Research Methods in Education

Research Methods in Education

Wang Li

Professor and Deputy Director, UNESCO International Research and Training Centre for Rural Education (INRULED), Beijing, China

Peng Liping

Associate Professor in the Faculty of Education, East China Normal University (ECNU) and Director, International Teacher Education at ENCU, Shanghai, China

Qutub Khan

Advisor, UNESCO International Research and Training Centre for Rural Education (INRULED). Formerly, Associate Professor of Economics at Jamia Millia Islamia and Delhi University.

United Nations Educational, Scientific and Cultural Organization

International Research and Training Centre for Rural Education

SAGE | TEXTS

Los Angeles | London | New Delhi
Singapore | Washington DC | Melbourne

First published in 2018

SAGE Publications India Pvt Ltd
B1/I-1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044, India
www.sagepub.in

SAGE Publications Inc
2455 Teller Road
Thousand Oaks, California 91320, USA

SAGE Publications Ltd
1 Oliver's Yard, 55 City Road
London EC1Y 1SP, United Kingdom

SAGE Publications Asia-Pacific Pte Ltd
3 Church Street
#10-04 Samsung Hub
Singapore 049483

Published by Vivek Mehra for SAGE Publications India Pvt Ltd, typeset in 10/12.5 pt Minion Pro by AG Infographics, Delhi and printed at _______.

Library of Congress Cataloging-in-Publication Data Available

ISBN: 978-93-528-0624-9 (PB)

SAGE Team: Amit Kumar, Indrani Dutta, Apoorva Mathur, Madhurima Thapa and

Contents

Detailed Contents

List of Tables

List of Figures

List of Boxes

Preface

Educational research has experienced remarkable growth in the past three decades, providing students with an increasingly sound base of knowledge from which to practice. Yet as we proceed into the twenty-first century, many questions endure and much remains to be done to incorporate research-based knowledge into educational practice. Every step of the way, the educational research process is fraught with decisions, trade-offs and uncertainty. It is not an easy route to traverse. This book is designed to serve as a road map that can inspire the direction if not give absolute instructions. As the field of educational research grapples with its identity, the book uses better tools, more transparent practices and more open conversations to improve our understanding of human behaviour. The different approaches used by grounded theory researchers, phenomenologists and ethnographers are noted throughout.

The book is designed to introduce postgraduate and doctoral students to the process of scientific research in the social sciences, business and related disciplines in general and in education in particular. The book is based on the authors' lecture materials developed over a decade of teaching postgraduate and doctoral-level classes. It acknowledges the international efforts and gives recognition to the contributions of interdisciplinary research studies commissioned from around the globe by UNESCO and its International Research and Training Centre for Rural Education (INRULED) of Beijing, in particular.

The target audience for this book includes students of MA programmes in research methods in education, as well as those embarking upon their MPhil or PhD in education, novice researchers and professors teaching courses on research methods. In addition, senior researchers who carry on research as professional activity can also benefit from this book.

The book is different from other textbooks on the market in several ways. First, unlike other books, this book is not about describing the 'research designs' (empirical data collection and analysis), but more importantly, its purpose is to inform action from beginning to end.

Second, the book is comprehensive, straightforward and jargon-free. It has been designed with the purpose to encourage students and researchers to contextualise their knowledge and understanding of quantitative and qualitative research designs by drawing links between the research question and its theoretical framework with the choice of the quantitative/qualitative methodology and analysis. It supports a theoretical appreciation of the field of inquiry and encourages them to balance the emphasis on methods with review of the literature and understanding of the wider context of research. It provides students with the theoretical understanding, practical knowledge and skills they need to carry out independent research. The book identifies key research methodologies, data collection tools and analysis methods, and focuses on the direct comparisons between them. Each chapter highlights the strengths and weaknesses of a key research method.

Finally, each chapter comprises user-friendly features, such as key terms, summaries, case studies, practical examples and proposed further reading on key topics discussed in this book. Self-test exercises are interspersed throughout the text rather than clustered at the end. This not only places the exercises in more relevant locations but also makes for a lovelier presentation by encouraging the students and researchers to look at easily accessible data sets and use prompts to formulate research questions and produce quality research evidence.

The content of this book is structured into six main themes. However, professors or instructors can add, drop, stretch or condense topics to customise the book to their specific needs.

> **Theme 1—Foundation of Educational Research:** Chapter 1 introduces fundamental concepts in educational research. It presents an overview of major educational paradigms shaping contemporary educational research and summarises how epistemology—a branch of philosophy dealing with the study of the nature, origin and validity of human knowledge—relates and impacts the process of educational research. Chapters 2, 3 and 4 introduce readers to the topic of research evidence for education practice, discuss the philosophical underpinnings of qualitative and quantitative research designs and other methods and describe the major purposes of educational research, respectively. The chapters suggest a head start on how to use these research designs and methods in research inquiries and improve our experiences in conducting scientific research.
>
> **Theme 2—Conceptualising a Research Study:** Chapter 5 deals with the subject matter related to conceptualising a research study and further sets the stage for learning about the research process by discussing issues relating to a study's conceptualisation: the formulation of research questions and hypotheses, the review of relevant research and the development of theoretical and conceptual contexts.
>
> **Theme 3—Designs for Educational Research:** It presents material on qualitative and quantitative research studies. Chapter 7 attempts to highlight the significance and centrality of data in the context of educational research designs. It describes issues important to the design of research that is ethically sound. Chapter 4 discusses mixed method research designs in which the method for qualitative and quantitative inquiry is blended. Chapter 6 describes sample survey techniques and strategies for selecting samples of study participants. Chapter 8 defines the concept of an 'education indicator' and explains the key role of indicators as an integral part of educational research. Chapters 9 and 10 highlight the mechanism for research control for quantitative and qualitative studies and examine both types of designs with different purposes. The chapters highlight some of the critical problems in communicating with the users of research outcomes (audiences).
>
> **Theme 4—Measurement and Data Collection:** It describes gathering of information for a research study. Chapters 9 and 10 discuss the overall information gathering plans and techniques, and present materials on specific data collection methods such as observation, interviews, tests questionnaires and surveys. The chapters discuss the concept of measurement and then focus on methods of assessing data quality.

Theme 5—The Analysis of Research Data: It explains methods of analysing quantitative and qualitative data. Chapters 12 and 13 present the univariate method and construction of index numbers and inferential methods and analysis of time series, respectively. Both describe the development of an overall analytic strategy for quantitative studies. Chapter 9 discusses methods of doing qualitative analyses, with specific information about grounded theory, phenomenological and ethnographic analyses.

Theme 6—Communicating Research and Using Research Results: It emphasises two types of research communication. Chapter 14 outlines the role of text, tables, graphs and charts as formats for presenting numerical data. It discusses how to write about research and prepare a research report, and prepare a thesis or dissertation. The chapter is intended to sharpen the critical awareness of educational researchers with regard to the use of research findings. Finally, the second part of the chapter presents the interpretation and appraisal of research reports and conclusion offers suggestions on utilising research to build an evidence-based practice.

It is our hope that the content, style and organisation of this book continue to meet the needs of a broad spectrum of students, researchers and other readers. We also hope that the book will help a new generation of researchers and users of research designs to solve complicated problems with greater understanding and foster enthusiasm for the kinds of discoveries that research can produce and for the knowledge that will help support an evidence-based education practice.

About the Authors

Wang Li is currently Professor and Deputy Director at UNESCO International Research and Training Centre for Rural Education (INRULED) in Beijing, China. His areas of interest include rural education and rural development, water resources, rural and town planning and geothermal utilisation. Dr Wang Li graduated from Northern China Institute of Water Conservation and Hydropower in 1982 with a Bachelor of Engineering (water resources). In 1989, he got a certificate in Geothermal Utilisation from the United Nations University in Iceland. Beginning from 1996, he began to serve as Deputy Director of UNESCO INRULED till now, during which he got a Doctor of Philosophy (comparative education) from Charles Darwin University, Australia with a study period between 2001 and 2004 as a full time student while keeping the title of Deputy Director of INRULED in the meantime. He has contributed to more than 60 books and papers, undertaken about 50 international research projects and participated in more than 100 international conferences, symposia, forums and workshops. Professor Wang Li has represented INRULED and Chinese Ministry of Commerce in projects in Cambodia and Pakistan and in many international conferences, forums and meetings.

Peng Liping is currently Director, International Center for Teacher Education, and Deputy Chief, Basic Education and Further Education, East China Normal University. A resident of Shanghai, Dr Peng Liping has been handling varied academic and administrative positions over the last 30 years. He is the former editor of the Chinese version of *Prospect* published by the UNESCO International Bureau of Education. He has also been involved in a number of prestigious translation projects involving translation of classic English texts in various subject areas to Chinese.

Qutub Khan (? –2017) started his career as an Associate Professor of Economics in Jamia Millia Islamia, New Delhi, and Delhi University in 1967. After working for about 10 years, he joined UNESCO International Institute for Education Planning (UNESCO–IIEP) in 1979 as a Consultant. He worked at UNESCO headquarters and its regional and field offices and retired from UNESCO in 2009 as the Regional EFA and Literacy Coordinator for the Arab States. During his service with UNESCO, he participated in several international and national conferences, prepared policy documents and provided technical assistance to the Member States of UNESCO. Dr Khan used to live in New Delhi and was an advisor to INRULED before his untimely demise in 2017.

1

CHAPTER

Educational Research: The Basics

Introduction

The world of education is dynamic where theories come and go, and any single theory is unlikely to operate and function in isolation. Learners are diverse and learning process is complex. Learners' success measured in terms of students' learning achievement does not depend solely on the teaching–learning process; it is rather influenced and determined by a multitude of factors such as social backgrounds, family background, personality, age, gender, location and so on. Thus, for allowing us to adapt to suit local and personal environments and conditions, education theories and practices need to be combined, tested and challenged (Babbie 2010).

We have all been to school and we know how and what we were taught in the classroom. But the problem is that we were taught the traditional theories and practices. New theories and technological advances have taken, and are taking, place. If our education practices are based exclusively on our learning experiences, without due reflection and recognition to these dramatic transformations, then our education systems run the risk of being outdated and not being forward-looking.

For making solid and meaningful decisions about education practices and policies, it is vital for us to better understand the fundamentals of teaching–learning process that guide the success of our education systems. Researchers are often asked or required to unearth ideas by undertaking valid and reliable research, which guides us to uncover these fundamentals to improve education standards and the quality of teaching.

Education research helps to shape our education systems. It helps us to understand how we raise our children and how we treat our teachers, managers, planners and policymakers. But what is education research and who conducts it? This chapter explores the aims and objectives of research in education. The chapter offers the reader the opportunity to learn more about research methods and data in both an academic context and professional context in order to give the reader a better understanding of the role and uses of research within the education sector.

The chapter presents an overview of the major educational research paradigms, shaping contemporary educational research. The chapter explains how epistemology—a branch of philosophy dealing with the study of the nature, origin and validity of human knowledge—relates and impacts the processes of policy formulation and research in the field of education. It examines the misconception of knowledge as education, and highlights how knowledge is an aspect of education and sources of knowledge together with their implications on the education research.

