIPAK CHAULAGAIN** is a distinguished neurosurgeon, educator, and researcher with a diverse career spanning clinical practice, academic leadership, and global mentorship. He holds degrees in Medicine (MD), Surgery (MS), a Fellowship in Minimal Access Surgery (FMAS), and a Ph.D. in Neurosurgery. Currently, he serves as an Associate Professor at Uzhhorod National University, Ukraine.

Dr. Chaulagain specializes in brain and spine surgery, with research interests in glioblastoma, neuroimaging, endoscopic neurosurgery, and traumatic brain injury. His work is widely published in peer-reviewed journals, contributing to advancements in neurosurgical techniques and patient outcomes.

Beyond clinical and academic achievements, he is dedicated to mentoring over 2,000 international medical students for licensing exams such as USMLE, PLAB, and FMGE. As the Founding Editor-in-Chief of the *Euroasian Journal of Scientific and Multidisciplinary Research (EJSMR)*, he promotes global scientific exchange. Through his platform [www.doktordeepak.com](http://www.doktordeepak.com), he offers medical consultations, surgical tutorials, and academic workshops, bridging medicine, education, and innovation.

---

**DR. HARSHITHA Y S** is the Program Coordinator for BBA Finance and International Business at CHRIST University, Bengaluru—one of India’s premier business schools. She holds a Ph.D. in Marketing and brings over eight years of teaching experience at both undergraduate and postgraduate levels. Dr. Harshitha has actively contributed to academia through the presentation of research papers at national and international conferences, publication of book chapters, and participation in numerous workshops, seminars, and faculty development programmes. Her commitment to education was recognized in 2017 when she received the Best Teacher Award from the Rotaract Club of Mysuru for her outstanding contribution to student development and academic excellence.

---

# Fundamentals of Research Methodology

## A Practical Approach

### Dr. Purna Chandra Pradhan



Dr. Purna Chandra Pradhan is an accomplished professional with 20+ years of experience across baby foods, finance, automotive, and e-commerce. He has held leadership roles at Citi Finance, Amararaja, UTI, and others. A Gold Medalist with a Ph.D., MBA, and M.Phil, he teaches at top B-schools and leads a software consulting firm. A certified Black Belt and U.S.-trained in project management, he has authored books and consults in e-Governance across agriculture, health, and education.

### Dr. Prosper Kwame Klu



Dr. Prosper Kwame Klu is a Lecturer in Agricultural Engineering at Ho Technical University, Ghana, and a researcher at Nanjing University of Science and Technology, China. With a Ph.D. in Environmental Science and Engineering, his work focuses on nano-catalytic water treatment for endocrine-disrupting chemicals. He holds degrees in nuclear and environmental protection and science education. An experienced educator and author of Fundamentals of Research Methodology, Dr. Klu has published in top journals and is dedicated to advancing research and student mentorship.

### Dr. Dipak Chaulagain



Dr. Dipak Chaulagain is a renowned neurosurgeon, educator, and researcher with a PhD in Neurosurgery and advanced degrees in Medicine and Surgery. An Associate Professor at Uzhhorod National University, Ukraine, he specializes in brain and spine surgery, neuroimaging, and glioblastoma research. He has mentored over 2,000 international medical students for exams like USMLE and PLAB and serves as Founding Editor-in-Chief of the Euroasian Journal of Scientific and Multidisciplinary Research. Dr. Chaulagain also shares resources via his platform [www.doktordeepak.com](http://www.doktordeepak.com).

### Dr Harshitha Y S



Dr Harshitha Y S is Program Coordinator of BBA Finance and International Business at CHRIST University, Bengaluru, which is one of the leading Business Schools in India. She did her PhD in Marketing and has 8 years of teaching experience at both UG and PG levels. She has presented research papers in National and International Conferences, published chapters in books and has actively participated in several Workshops, Seminars and Faculty Development Programmes. She was awarded as best teacher by Rotaract Club Mysuru in 2017 for her exemplary work in education and enriching the lives of students.

### Publisher:



Department of Research and Publications  
**A2Z EduLearningHub**  
Kerala, India



[editor@edulearninghub.com](mailto:editor@edulearninghub.com)



[www.edulearninghub.com](http://www.edulearninghub.com)

Publication Month: **June 2025**