It is anticipated that the chapter will help orient students, educators and researchers to the newer paradigms to undertake uniquely powerful and insightful inquiries that contribute to transforming the landscape of education. The chapter will familiarise them with the methods used to conduct reliable research in education and to better recognise standards needed to ensure the validity of the research findings. The chapter is intended to help them distinguish between independent, peer-reviewed and unbiased research. Further, the chapter will help to better assess whether or not claims made about education policies are validated with proper evidence.

The discussion presented in the chapter furnishes a self-guided overview of educational research methodology. The thrust of resources outlined in the chapter is primarily on the relevance and quality of education statistics and the applicability of research findings. Their main emphasis is to address the most fundamental questions quite often asked about educational research, namely:

- How do we understand and interpret what the research in education says?
- How do we conclude or decide if the research in education is valid, reliable and trustworthy?
- How do we ensure if education policy claims are based on valid research evidence?

The chapter addresses each of these questions thoroughly in simple, non-technical, non-academic language and interactive formats. It provides to novice researchers an opportunity to advance their understanding of research in education through critical exploration of research language, ethics and approaches. The theoretical underpinnings contained in this chapter will help them critically review literature relevant to their field of interest and how research findings are useful in forming understanding of their work. It concludes by assessing how knowledge and education have affected the development and relevance of our education systems.

Paradigms of Research

The concept of social construction of rationality can be best understood by considering the seminal work of Thomas Kuhn's scientific paradigm. Kuhn, a science historian, developed this philosophical concept of science. The focus of his philosophy resulted in two approaches of thinking about what could be considered as 'scientific': the Aristotelian traditional teleological approach and the Galilean causal and mechanistic approach.

Definition and Meaning of Paradigm of Research

The term paradigm needs clarification. Thomas Kuhn introduced the concept of 'paradigm' into philosophical debate. He referred the term 'paradigm' to an overall theoretical framework. Kuhn (1962) defined paradigm in his book *The Structure of Scientific Revolutions* published in 1962 as.

> [A]ccepted examples of actual scientific practice, examples which include law, theory, application, and instrumentation together—[that] provide models from which spring particular coherent traditions of scientific research.... Men whose research is based on shared paradigms are committed to the same rules and standards for scientific practice.

He characterises a paradigm as 'an integrated cluster of substantive concepts, variables and problems attached with corresponding methodological approaches and tools…' A paradigm provides a conceptual framework for seeing and making sense of the social world. Bodgan and Biklen (as cited in Mackenzie and Knipe 2006) define a paradigm as 'a loose collection of logically related assumptions, concepts or propositions that orient thinking and research'.

> Paradigms are 'universally recognised scientific achievements that for a time provide model problems and solutions to a community of practitioners. Something like a paradigm is a prerequisite to perception itself' (Kuhn 1962).
>
> 'A paradigm is thus a comprehensive belief system, world view, or framework that guides research and practice in a field' (Willis 2007).

The underlying philosophy of paradigm is that it directs attention to recognised patterns of commitments, questions, methods and procedures that underlie and demonstrate direction to scientific work. Kuhn emphasises paradigmatic elements of research when he suggests that science comprises emotional and political as well as cognitive elements. By viewing its discourse as having different layers of abstraction, one can distinguish the underlying assumptions of a paradigm.

A paradigm, thus, determines the way for the formulation of a problem and suggests methods to handle it. It offers a way to consider the divergence in vision, customs and tradition. It enables us to regard science as having different sets of assumptions, commitments, procedures and theories of social affairs. It establishes the criteria according to which one selects and defines problems for inquiry and how one approaches them theoretically and methodologically.

From the viewpoint of cultural dimension, a paradigm is a man-made cultural object reflecting the dominant notions about scientific behaviour in a particular scientific community. Paradigm generates scientific approaches and procedures exemplary to the new generation of scientists as long as they do not oppose the paradigm.

Evolution of Paradigms in Research

Education is an interdisciplinary subject. Scholars with many disciplinary affiliations generally conduct research largely with background in psychology or other behavioural sciences. Only quite a few of them have a humanistic background in philosophy and history. Thus, there cannot be any prevailing paradigm in the very multifaceted field of educational research. Consequently, we present here educational paradigms under the following main headings.

Traditional paradigms

- Positivist paradigms
- Post-positivist paradigms

Relatively new paradigms

- The interpretive paradigm
- Critical paradigm

Traditional Paradigms

Traditional paradigms are well-recognised, tried and trusted first generation paradigms of educational research. As here is a plethora of social science research methods textbooks describing these paradigms, we, therefore, will not spend much time on these paradigms. We will just describe their salient features and the areas where they can be effectively used in educational research.

The traditional paradigms focused primarily on order and stability in society. They regarded society as a system of interrelated, interdependent parts, for example, social institutions or structures, population groups, organisations, a part may be family, education, economic, religion, women, men, the government and so on. Each part functions to maintain an orderly and predictable system, preserving social order. In traditional paradigm, there is normative consensus where members of society share a set of values and norms. The key tenants of traditional paradigms were Emile Durkheim, Talcott Parsons and Robert Merton.

The traditional paradigms have been criticised on several grounds. The paradigm may justify and legitimise the existence of a part of society, for example, poverty or unemployment. It helps to preserve status quo by overlooking or downplaying sources of tension and inequality. It does not allow any kind of questioning when it comes to the efficiency of a part. Finally, origins of social conflict and instability are not accountable or are considered dysfunctional.

Positivist Paradigm

By the mid-nineteenth century, philosophers like Auguste Comte developed positivism in sociology and John Stuart Mill empiricism in psychology. Comte emphasised observation and reason as means of understanding human behaviour. He argued that true knowledge is based on experience of senses and can be obtained by observations and experiment. They stated that certain knowledge is based on natural phenomena and their properties and relations. Positivism holds that valid knowledge (certitude or truth) is found in this a posteriori knowledge. Verified data (positive facts) received from senses are known as 'empirical evidence'; thus, positivism is based on 'empiricism'.

Positivism is well established in almost all universities and in specialised institutions engaged in advanced educational research around the world. Being 'scientific research paradigm', it endeavours to investigate, validate and predict patterns of behaviour. Students preparing their graduate dissertation or writing terminal research papers commonly use it to test theories and/or hypotheses. The concept is significantly useful in natural science, physical science and relatively less in the social sciences particularly in research studies involving very large samples. Essentially, the focus of the concept is on the objectivity of the research process (Creswell 2008). It primarily relies on quantitative methodology, employs experimental research methods comprising treatment and control and administration of pre- and post-tests to measure gain scores. Here, the researcher does not belong to study (research) site and he or she is the one who controls the research process.

Post-positivist Paradigm

Willis (2007) considers post-positivism as 'a milder form of positivism. It adheres more or less to the principles of positivism except that it permits more interaction between the researcher and his participants'. In addition to the research methods used in positivism, this post-positivism also employs methods such as survey research and qualitative methods such as interviews and participant observation (Creswell 2008). In fact, this paradigm is a modified version of the scientific method widely used in social sciences. An important strength of the paradigm, according to Taylor and Medina (2013), is that

> [I]t helps the researchers to analyse and produce objective and generalizable knowledge about social patterns, seeking to affirm the presence of universal properties/laws in relationships amongst pre-defined variables. This epistemology is manifested by quasi-experimental research designs that utilise treatment, outcome measures and experimental units but do not use random assignment to create comparison from which treatment caused change is inferred.

'Objectivity', 'validity' and 'reliability' are the key quality standards of this paradigm, which can be modified during the process of research with the help of 'triangulation' of data, methods and theories.

Relatively New Paradigms

Empirical research conducted by behavioural scientists, particularly in the Anglo-Saxon countries, in the 1960s and the early 1970s was criticised on the grounds of its dominance on a positivist quantitatively oriented paradigm. This did not allow other paradigms of a humanistic or dialectical nature to be used in educational research. The criticism was directed at those with a behavioural science background.

Interpretive Paradigm

Interpretive paradigm originated in educational research during the 1970s. Influenced strongly by anthropology, the aim of the paradigm is to understand other cultures from the inside—the culturally different 'others', by learning to 'stand in their shoes', 'look through their eyes' and 'feel their pleasure or pain'. The same is expressed in two seminal reports of United Nations Educational, Scientific and Cultural Organization (UNESCO)—Faure (1972) 'Learning to Be' and Delors et al. (1996)—'Learning: The Treasure Within' which proposed an integrated vision of education based on two key concepts, 'learning throughout life' and the four pillars of learning: to know, to do, to live together and to be.

The basis of this paradigm is construction of inter-subjective knowledge, especially with regard to its methods, validity and scope, and the distinction between justified belief and opinion. The paradigm uses ethnographic methods of informal interview, participants' observation and establishing ethically sound relationships for constructing trustworthy and authentic accounts of other cultures. When applied in educational research, it helps researchers to understand teachers' and students' life experiences and the cultures of classrooms, schools and the community they serve.

The interpretive knowledge construction is regulated by standards of trustworthiness and authenticity. In the authenticity criteria, Guba and Lincoln (1989) suggest:

- Credibility: Did the researcher undertake prolonged immersion in the field, check his or her interpretations with his or her informants and display a process of learning?
- Dependability: Did the researcher engage in open-ended or emergent inquiry?
- Transferability: Is there sufficient rich description for the reader to compare his or her own social context with the social setting of the research?
- Confirmability: Can the research data be tracked to their source?

In the authenticity criteria, they relate the ethics of the relationship established by the researcher with his or her participants. It includes:

- Fairness: Are the informants represented fairly?
- Educative: Did the participants benefit by learning about their social world?
- Catalytic: Did the participants benefit by identifying problems associated with their social work?
- Tactical: Did the research empower the participants to improve their social situation?

In educational research, this paradigm asks teachers to be reflective for knowing their students, with particular emphasis on the social, political, historic and economic forces shaping the pedagogies, curriculum policies and schooling system in which the teachers and students are deeply involved. It gives direction to them to adopt more student-centred pedagogies.

Critical Paradigm

In recent years, a key challenge facing our education system is how to empower our students and colleagues to become imaginative and critical thinkers capable of taking informed decisions for addressing question such as: Whose interests are not being served by particular social policies and practices? The critical research paradigm addresses this issue by equipping teachers to identify and transform socially unjust social structure, policies, beliefs and practices. It aims at identifying ‘imbalances, inequalities and injustices’ that plague our education systems.

According to Taylor (2010),

> [A]pplied to education, critical inquiry in education focuses first on raising the conscious and awareness of teachers about established values and beliefs that underpin their seemingly natural teacher-centred classroom roles. Once this process is underway, critical theory is introduced (e.g., critical pedagogy, cultural inclusiveness, social justice) that stimulates teachers’ creative thinking about designing curricula and assessment that are more student-centred, inquiry oriented, culturally sensitive, community-oriented, socially responsible, etc.

Paradigms in Educational Research

Since research paradigms influence both the strategy and the manner the researchers construct and interpret the meaning of the reality, they constitute a crucial element in the research study. The research paradigms have solid philosophical foundation and dispose the researcher’s point of view on the reality as given by nature or constructed by human agency. Researchers have been divided for long into two camps depending on the research paradigm: the followers of the quantitative methods and the qualitative ones. Supporters of each camp are arguing on the superiority of each method. The quantitative methods rest on numbers while the qualitative camp uses words or narratives. It appears an ongoing debate on numbers versus words. In recent years, a third camp supports mixed methodology as more pragmatic approach to research.

This section attempts to describe methods that are generally used in each strand of research. Each approach has its advantages and detriments, and is more suitable to answering certain kinds of questions.

Educational researchers use diverse scientific methods to study the best practices in teaching. Let us consider this case: ‘When given positive encouragement, do students learn better or worse?’ For answering this particular question, the researcher compares a group of students who are given positive encouragement (experimental group) with a group of students given no encouragement (control group). He or she then compares each group of students on important learning variables such as their performance, effort and motivation. Upon the completion of data analysis for both the groups, the researcher suggests a concrete and definitive answer about the impact of positive encouragement, negative criticism or no encouragement on the academic performance of experimental group of students. Educational research that uses the scientific method to collect data enables researchers to offer a definite answer about best practices in teaching. The underlying objective of educational research is to collect reliable information, analyse it and develop new knowledge about the teaching–learning situation to improve educational practice.

In an educational setting, researchers basically use two major types of research models or research paradigms: qualitative and quantitative paradigms. The qualitative approach to research helps the researcher understand a phenomenon from a closer perspective. The quantitative approach inclines to approximate phenomena from a larger number of individuals by means of survey methods.

In educational research, there are several reasons that you would use for choosing a paradigm. These reasons are:

- World view or assumptions of each paradigm
- Training and experience
- Psychological attributes
- Nature of the problem
- Audience for the study

Each of these reasons is discussed separately for qualitative and quantitative research methods in the following.

A: Qualitative Research Methods

Qualitative research paradigms are also known as constructivist, naturalistic, interpretive, post-positive or post-modern perspective as advanced by Ditthey, Kant, Wittgenstein (latter), Foucault, Miles and Huberman.

The thrust of qualitative approach is on describing a phenomenon in a deep comprehensive manner. Qualitative research relates to analysis of human behaviour and the underlying reasons behind it—what people do and why? The method uncovers the deeper meaning, significance and seriousness of human behaviour and experience. Examples include contradictory beliefs, behaviours and emotions.

The analysis of qualitative research is basically non-quantitative. Here the researcher observes people: what they do and say. Then he or she collects, analyses and interprets data. The research is relatively subjective and as such relies primarily on very different methods of collecting information such as individual in-depth interviews and focus groups discussions.

Mostly, the approach is used for a small number of participants taking part in the research study. This is obviously for reasons of its high demand for both resources and time. Interviews can take the form of highly structured and guided by open-ended questions or less structured and guided by conversational interview. The research gives a rich and complex understanding of people's experience but not in obtaining information that the researcher can use to generalise the whole population or other larger groups. However, qualitative research leads to larger studies and deeper understanding that can inform theory, practice and specific situations.

Benefits of Qualitative Methods

The use of open-ended questions and interviews permits researchers and practitioners to understand how individuals are doing, what their experiences are and recognise important antecedents and outcomes of interest that might not surface when surveyed with predetermined questions. Although qualitative research can be thought of as anecdotal, when pooled across a number of participants it is a very important approach that equips researchers to acquire a conceptual understanding and evidence that certain phenomena are occurring with particular groups or individuals. The method:

- Permits identification of new and untouched phenomena
- Provides a deeper understanding of mechanisms
- Offers one-on-one and anecdotal information
- Offers verbal information that may sometimes be converted to numerical form
- Reveals information that would not be identified through predetermined survey questions

Limitations

- Results and findings cannot be generalised to the general population
- Very limited scope in applying statistical methods
- Difficult to assess relations between characteristics

B: Quantitative Research Methods

Quantitative research methods are also known as traditional, positivist, experimental or empiricist as advanced by authorities such as Comte, Mill, Durkheim, Newton and Locke.

The focus of quantitative approach to gathering information is on explaining a phenomenon across a large number of participants of the research study thereby allowing an overall view of summary characteristics across groups or relationships. Quantitative research deals with a systematic investigation of quantitative properties—how many people do these things and how often? In this type of research, the researchers collect and transform data into numbers for statistical computations and finally for drawing conclusions.

The basis of quantitative research is numeric figures and numbers, and the underlying idea is to compare past records with the present records and prepare quantitative forecasts and projections for the future period. Quantitative research uses extensively statistical and mathematical models, theories or hypothesis related to phenomena.

In quantitative research, the researcher uses extensively the structured questions for which he or she predetermines the response options for large number of respondents. For instance, the researcher can analyse primary enrolment in a province or in the country in 2015. He or she can compare this figure with the enrolment for the last five years and prepare projection for the period 2016–2025.

The survey in this approach comprises large number of participants and a scientific application of statistical methods to pinpoint the overall patterns in the relation of processes. The survey can be used across the groups. For example, the same survey can be used for both the experimental group and the control group. This makes comparison between these groups simple to assess the impact of outcomes of interest and determine the influence of an intervention. The researcher can easily survey people a number of times, thereby allowing the conclusion that a certain trait influence specific outcomes.

Benefits of the Quantitative Approach

The approach is highly scientific as it allows the researcher to generalise the findings across a large group of individuals. For example, if policymakers wanted to know the impact of an intervention, they would likely require some evidence that this intervention actually works. By conducting interviews of few individuals or focus group with 30 matches of control and experimental participants might be reflective of specific cases in which the intervention worked, it would not provide strong evidence that this particular intervention is overall meaningful or beneficial. The approach can provide evidence to support whether an intervention is successful or not. The approach:

- Helps in gathering information from a relatively large number of participants
- Can be carried out for large samples
- Allows comparison
- Allows generalisation to broader population
- Provides numerical or rating information
- Informative for initiating policy or guidelines
- Lends to statistical techniques that allow determining relations between variables

Limitations

- Difficulty in recognising new and untouched phenomena
- Caution in interpretation without a control group

C: Mixed Research Methods

When analysing a situation and its possible causes or solutions, it is often beneficial for the researcher to use multiple perspectives whenever possible. Researchers can do this by combining both the quantitative and qualitative methods in their research endeavours. A survey that collects qualitative data can be done together with research that collects quantitative data. The underlying purpose of using mixed methods is that it enables the researcher to see a bigger picture of not only what the causes or results are but may also probe and identify into what overall factors and influences surrounding the issue may be. Thus, this method allows the researcher to see both objective (quantitative) and subjective (qualitative) sides of his or her research problem. While some research topics may be inclined to use one type of method or another, experts strongly suggest that using a mixed methods analysis, when possible, can lend itself to superior results over using one sole type of research.

In summary, the qualitative and quantitative approaches to research bring a different perspective of situations or phenomena. These two main approaches to educational research are extremely informative, especially if used in unison. Each approach has its benefits and detriments, and being aware of the methods used to gather information can help practitioners and policymakers understand the extent to which research findings can be applied.

Readers are informed that the earlier description presents a very brief account of the tremendous and fast-changing field of educational research but the description is not exhaustive. For instance, approaches such as 'mixed methods' widely used by post-positivists into their predominantly objective research have not been dealt with. The critical paradigm helps them recognise that education neither works in isolation nor it takes place in a social vacuum. They have begun to understand that educational practices largely depend on cultural and social contexts in which they operate. They are also not neutral to educational policies. Thus, the two main paradigms are not exclusive, but complementary to each other.

Interdependence of Paradigms

We have noticed the distinction between two key paradigms of educational research with different basis of knowledge. The positivist paradigm is functional, structural, objective, rational, goal-directed, manipulative, hierarchical and technocratic whereas the post-positivist paradigm is humanistic, consensual, subjective and collegial. The first approach emanates from the classical positivism. The second one, more popular now, is partly derived from the Habermas' (McCarthy 1979) critical theory of communication of the Frankfurt school. The first approach is linear and consists of a straightforward rational action towards a preconceived problem. The second approach leaves room for reinterpretation and reshaping of the problem during the process of dialogue prior to action and even during action.

Keeves (1988) argues that the various research paradigms employed in education, the positivist, the phenomenological and the ethnographic-anthropological are complementary to each other. Describing the unity of educational research, he distinguishes between paradigms and approaches, and contends that there is, in the final analysis, only one paradigm but many approaches.

For example, the researcher can observe the teaching–learning process, record on paper or prepare a video. The observations can be quantified and the data can be analysed using a statistical method. He or she can study the content within the context of national traditions and the philosophy underlying curriculum construction. The researcher can design a comparative, cross-national perspective for studying both the teaching–learning process and its outcomes.

Depending upon the objective of a particular research project, the researcher can lay more emphasis on either the qualitative or quantitative paradigm. It means that qualitative and quantitative paradigms are more often than not complementing each other. For example, the researcher cannot arrive at any reliable and valid information about a school regarding the achievement of competence level in a subject simply by visiting a number of classrooms and thereby trying to collect impressions. In this case, the researcher can use sample survey technique of International Association for the Evaluation of Educational Achievement (IEA). But when the purpose is to do accounting for factors responsible for differences between schools, such surveys are not much useful. In situation such as this, the research needs different kind of qualitative information.

Educational planners, policymakers and administrators always look for generalisations, rules and regulations that they can apply in schools, colleges and universities having children from diverse socio-economic backgrounds. They are very much concerned with collectivity rather than in an individual child. On the other hand, teachers and other classroom practitioners are hardly interested and much helped by generalisations which in principle apply to the entire system rather than a given child or a given education institution.

Knowledge and Education: The Relationship

As an important branch of philosophy, epistemology studies and examines the nature, origin and limits of human knowledge. Its importance is not only recognised by philosophers but it is fundamental to education that transmits and imparts knowledge to all of us. Knowledge is significantly linked to intellectual development, acquisition of human skills (both technical and life skills) and, to a large extent, the development of appropriate value system. This is the basic essence of the business of education. Seen from this perspective and within the context of educational research, it is quite meaningful and essential to understand and ascertain the essence of knowledge, its nature and how best it could be used in the process of educational research. According to White (1982), one main purpose of the examination of the nature of knowledge by philosophers of all ages has been to enquire into the original, certainty and extent of human knowledge. The extent of knowledge does not imply 'how much we actually know' but 'how much we can know'. The answer to this question depends on what do we mean by the term 'knowledge'. This is explained next.

The diversified opinions of epistemologists have guided theories and concepts of knowledge. As a consequence, we do not have a universally acceptable standard definition of the concept of knowledge. However, concerted efforts have been made by renowned philosophers, educationists and international organisations actively involved in the field of education, such as UNESCO, UNICEF, the World Bank and OECD, to differentiate knowledge from other cognitive notions such as belief, understanding, perception, sensation, truth and reason with a view to understand the concept of knowledge and accept and recognise their interrelationship. These efforts, as expected, have led to the formation of a common ideology that runs through the different activities that are involved in knowing, such as types, forms, sources, theories and criteria of knowledge.

It is education that equips the individual with knowledge, skills and sound morals so as to be able to function as a productive citizen. What a learner can know is certainly influenced by the different aptitude of the learner. This relationship between education and knowledge suggests for proper epistemological understanding of the educational system and the philosophical foundations of the society. Knowledge alone cannot reflect the aim, goals and aspirations of the society. Thus, our education systems should include in curriculum those contents that help knowledge creation essential for the sustainable development and modernisation of our societies. Educators who are well grounded in the concept of knowledge are in a better position to determine and deploy the proper methods of teaching knowledge at various levels.

It has been argued that knowledge is a necessary condition for being educated. So an educated man is a knowledgeable man but a man with a lot of knowledge could be miseducated if other aspects of education are not adequately developed. As such, knowledge is necessary but not a sufficient condition for education. An educated man needs more than just the acquisition of knowledge. An educated man must have a body of knowledge with the required level of conceptual scheme that will help him or her raise his or her knowledge above the level of an assemblage of disjointed facts and insights. He should understand principles and organisation of facts while his or her knowledge must reflect in his or her ways of looking at and the assessment of issues as they affect him or her and his or her environment.

An educated man must practise the virtues of knowledge. He should be committed as an insider to the knowledge, understand the processes of the acquired knowledge and practise the utilitarian values (Peter 1979). Knowledge and education should be deployed for positive and uplifting uses by the person and the society instead of deliberate manipulations for selfish and non-utilitarian gains, which is the hallmark of many societies whose 'educated' elites have become hindrance to positive development efforts.

For knowledge to be meaningful, it must always have a reference to the truth. Rene Descartes (cited in Cunning 2007) in his or her search for certainty devised a method of systematically doubting everything he thought he knows until he could reach a conclusion where he cannot doubt anymore. Knowledge should be firmly rested on truth that cannot be doubted. It should stand on its own strong foundation of truth. In addition to truth, knowledge must also be real. Our education systems should stand alone on the reality of the society properly encapsulated in the bodies of knowledge that reflects the truism of the society. The process of knowledge promotion in the education system should be consciously systematic and purposeful. We can say that knowledge is a cognitive-based activity within the social milieu that is facilitated by the human intellect, which seeks for the truth of reality.

Criteria of Knowledge

In common parlance, people use 'belief' and 'knowledge' interchangeably. We tend to use the word 'belief' when we do not know all what we should know about an object because our knowledge about the object is not definite. Belief is a prelude to knowledge that generally depends on social and psychological factors.

> The effect of the belief system on knowledge is such that in some cases it could extend the frontiers of knowledge by attempting to legitimize beliefs. On the other hand, the belief system can become restrictive by projecting negative attitude to the use of certain types of knowledge. (Dash 2005)

Since belief does not comprise objective truth, it cannot be considered as knowledge but it could be regarded as a source of knowledge development.

Truth is a necessary condition for knowledge.

> For those who see knowledge as static, teaching becomes 'telling' or 'dumping' of knowledge in the learners where they are bombarded by avalanche of knowledge while teaching becomes teacher-centred. On the other hand, those who see knowledge as evolutionary tend to engage learners in continuous search for certainty and truth with the active participation of the learners. Truth then is considered as one of the criteria for knowledge. (Omatseye 2004)

Another important criterion of knowledge is 'utility'. Knowledge can have utility value for social welfare if it helps in solving problems that could bring about improved living. Knowledge should always create further knowledge. The utility value of knowledge has implication. It should bring about positive change in the overall disposition of the individual. Also, the referent or object of knowledge should have existence or reality that can be attributed to it. For example, Mrs Vincent cannot be a male teacher because the attributes of a 'Mrs' do not belong to a male, although a 'Mrs Vincent' may exist. Knowledge must comprise certainty that makes it real to the learner, which helps in upholding the reliability of knowledge.

Sources or Types of Knowledge

In epistemology, a common concern with respect to knowledge is what sources of information are capable of giving knowledge. There are several sources of knowledge by which people acquire knowledge or justified belief. These sources are:

Perception/Senses: that which can be perceived through the experiences of the senses. The view that experience is the primary source of knowledge is called 'empiricism'.

This source refers to knowledge disclosed to man by the use of whatever he believes in or through supernatural means. This could be through visions, dreams, trance or meditation.

Knowledge so acquired cannot be investigated or tested. This source of knowledge is basically based on one's faith especially in religious or spiritual activities.

Reason: It can be considered a source of knowledge, either by deducing truths from existing knowledge, or by learning things a priori, discovering necessary truths (such as mathematical truths) through pure reason. The view that reason is the primary source of knowledge is called 'rationalism'.

Reasoning can be either in the form of deduction or induction. Basically, deduction is used in mathematics and is the more certain of the two as it involves 'drawing out' valid conclusions from previously known facts, for example, all cats are animals, Sam is a cat, so Sam is an animal. Induction, on the other hand, is usually used in science and is less certain. For example, I drop this book and it falls, so it is likely all dropped books (and indeed things) will fall. It is important to note that both forms are usually dependent on 'sensation' to give us the initial facts or ideas in the first place.

Introspection/Intuitive: knowledge of one's self that can be found through internal self-evaluation. This is generally considered to be a sort of perception. For example, I know I am thirsty or nervous.

This is a product of insight or inspiration. It is just a flash in mind as a result of the power of the intellect abstracting within the consciousness. This source of knowledge offers solution to a problem that one has been contemplating for a while. It is mostly random or accidental while some say is mystical power or a sudden 'mental spark'. This source cannot be evaluated scientifically; however, it still remains a source of knowledge through which many new knowledge have emerged.

Authoritative: Other people are continual sources of information. Such information, however, is always second-hand knowledge. It is all 'hearsay'. This type of knowledge is derived from sources that have consistently been verified to be true.

The implication of these sources of knowledge in the educational system is such that it is important first to reflect on which path and part of knowledge that is most vital. These sources of knowledge provide guidelines to education decision-makers and policymakers in the design of the curriculum that is a reflection of the type of society they envision in their philosophy of education. While divine knowledge is more of spiritual of development, empirical and rational knowledge are more germane to modern life for they challenge us to think beyond the surface and equip us to create new possibilities that enhance life. While conducting educational research, a good knowledge of the sources of knowledge is important as it helps in understanding the structure and how best that knowledge could be taught. This is important because the educational system is dominated by intellectual activities of 'knowing how' and 'knowing that'.

Educational Research as Scientific Enquiry

As a student, teacher or administrator, consider how many times you have heard, 'evidence-based practice' or 'according to the research'. It seems that every new idea in education is research-based, but what does that really mean?

Not all educators have an understanding and appreciation of research. For some, research is something that is important only for faculty members in colleges and universities. Although it is true that college and university faculty members value and conduct research, personnel such as school psychologists, principals, school board members, adult educators, college administrators and graduate students in other educational settings also read and use research.

It is for this reason that the concept of 'educational research' differs from one source to another. We have seen earlier that education in its simplest form refers to a search for knowledge. Some researchers conduct a systematic search for reliable and pertinent information on a specific topic. But in reality, educational research is a scientific enquiry. It is an art of scientific investigation. Research refers to a careful investigation or inquiry through special search for new facts in any branch of knowledge. It is a movement from the known to the unknown—a journey of discovery (Siniscalco and Auriat 2005).

> 'Research comprises defining and redefining problems, formulating hypothesis or suggested solutions; collecting, organizing and evaluating data; making deductions and reaching conclusions; and at last carefully testing the conclusions to determine whether they fit the formulating hypothesis' (Woody 2011).

> 'A cyclical process of steps that typically begins with identifying a research problem or issue of study. It then involves reviewing the literature, specifying a purpose for the study, collecting and analysing data, and forming an interpretation of information. This process culminates in a report, disseminated to audiences that is evaluated and used in the educational community' (Creswell 2002).

A more specific definition is given by UNESCO-IIEP (Postlethwaite 2005). Accordingly,

> 'Education research implies research activities designed and undertaken towards unravelling educational problems or bringing about improvement in the teaching–learning process. Educational research is an activity directed towards the development of an organised body of scientific knowledge about the events with which educators are concerned' (ibid.).

UNESCO-IIEP further states:

> [E]ducation research examines teaching and learning processes and the human attributes, interactions, organisations and institutions that shape educational outcomes. It seeks to describe, understand and explain how learning takes place throughout a person's life and how formal, non-formal and informal contexts of education affect all forms of learning. It embraces the full spectrum of rigorous methods appropriate to the questions being asked and also drives the development of new tools and methods. (Ibid.)

Research is:

R = Rational way of 'thinking' on an issue or concern
E = Expert and exhaustive 'treatment' of the concern
S = Search for 'solution'
E = 'Exactness'
A = Analysis of 'data'
R = Relationship of 'facts'
C = Careful 'recording', critical 'observation', constructive 'attitude'
H = 'Honesty' (Truth)

It is comparing and contrasting objects, gathering original (first-hand) information, examining critically the issue(s) or concern(s) and analysing, interpreting information pertaining to individuals or organisations and identifying solutions. Educational research is a careful in-depth analysis of a specific problem, concern or issue using both qualitative and quantitative scientific methods.

Educational research signifies thinking and reflection on a given issue in an education setting.

It is not simply theorising; it comprises several activities. It defines and redefines educational problems, formulates hypothesis or suggests solutions, collects, organises and evaluates data, makes deductions and suggests conclusions and finally tests carefully the conclusions to determine whether the conclusions fit the stated hypothesis.

However, research in education is governed by multiple and complex factors. Evidence shows that there has been an enormous increase in educational problems in recent years as a result of rapid changes in technology. Table 1.1 highlights the several problem areas and gives examples of the nature of problem. It is possible only through systematic research to address these problems and other critical educational challenges of society that are highly technical and changing fast.

In education, our main concerns are as follows: What is the best way to teach? How do we know that our concerns and beliefs are justified and correct though? Is there any way to support our beliefs with more valid evidence than just that we believe it to be true? That is, is it true from our own experiences, or it is stated by an expert that it is true? Thus, there are always two opposite beliefs and concerns; so which is true?

Educational research seeks to answer to all of these questions. Research in an educational setting is of paramount importance because it is the 'sovereign solvent' of educational problems. Through educational research, we get better and responsible children, better curriculum, better methods of

TABLE 1.1 Categories of Education Problems

Problem Category	*Example*
Psychological	Learning theories, factors affecting learning, remembering and forgetting, motivation, maturation, child development and growth and so on.
Philosophical	Logical analysis of concepts, values and worthiness in education, educational aims, moral judgements, methods of reasoning, meaning, nature and sources of knowledge and so on.
Evaluation	Continuous assessment, tests instruments, examinations (both internal and external), item analysis, students' report cards and so on.
Content	School subjects, factors affecting choice of content, mode of organising content and so on.
Methodological	Methods of teaching, a comparison of these methods, teaching styles, teacher effectiveness, teaching practice, micro-teaching, institutional media and so on.
Administrative	School financing, discipline, school records, classroom management, leadership styles, recruitment and deployment of staff and so on.
Sociological	School–community relationship, teacher–pupil interaction, interpersonal relations within school, classroom behaviour of students, students' unrest, cultism and so on.
Historical	History of institutions, programmes, places or persons of educational interest.

Source: Adapted from Nwana (2005).

teaching, better school administration, organisation and supervision and above all a better understanding of the world in which we live. A well-planned and well-executed educational research programme saves time, money and energy, and avoids a lot of frustration and takes the nation to the path of progress and glory.

Educational research, thus, implies a systematic and scholarly application of the scientific method to the solution of educational problems. Educational research is an activity, which is directed towards development of a science of behaviour in any educational institution. The ultimate aim of such a science is to provide knowledge, which will permit the educator to achieve the stated goals by using the most effective method. Education must therefore keep pace with the rapid development in the world.

Education research examines the teaching–learning process in relation to human attributes, interactions, organisations and institutions that shape educational outcomes. Scholars in the field describe, understand and explain how learning takes place throughout a person's life and how formal and informal contexts of education affect all forms of learning. Education research encompasses the full spectrum of rigorous methods appropriate to the questions being asked and drives the development of new tools and methods. In doing so, it attempts to answer educational questions that can be investigated in a satisfactory manner by using tools and methods, which permit such satisfactory investigation and the utility of results generated from such inquiries. Since most questions in education come from people with different conceptions of social reality, a multitude of paradigms (see earlier) have been suggested for practitioners to determine the criteria according to which one would select and define problems for inquiry (Dash 2005).

In simple terms, educational research comprises those activities or processes which allow one to systematically test and/or obtain a body of information, data or knowledge about teaching/learning or conditions which affect teaching and learning.

Research Methods and Research Methodologies

It has been noticed that students preparing their master's dissertation or doctoral thesis tend to confuse the difference between research methods and research methodology. Key differences between research methods and research methodologies are explained next.

Research Methods

Research method is a broad term. It refers to all those methods and techniques that are followed to conduct an investigation. These are the methods researchers use in performing research operations. It is a procedure adopted for research, which includes sampling techniques, development of data collection tools, administration of tools, organisation of data, procedure of data analysis and so on.

Methods (also known as techniques) are used to reveal the existence of, identify the 'value' significance or extent of, or represent semantic relationships between one or more concepts identified in a model from which statements can be made. Thus, in order to know the range of available methods that can be selected, you must know about the theory being applied. We must realise that behind every method there is always a theory (Creswell 2003).

Since the objective of research is to explore or arrive at a solution for a given problem, the available data and the unknown aspects of the problem have to be associated or related to each other to make a solution possible. Keeping this fact into consideration, Kothari (2004) classifies research methods into the following three distinct groups:

Group I: includes those methods dealing with data collection.
Group II: includes statistical techniques establishing relationships between the data and the unknown variables.
Group III: includes techniques for assessing accuracy of the results obtained.

Research methods in Groups II and III are generally regarded as the analytical tools of research.

There are several methods researchers use to get information. The most frequently used methods are: literature searches, talking with people, focus groups, personal interviews, telephone surveys, mail surveys and Internet surveys. Here we explain them briefly. A detailed description of each of these methods is given in subsequent chapters.

Literature search: It refers to reviewing all readily available materials on topic of research. These materials can include internal organisation information, relevant publications, newspapers, magazines, journals, annual reports, online databases and any other published materials. This is very inexpensive method of gathering information. The method often does not yield timely information.

Talking with people: It is a useful way to get information at the initial stage of the research project, particularly the information which is not publicly available or that is too new to be found in the literature. Meeting with prospects, exhibitions and shows, seminars and so on are good examples.

Focus groups: This research technique is a preliminary technique which helps the researcher to explore people's ideas and attitudes. Focus groups can be carried out within a couple of weeks with relatively less cost. But a key limitation of this technique is small sample which may not represent the population in general.

Personal interviews: This method is a relatively better method to gather in-depth and comprehensive information. Questions are put in a written questionnaire and the interviewer records the answers. Personal interviews are generally used only when participants of the research study are not likely to respond to other survey methods.

Telephone surveys: This is the fastest information gathering method from relatively large samples (100–400 respondents). The method uses prepared script for the interview and more or less the same as written questionnaire. It allows the opportunity for some opinion probing. They usually last less than 10 minutes. They can be completed in two to four weeks.

Mail surveys: It is a cost-effective technique of gathering information. They are quite useful to gather information for large sample sizes or when the sample comes from a wide geographical area but they take over twice as to complete (8 to 12 weeks). Because there is no interviewer, the possibility of interviewer bias does not exist. The main limitation, main disadvantage, is its inability to probe respondents for more detailed information.

E-mail and Internet surveys: These surveys are relatively new and little is known about the effect of sampling bias in Internet surveys. While it is clearly the most cost-effective and fastest method of distributing a survey, the demographic profile of the Internet user does not represent the general population, although this is changing.

The selection of the research method is critical for what conclusions the researcher can draw/make about a phenomenon. It affects what he or she can say about the cause and factors influencing the phenomenon. It is, therefore, important to choose a research method, which is within the limits of what you can do. Time, money, feasibility, ethics and availability to measure the phenomenon correctly are examples of issues which can constrain the research.

Research Methodology

Research methodology compared to research methods is a larger canvas which subsumes method of research: it includes the instance and the philosophy of the researcher, that is, the perspective on social reality and truth and generalisations of research finding and so on which would justify the selection and use of the method of research. It is defined as a highly intellectual human activity used in the investigation of nature and the matter and deals especially with the manner in which data are collected, analysed and interpreted.

Research methodology means the manner in which a research problem is solved systematically. It refers to the science of studying how research is being carried out scientifically. Here we study the several methods a researcher generally adopts in studying his or her research problem along with the logic behind them. It is, therefore, necessary for you to know both: the research methods and the research methodology. As such, an understanding of research methodology will facilitate our understanding of basic statistics.

Research methodology tells us how to develop certain indices or tests, how to compute the several measures of central tendency (such as mean, median and mode) and measures of dispersion (such as range and standard deviation [SD]) or chi-square and how to apply particular research techniques. But the important things that we must know fully are: (a) which particular method to choose and (b) the interpretation—what would they mean and indicate and why.

Further, it is also necessary to know the assumptions associated with various techniques and the rationale and criteria for selecting certain techniques and procedures. It means that it is critical for the researcher to design his or her methodology for a given problem as the same may differ from problem to problem. Suppose we wish to design a school building. We cannot start the construction of the building just like that. The basis of our choice and decision for the selection of a particular design should be based on the systematic and conscious assessment and evaluation of all key school building design standards. In other words, we have to evaluate why and on what basis we select the particular location of school (school mapping), its size;, number and location of doors, windows and ventilators; the particular construction materials to be used and not others; and so on. Likewise, in our research, we have to expose the research decisions to evaluation before they are implemented. We have to specify very clearly and precisely what decisions we opt and why so that they can be evaluated by others also. That is:

- Why should we collect certain data?
- What data should be collected?
- From where we should collect it?
- How should we collect it?
- How should data be analysed?

When we commission a research study to identify answers to a given question, we must ensure that the process:

- is undertaken within a framework of a set of 'approaches';
- uses procedures, methods and techniques that have been tested for their 'validity' and 'reliability' and
- is designed to be 'unbiased' and objective

Here, approaches refer to, for example, qualitative, quantitative and the academic discipline in which we have been trained. Validity implies that standard research methods and tools have been used, and correct procedures have been followed in the investigation to find answers to a given research question. Reliability means the quality of a measurement procedure that provides repeatability and accuracy. Unbiased and objective analysis mean that the analysis carried out in the research enquiry is unbiased and conclusion drawn to the best of our ability and without introducing our own vested interest. 'Bias is a deliberate attempt to either conceal or highlight something'. Adherence to these three criteria enables the research to be called 'research'.

We should remember that research methods do constitute a part of the research methodology. The scope of research methodology is wider as compared to research methods.

Research methodology is the way in which research problems are solved. It is a science of studying how research is conducted scientifically. Under it, the researcher acquaints himself or herself with the various steps generally adopted to study a research problem, along with the underlying logic behind them.

Table 1.2 describes the purpose of education research methodologies.

TABLE 1.2 Purpose of Education Research Methodologies

Exploratory	*Descriptive*	*Analytical*	*Predictive*
Exploratory research commits the researcher and begins when few or no previous studies exist to analyse a particular problem or issue. Such a study explores patterns, hypotheses or ideas that can be tested and as a basis for further research. The research tools commonly used in such studies include case studies, observation and reviews of previous related studies and data.	Descriptive research is undertaken to identify and classify the key elements or the salient features of the subject, for example, number of school days lost because of teachers' strike or flood. Quantitative techniques are most often used to collect, analyse and summarise data.	Analytical research often extends the descriptive approach to suggest or explain why or how something is happening, for example, underlying causes of dropouts and repetition in primary schools. A typical feature of this research is that the research enables to locate and identify the different factors (or variables) involved in the study.	The predictive research enables the researcher to hypothesise and speculate intelligently on future possibilities. These studies are grounded on a close scientific analysis of available evidence of cause and effect, for example, predicting in which year and in which (province/state) Education for All (EFA) goals might be achieved.

Source: Authors.

It is clear from Table 1.2 that different kinds of research questions require different kind of approaches. Various methodologies have emerged to deal with different types of research. They can be broadly categorised as follows:

Type of research	Methodology
• Historical	Qualitative
• Comparative	Qualitative
• Descriptive	Qualitative
• Correlation	Quantitative
• Experimental	Quantitative
• Evaluation	Qualitative
• Action	Qualitative

A question that immediately comes to our mind is: How do I know which research methodology to choose? A reasonable answer could be that you ask your supervisor(s) or others who have worked directly in the application domain that you are interested in studying, and/or look at significant papers in your field and attempt to determine what methodology they are using.

In conclusion, research is a structured enquiry that utilises acceptable scientific methodology to solve problems and create new knowledge that is generally acceptable. Mere collection of facts, from reference books, readings in the library, historical documents, questionnaires or even the Internet, is not research, unless the findings and conclusions obtained from the analysis of the data are used meaningfully to address and solve problems.

In addition to these two main concepts—research methods and research methodologies—there are other key concepts often referred to in the process of research. These concepts are explained in the following.

Research Philosophy

'Research philosophy' is associated with clarification of assumption about the nature and the source of knowledge. The basis of all studies is some kind of assumptions about the world and the ways of understanding the world. There is no agreement among philosophers about the most appropriate ways of understanding the world; therefore, it is necessary that you clarify the philosophy you have chosen to understand your research problem.

'Positivism' and 'phenomenology' are the two main contrasting research philosophies (see earlier) related to research in social sciences. Positivism is an objective approach. It relies on facts and quantitative data. Phenomenology, on the contrary, takes into account subjective human interests and focuses on meanings rather than hard data. You have to specify in your dissertation which philosophy you are following.

Research Approach

'Research approach' is basically divided into two categories: inductive and deductive. If your aim was to find answer to specific research question(s) formulated in the beginning of the research process, you would be following an 'inductive approach'. Alternatively, if you choose to achieve research objective(s) via testing hypotheses, your research approach can be specified as 'deductive'. Your choice of selection between the two will depend on a set of factors such as the area of study, research philosophy, the nature of the research problem and others.

Research Design

Research design can be exploratory or conclusive. If the purpose of your research is merely to explore the research problem and not to produce final and conclusive evidences to the research problem, your research design would be 'exploratory'. On the contrary, the aim of 'conclusive research design' is to provide final and conclusive answers to the research question.

Conclusive research can be further divided into two sub-categories: descriptive and causal. As the name suggests, descriptive research design describes specific elements, causes or phenomena in the research area. Causal research design, on the other hand, is carried out to study cause and effect relationships.

Strengths and Weaknesses of Research Methodologies

Each research methodology has both strengths and weaknesses. Some of these differences are apparent while others are not. Thus, when we are planning to choose a given research methodology, it is imperative for us to consider several issues. In particular, there are several key questions that can serve as a framework for evaluating research designs. The main questions are discussed in the following.

Research Question

The research question is by far the most important question to be considered when selecting and evaluating a research design. Thus, the key research question should be written appropriately prior the researcher selects a given research methodology. Once the key research question is articulated, only then the appropriate methodology should be chosen based on that question. The researcher should be careful when selecting a particular research methodology. He or she should avoid and discard a preferred methodology as a framework to guide his or her research.

For example, a researcher wants to know if students' pronunciations in Chinese and English as foreign languages are better in private schools after two years of studying. For this particular study, his or her research question will be: What is the relation between studying Chinese versus studying English and foreign language pronunciation after two years of study?

For investigating this particular research question, the immediate task of the researcher then would be to select and decide the research design most appropriate for the study. For this particular example, the researcher can easily ignore and eliminate several options. An experiment would be impossible since students cannot be randomly assigned to Chinese or English classes. Or the researcher might decide that qualitative and micro-genetic studies are not suitable since he or she is not interested in the processes or developmental trends that occur over time. Thus, the researcher has to carefully frame such questions and must consider their final selection prior to the selection of a research design.

Sample Design

When selecting a particular research design, it is also necessary for the researcher to reflect and pay attention to the nature of his or her sample. This question is highly significant because some methodologies are challenging to implement with certain populations.

Most studies that use survey-based methodologies presume that the respondents have sufficient ability to read, understand and respond to the items contained in the survey (Chinn 2006). Consider this example: if the sample of the researcher consists of young primary schoolchildren (Grade 1 or Grade 2 children), or their parents with impaired visual abilities, then this might prevent the use of a self-administered survey. In addition, if the researcher was studying a large sample, with more than 1,000 participants, in many cases this would prohibit the investigator from implementing single-subject designs, since the sample is so large.

Availability of Resources

Diverse resources are required for completing research studies. A general tendency among researchers is to underestimate often the cost, personnel and time required for completing educational research studies. For instance, a college student doing a small study for a research methods course (master's degree dissertation) will certainly not have the same resources available as an experienced investigator with a multimillion dollar grant. Similarly, some research methodologies require more personnel than others. Time is another equally important resource that often affects the type of methodology the researcher selects for a particular study. Thus, if fewer resources are available, a researcher might not be able to use the ideal methodology to conduct a study.

Intended Audience

The benefits that accrue from different kinds of research studies vary from one type of audience to another. For instance, action research is considered to be highly appropriate for practitioners. Teachers can be directly involved in action research studies. Other educators might trust the results obtained from one of their colleagues doing action research than from unknown researchers.

Why Educational Research

We all know that education problems occur within or outside the school system. In an education system, problems occur at various levels and types of education such as pre-primary (or early childhood), primary, secondary or tertiary levels, teacher education, vocational and technical education and so on. What does this mean? It means that in educational research, researchers must use extensively scientific methods to find out how quality of education (teaching–learning process) can be improved, how endogenous and exogenous conditions responsible for knowledge creation can be tested and verified and the conditions under which they might occur can be improved and made conducive to learning. It means that educators need to be both consumers and producers. Creswell (2002) suggests reasons for which educational research is carried out:

Improved practice: Research can suggest ways of improving practice that have been verified with many applications and by many different types of people, which is difficult for practitioners.

Add to knowledge: Research can add to what we know about how people learn and what we can do help facilitate the learning process.

Address gaps in knowledge: Research can address areas in which little is known.

Expand knowledge: Research can allow us to extend what we know in ways we never conceived.
Replicate knowledge: Research can act as a test to verify previous findings.
Add voices of individuals to knowledge: Research can add an important perspective for different learning types. Much of the educational research prior to the 1980s is based on able, white, middle-to-upper class males. This is certainly not reflective of our increasingly heterogeneous students, and research helps revise theory and practice to reflect different student needs.

Objectives of Educational Research

The terms 'goal' and 'objectives' are often confusing. The difference between a goal and objective can easily be understood from the taxonomy given for the two terms in Table 1.3 and Figure 1.1.

Education is mainly concerned with the processes, which deal with deliberate change in the behaviour of people through the acquisition of knowledge, skills, attitudes, interests and appreciation. Therefore, the key goal of educational research should be to identify basic principles on which the researcher can

TABLE 1.3 Goals and Objectives: The Difference

Goals are	*Objectives are*
Broad statements	Specific
General intentions	Precise
Intangible	Tangible
Abstract	Concrete
Generally hard to measure	Measurable

Source: Author.

FIGURE 1.1 Goals and Objectives: The Difference

Source: Author.

base his or her interpretations, predictions, explanations and control of behaviour. It implies that any research in the domain of education should contribute to some aspect related directly or indirectly to the teaching–learning process prevalent at a given point in time in the education system. The topic(s) selected for investigation should be truly educational, preferably a problematic topic whose findings will add to the knowledge bank of education and lead to some solution of a societal problem.

Fraenkel and Wallen (2006) state that well-conducted research is the key to the success of global education endeavours such as Education for All (EFA), Education for Sustainable Development (ESD) and Millennium Development Goals (MDGs). Not only does it form the foundation of programme development and policies globally but it also translates into effective global education programmes.

Educational research will only be useful and effective in planning and policy formulation if it is empirical. Mere theorising about what might be effective or what could work will not solve the persistent problems our education systems are facing currently. Researchers' focal point should be grass roots people (local communities). They should work closely with local community and identify problems their children or schools are facing. It is only after this interaction, attempts should be made in designing research studies so that policymakers get hard data on which they can base their decisions. Using standard research methods and tools for producing results that can be examined and assessed by peers, methodologies that can be replicated and knowledge that can be applied to real-world situations, our research studies should be scientific and systematic. Researchers must not work in isolation but as a team to enhance our knowledge of how to best address the world's educational problems.

Educational research seeks to develop new knowledge and skills about the teaching–learning process to improve and transform educational practice. The main concerns are:

- Learning: How to assess students' learning and performance? Are students learning what they are supposed to be learning?
- Teaching: Are our teaching methods effective? What we can do to improve student learning?
- Motivation: What are the best practices teachers can use to motivate students to enable them actually understand: here, now, today, in this room? How can teachers motivate students achieve creativity, problem-solving abilities, a passion for learning, a dedicated work ethic and lifelong learning opportunities?
- Development: How do children and adults change over time?
- Classroom-based management: What classroom or school practices make the classroom optimal for student learning?

The aim of educational research should be to establish doable, affordable and sustainable practices that teachers, managers, administrators and students should use for improving the students' learning achievement. And for this to realise, we will have to collect data scientifically about these important topics in education. Therefore, as an educational researcher, we should focus on practical research projects that will have significant implications for our education systems. A strong foundation for the research is the querying attitude. It is, thus, important to recognise that everything and anything is questionable and these questions make way for a research.

Educational researches have a wide spectrum of objectives, right from doing evaluations, making interventions, comparing and developing learning material, streamlining policies, planning managing and evaluating education systems, finding cause and effect relationships to finding correlations among variables, solving problems and many more such objectives. It offers guidelines and direction to policy formulation basically in five key areas: explanation, prediction, monitoring, discovery (new theories) and hypothesis testing.

These are only a few of the many reasons research is important, particularly to educators. In an increasingly data-driven society, it is vital that educators know how to locate, find and interpret research on their own. Further, educators need to be able to conduct quality research to examine issues within their own contexts. Given in the following are some of the objectives of educational research:

Objectives of Education Research

- Identify education stakeholders' (student, parental, community and teacher) behaviour, opinions, attitudes, trends, patterns and perception, and explain why something is happening about education and significant impediments and facilitators of student learning achievements and outcomes.
- Identify a given education situation and predict what may happen in the future.
- Monitor many decisions made by policymakers and managers of the education system to insure that goals are being attained.
- Identify the relevancy of education and its relation to society and economy (income and employment).
- Describe the role and influence of both internal and external factors in shaping the education system as well as students' academic interests and pursuits.
- Explain which education offered is quality education, relevant and accessible to all.
- Identify strategies to improve students' participation and achievement in the education system.
- Understand society and social processes, as well as to test and or create theories in order that we are better able to inform about social action and potentially 'improve' social conditions.
- Test hypothesis about some educational issues.

Source: Postlethwaite (2005).

Functions of Educational Research

When it comes to the applications of research, the educational research can be classified into basic and applied research. The novice researchers often do not know or are confused about the possibility of the application of their research and what could be the scope of their research. Most of the research that is conducted by university students is applied research while researches conducted by scientists and experienced researchers may be basic or applied research.

This section looks at the difference between basic and applied educational research and attempts to discover why there is a separation.

> Basic Research: Basic research (also known as pure or fundamental research) is mostly very comprehensive and is conducted to obtain a deep insight into a field of study, a phenomenon or a theory. It is driven by a researcher's curiosity or interest in a specific question. Its main motivation is to 'expand our knowledge', not to invent or create something. There is no apparent commercial value to the discoveries that result from basic research. Therefore, the

purpose of basic research is not to impact the society at present but in the future. It is universities which largely benefit from basic research as it enhances the knowledge base to do further studies. Some basic researches are conducted to find out new techniques, procedures and tools to conduct the research itself. In short, basic research lays down the 'foundation' for applied sciences, which follow from basic research.

For example, basic research probes for answers to questions such as:

- How did the universe begin?
- What are protons, neutrons and electrons composed of?

The basic difference between basic research and applied research is what the two approaches will be used for. Will the research be deployed to enable us understand better a real-world problem or will it be used simply for further our general information?

Governments conduct most basic research studies and big organisations invest in it. Basic research is quite costly and time-consuming and as a result it often brings new ideas, innovations and develops theories that can bring forward entirely new possibilities.

Applied research intends to link research with action in a form that generates applied knowledge. It attempts to solve 'practical problems' of the modern world rather than to acquire knowledge for knowledge's sake. It is designed to 'improve the human condition'. This intention is evidenced in the processes of designing, carrying out and validating the research findings. It does not necessarily imply a preference for a particular type of knowledge or the methodology associated with its production. It is a type of research that is used to answer a specific question that has direct application to the world. This is the type of research that solves social and real-life problems.

The aim of applied research is to develop new products, ideas and goods to help the community at large. This type of research is most suitable to solve a problem or study a situation, phenomenon or idea. The scope of applied research is for the present and not for sometime in the future.

Applied research can be conducted at any level, such as by government organisations, agencies, institutions and even on a personal level too. There are various types of applied research: explanatory, exploratory, descriptive or confirmatory. In colleges and universities, researchers conduct applied research to examine and search the facts that are already present, but hidden.

For example, applied researchers may investigate ways to:

- Improve teaching–learning environment in schools
- Solve a specific problem facing the education system

Applied research is also classified as problem-oriented and problem-solving research. **Problem-oriented research** in an educational setting is done generally by the Ministry of Education (MOE) to sort out problems faced by the education system. As the name suggests, problem-oriented researches are undertaken to know the exact nature of the problem that is required to be solved. For example, why did the number of out-of-school children increased by 5 per cent or why did 10 per cent primary school teachers left schools last year.

A given educational institution for the problem faced by it organises and undertakes **problem-solving research**. For example, in order to identify the factors leading to a continuous decrease in girls' attendance in school, the principal may conduct the research to explore ways and means to solve this problem. In short, the main aim of problem-solving research is to discover some solution for some practical pressing problem(s).

Applied research is generally done on large-scale basis and as such it is quite expensive. It is often conducted with the support of some multilateral and bilateral financing agency like the World Bank, UNESCO, UNICEF, USAID and so on or by the government or public corporation. In an educational setting, applied research is carried out primarily for improving the efficiency, quality, relevance, effectiveness and communication.

Evaluation Research

Evaluation research is defined as a type of study that uses standard social research methods for evaluative purposes, as a specific research methodology and as an assessment process that employs special techniques unique to the evaluation of social programmes. It is used to determine the impact of a social intervention (an action taken within a social context) designed to produce an intended result. Evaluation research thus analyses the impact of a particular programme on a certain social problem the programme is trying to solve.

Appropriate topics for evaluation research are diverse and extensive as any other social research. In the field, there are three types of studies:

Needs assessment studies: they refer to particular studies directed to determine the existence and extent of problems, usually pertaining to a specific population.

Cost-benefit studies: they include studies that decide whether the results of a programme justify the expense. The cost could be financial or non-financial.

Monitoring studies: these studies provide a steady flow of information about a topic of interest. These are usually conducted over an extended period of time. In some cases, monitoring studies require incremental interventions, meaning the results may change slightly as monitoring methods alter and changes within the topic being studied are made.

Evaluation research makes use of several types of research designs. Important among them are:

Experimental design: This research design depends on random assignment and laboratory controls to ensure the most valid, reliable results. The designs produce the strongest, most valid and reliable results. However, experimental design is often not practical for many studies in social science and education because researchers cannot, in many instances, exercise laboratory controls in natural settings or randomly assign subjects.

Quasi-experimental design: Another evaluation research design type is quasi-experimental design. These research designs are non-rigorous inquiries that somewhat resemble controlled experiments; however, they lack key elements that are indicative of experiments, including pre- and post-testing and/or control groups.

Time series design: Time series design is an additional evaluation research design type. It measures a problem or an issue over a fixed time period such as the study of school attendance rates before and after the introduction of school feeding programme.

Evaluation research should be a rigorous, systematic process that involves collecting data about organisations, processes, programmes, services and/or resources. Evaluation research should enhance knowledge and decision-making and lead to practical applications.

In view of the emergence of several global initiatives and the problems resulting from global overpopulation, environmental degradation and the overuse of the earth's natural resources, experts and scholars feel that the time has come for a paradigm shift in emphasis away from purely basic research and towards applied research.

Whatever may be the choice of selection of a research method, for educational research to be meaningful, it is important that the topic of our research should focus on individuals in an educational context. The main purpose of our research should be to examine and analyse the key variables about a human being: how they behave, act, think, feel, believe and so on. The research, as such, should examine individual characteristics, not group structures. In an educational setting, the purpose of the research should be geared towards examining student performance, motivation, attitudes, beliefs, teachers' knowledge and so on. School-level variables such as school facilities, government funding and so on should not be the subject matter of educational psychology study. 'The key purpose should be to improve educational practice'.

Therefore, it is necessary that the researchers design and develop models that help in identifying key issues and challenges facing our education system on the one hand and suggest policy measures that can have a significant impact on improving the teaching–learning process.

As these research approaches are widely used in educational research, they have been discussed in details in subsequent chapters.

Educational Planning and Educational Research

According to UNESCO-IIEP (Postlethwaite 2005), there are primarily three types of educational research questions directly related to educational planning. These questions are descriptive, correlational and causal.

Descriptive Questions

In an educational setting, research on descriptive questions compares basically existing conditions of schooling with respect to: (a) legal standards and guidelines, (b) prevailing conditions in several other school systems or (c) conditions operating in several sectors of a single school system.

Correlational Questions

These type of questions assume an association between variables. If the variables are associated with each other, then this is an evidence of causation. However, the researchers should be very careful when they intend to understand the difference between the notions of association. For instance, the researcher can find an association between the incidence of classroom libraries and average class reading scores. But

> [I]n reality the higher score of students may be the cause that the students come from the high socio-economic strata, while they tend to be in classes with classroom libraries, read better than other students because their home environments (in terms of physical, emotional, and intellectual resources) facilitate the acquisition of reading skills. (Postlethwaite 2005)

Some other examples of correlational questions suggested by UNESCO-IIEP are as follows:

- Is it true that students in classes with high pupil–teacher ratios (PTRs) have lower achievement scores than those in low PTRs?
- Is it true that teachers are less motivated to teach in rural schools than those in urban schools?
- Is it true that students commuting longer distances to reach school have low attendance rates than those living within the vicinity of schools?
- Is it true that students in schools having trained/qualified teachers have higher test scores than students in schools where teachers are not trained/qualified?
- Is it true that boy students perform better in science subjects than girl students?

Causal Questions

Educational planners widely use causal questions. For example, are there schools with well-equipped classrooms and desk for children to sit? Although there are schools lacking both: equipped classrooms and sitting desks. It is important for educational planners to know whether schools (with a particular socio-economic context of children) with a shortage of chairs and desks score less than schools (with a similar socio-economic context of children) with an adequate supply of desks and chairs. In other words, is it only the desks and chairs, or something else, which is the real cause for the students' better achievement? The other factors that may cause better achievement of students may be a better availability of textbooks/exercise books, allocation of financial resources or better qualified/trained teachers. The important point that the planner has to bear in mind is, therefore, to extract the relative significance (impact) of each possible input and process factors on achievement.[1]

Educational researchers can use both surveys and experimental research methods for assessing the relative influence of many input and process factors on educational achievement. We all know that students' educational achievement is not influenced exclusively by one single factor. It is rather the case that several factors operating within and outside the school impact how well or poorly students perform in school. According to UNESCO-IIEP,[2] causal questions take one of two forms.

- Other things remaining the same, do students with Textbook A achieve better than students with Textbook B?
- What is the relative influence (effect) on school achievement of the following factors:
  - The socio-economic level of students in the school
  - The general parental help given to children with their homework
  - Peer group pressure
  - The condition of school buildings
  - The supplies and equipment in the classroom
  - The curriculum
  - The quality of teaching and so on

Knowing the fact that students' achievement is affected by many factors, the main task of the researcher is to discern those key factors that significantly influence student achievement. Once the planners

[1] See http://unesdoc.unesco.org/images/0018/001824/18 (accessed on 17 May 2017).
[2] Ibid.

successfully identify the key factors, then they can decide on the action they wish to take (Postlethwaite 2005). Let us assume that sufficient evidence suggests that the availability of adequate number of desks and chairs does have a major influence on achievement. Then it is up to the MOE to ensure that sufficient desks and chairs are made available. But the task of the Ministry does not end here. Then the Ministry will have to organise further research to find out what is the next most important factor impacting student achievement.

Characteristics of Research

For a research design to be good, it should (a) minimise bias and (b) ensure utmost reliability of the data collected and analysed. The experimental error should almost be minimal. At the same time, a good research design should be capable of generating and providing maximum information and an opportunity for understanding completely the different aspects of the problem analysed in the research study.

As a matter of fact, there are no hard and fast rules because the question of a good research design directly relates with the purpose and objective of the research problem under investigation. Nevertheless, to be considered and qualify as research, the process must contain certain characteristics. According to Leedy (1996), there are eight characteristics of research. A research inquiry should:

1. Originate with a question or a problem
2. Have a clear articulation of a goal
3. Follow a specific plan of procedure
4. Usually divide the principal problem into more manageable sub-problems
5. Be guided by the specific research problem, question or hypothesis
6. Accept certain critical assumptions
7. Include assumptions about the underlying theories or ideas about how the world works
8. Collect and interpret data in attempting to resolve the problem that initiated the research

It follows from this that research, as far as possible, must be:

- Empirical: Research must be based on observations and experimentation on theories by the researcher. It means that conclusions drawn from research investigation should be based on hard facts and evidences.
- Objective: The research design should allow the measuring instruments to be fairly objective. This must ensure the objectivity of the collected data. If required, the data must be capable for further analysis and drawing generalisations.
- Logical: Research should be based on valid procedures and principles. It must be objective in order to be used for anything.
- Cyclical: Research should start with a problem and should end with a problem.
- Analytical: Research should utilise proven analytical procedures in gathering data disregarding whether the research is historical, descriptive, experimental or a case study.
- Controlled: In real-life situation, there are a multitude of factors that have direct bearing on an outcome. Control in research means exploring causality in relation to two variables (factors) and minimising the effects of other factors affecting the relationship.

- Control could not be easily realised in social sciences as compared to physical sciences. In pure sciences, research is carried out in laboratories where the researcher can exercise controls relatively easily. But this is not possible in education and social sciences because educational research studies human behaviour, attitudes and perceptions all of which a researcher cannot control. Since the researcher cannot control external factors, the impact of these factors is difficult to quantify.
- Methodical: Research should be conducted in a methodical manner without any bias using systematic method and procedure.
- Replicable: Research design and procedures are repeated to enable the researcher to arrive at valid and conclusive results.
- Rigorous: For drawing inferences and finding appropriate solutions to research questions under consideration, research design must rely on standard, relevant, appropriate and justified procedures. However, the degree of rigour will vary significantly between the physical and social sciences and within the social sciences.
- Systematic: Research is a systematic activity. It should follow and adhere to certain logical sequencing in procedures. The several stages of research should not be organised in haphazard and piecemeal fashion. Some procedures must follow others.
- Critical: Research procedures and methods should be critically identified, scrutinised and sorted out well before conducting the research. The process of investigation must be free from all kinds of drawbacks. The methods and procedures must be able to withstand critical scrutiny. The research should exhibit careful and precise judgement.
- Reliable: It refers to consistency throughout a series of measurements. For example, a respondent should give the same answer for a particular research question even if he or she is asked the same question repeatedly by several investigators. If he or she is changing his or her response to the same item, it means the information included in the questionnaire is neither consistent nor reliable.
- Valid: Any measuring device or instrument is said to be valid only when it measures precisely what it is expected to measure. For instance, an intelligence test conducted for measuring the IQ should measure only the intelligence and nothing else. And to ensure the validity, the questionnaire shall be framed accordingly.
- Generalisability: It means how best the data collected from the samples can be utilised for drawing certain generalisations applicable to a large group from which sample is drawn. Thus, a good research design is one which enables and helps an investigator to generalise his or her findings provided he has taken due care in defining the population, selecting the sample, deriving appropriate statistical analysis and so on while preparing the research design. Thus, for research design to be a good research design, it should be prepared methodologically and should ensure that:
  - The measuring instrument can yield objective, reliable and valid data
  - The population is clearly defined
  - An appropriate sampling technique is used to draw a representative sample
  - A scientific statistical analysis has been carried out
  - The findings of the study are capable of generalisations

For a process to be considered research, it is imperative that it should have the earlier-mentioned characteristics.

Qualities of a Good Researcher

Can you be a good researcher? Do you possess right properties to be a good researcher? If not, then you are required to go for intensive training to acquire and build yourself up on those qualities that you find yourself wanting.

For you to be a good researcher, the first essential quality that you must possess is your 'strong intention' to be grossly involved in research. Thereafter you have to demonstrate a highly 'dedicated interest' to do the best research possible. From there, you must 'accumulate the knowledge' needed to advance the current ideas already existent in the research world. Thus, these are key and innate talent, skills and qualities that you as a researcher must possess or acquire immediately to succeed in this challenging task that requires a lot of imagination and perseverance.

What then are the qualities of a good researcher? The notable attributes of good researcher who tread the path towards discovery include both general qualities and specific qualities. They are described in the following.

General Qualities

Scientific Mind and Thirst for New Information

A good researcher looks and manifests thirst for new information. His or her approach is critical, scientific and analytical. He or she has a scientific open mind and is not influenced or guided by superficial facts and considerations and gives up personal likes and dislikes. He or she performs research along scientific lines. He or she does not just take things for granted or by themselves but explores new grounds. He or she thinks differently, unconventionally or from a new perspective, leaving out the conventional for something innovative. A good researcher treads the unknown frontier. He or she discards superstitions and avoids all types of bias. He or she manifests practical experience and emotional maturity and utilises fully his or her efforts to gain scientific knowledge and discoveries.

A good researcher is someone who is intelligent and fully competent to express his or her ideas. How can you express your thoughts if you do not know how to write and express your viewpoint? How can people understand your point when it is not explicitly written and presented? Your ideas are very important; but of course, better if you present them in such a way that others understand well what you want to say. A good researcher is adept, skilled or proficient in the written language.

Tenacious Seeker of Truth and Knowledge of Things Around Him or Her

A good researcher has a keen sense of things around him or her. He or she is hard working, diligent, focused and devoted and a seeker of truth, that is, he or she sees something more out of a common occurrence around him or her. And he or she sees this quickly. He or she can see a minute worm or pollen inside a flower or the beautiful colour combinations of the rainbow or a wild plant or simply notice the small ant in his or her dish.

Alertness

A good researcher is always vigilant to appreciate even the minute changes in situations. He is willing and ready to work under all circumstances. He or she is highly precise and accurate in making observations, drawing conclusions, quick in perception and statement.

Clarity in Understanding and Explaining

A good researcher knows his or her academic subject exhaustively. He or she possesses good communication skills and is able to put across his or her ideas to others. He or she seizes and firmly holds ideas. He or she likes to reflect or consider about the things he encounters. He or she always looks for avenues for enhancing the knowledge that he or she gained, either formally in school or through his or her experience.

Trained and Educated

A good researcher is very well conversant in his or her area of research. He or she has adequate practical experience and training to understand and examine critically the problems and suggest solutions for resolving them. He or she is familiar with the latest techniques of research.

Patient and Educated

A research endeavour combines in itself success and failure. A good researcher does not feel defeated or frustrated if the expected result does not come forth at the expected time or he or she does not seek the cooperation of respondents. For a research project to be completed successfully, many years may be required. A good researcher has the courage, conviction and social qualities such as friendliness with respondents, pleasant manners, teamwork and so on.

Specific Qualities

Knowledge of the Technique of Research

A good researcher also possesses innate knowledge of the technique that is to be applied to the problem. With this knowledge, he is able to prepare questionnaires, interviews, schedule and getting essential and reliable information.

Further, a good researcher has the quality in assessing systematically diverse situations. He or she applies a systematic and objective thinking to arrive at something. In other word, he or she relies strictly on logical reasoning.

Personal Taste in the Study

A good researcher takes keen personal interest in the study and not worried about approvals or disapprovals of his or her work. He or she does his or her work with zeal and patience. A forced work often leads to monotony, tiresomeness and eventual failure. He or she does not do the work in haste and come up with invalid and wrong statements.

Unbiased Attitude

The good researcher is unprejudiced, unbiased and free from all preconceptions. He or she maintains an open mind towards the subject under study.

Familiarity about the Informants

The researcher is familiar and friendly with the respondents of his or her study. If he or she is familiar with them, he or she can gather information easily. Further, he or she appreciates their views properly.

Based on the given qualities of a good researcher, we can safely summarise the key characteristics of a good researcher as follows:

- Intellectual curiosity
- Prudence
- Healthy criticism
- Intellectual honesty
- Intellectual creativity

Do you have these qualities and characteristics? If not, then it is time for you to harness the hidden talents in you through training and continuous learning.

Ethical Practices in Research

Ethics in research means a code of conduct or expected social norm of behaviour while conducting research. A researcher is required to follow strictly ethical practices at each and every step of the research process. Practising ethics in research is a complex task that requires much more than merely following a set of rigid guidelines both from professional associations and from university/institutional review boards (IRBs). Of all of the steps in the research process, ethical practices relate very closely with the processes of data collection, reporting of research outcomes and distribution of reports than any of the other phase of research.

Practically, organisations and research-sponsoring members, the researchers who undertake the research and the respondents who provide to researchers the key data will have to follow an ethical conduct. The observance and follow-up of ethics starts with the person/institution/organisation instituting the research, who should do so in good faith, pay attention to what the results indicate and surrendering the ego, pursue organisational rather than self-interests.

One of the ethical considerations that the researcher has to keep in mind while conducting his or her research is to respect the site in which the research takes place. The researcher should demonstrate this respect by gaining prior permission to enter the site, by disturbing the site as little as possible during his or her study, and by viewing himself or herself as a guest at the place of study. For this, universities and other specialised institutions have illustrative guidelines to be followed by the researcher in the process of conducting his or her research.

Another code of conduct is seeking permission from all academic and non-academic staff/officials for gaining access to your research site. For example, if your study site is a high school classroom, then you have to seek permission from several individuals. These individuals include the following: members of school board responsible for ensuring the protection of the participants' rights, the research officials in the school district, the principal of the school, the teacher in the class, the actual students who will participate in the study and their parents.

In short, ethical behaviour is not required for any single stage of the research process. It rather permeates through all stages of the research process—data collection, data analysis, reporting and dissemination of information—provided such an activity is undertaken. Treatment of the subjects and confidentiality of information both are steered by research ethics.

Limitation of Educational Research

Despite the value and importance of educational research discussed earlier, there is a need to evaluate its contributions. Sometimes the results and findings of research studies are contradictory and contain vague findings. As such, only a handful of research studies suggest a 'declarative sentence'—which can be the basis of legislation or policy formulation. Policymakers in ministries of education and readers of educational research always look for some evidence that reveals a direct statement about an educational issue.

Most studies in education are observational studies. This means that investigators are absorbed in the reading or study of data previously collected by others. They seek correlations between different variables. Compared to other methodologies, this approach is far less expensive because it is easier and faster. With research budgets stretched thin, cost is a major consideration. The problem is that observational studies are subject to biases that sometimes make the results unreliable. If others cannot replicate results, the conclusions lose credibility.

The gold standard in education is experimental studies. These studies often are primarily randomised controlled trials or randomised field trials. They involve random assignment of students to groups that differ only by the treatment. There are two groups—one group is experimental that receives a particular treatment; the other is control group that does not get any treatment. Investigators draw inferences on the basis of the data collected. The problem is that experimental studies are costly and hard to implement because not all parents or schools are willing to let their children participate.

No matter which study is used, however, one can immediately refer to a common error: correlation is not causation. Just because two things are statistically associated (correlated) does not necessarily mean that one was responsible for the other (causation).

Another hurdle is that studies in the social sciences are viewed as less 'scientific' than studies in the physical sciences. This may lead to a conflict of opinions among different scholar over a given topic. They may examine the same data and may reach opposite or conflicting conclusions. In situation such as this, researchers often use 'substantively significant' in arguing in favour of the programme, which can be more ambiguous term than 'statistically significant'.

It is the ethical responsibility of the researcher to protect the rights and welfare of the participating subjects which could involve physical and mental discomfort, harm and danger. Also, most studies require that consent be obtained from parents, institutions and so on. Laws should also be in place to protect the confidentiality of the data as well as the privacy of the subjects. Since external environment influence education, it may be necessary to make additions, deletions or modifications. One cannot know immediately the effects of these changes because these effects often occur years later. The public nature of education also influences the kinds of research questions investigated. Some topics may be too controversial for conservative communities or too divisive for a particular staff.

Another equally important problem is the issue of questionable data. Evidence suggests that quite often the author of a particular research study may not have gathered information from intended people as well as from people who are able to understand and address the questions contained in data collection instruments used for the study. The sample size is dismally low to draw significant inferences or test the validity and significance of the research hypothesis. Further, from technical point of view, the researcher may have chosen an inappropriate statistic for analysing the data. Just because research is published in a well-known journal does not automatically make it a 'good' research.

Further to these issues, there is also a problem related to the ambiguous and unclear statement about aims and objectives of the study, the lack of a complete explanation of data collection procedures or ambiguous and vague statements of the research problem all of which drives the research inquiry into are other serious problems.

In short

> Research has limits, and you need to know how to decipher research studies because researchers may not write them as clearly and accurately as you would like. We cannot erase all 'poor' research reported in the educational field. We can, however, as responsible inquirers, seek to reconcile different findings and employ sound procedures to collect and analyse data and to provide clear direction for our own research. (Creswell 1994)

Summary

Educational planning is not a static process. The structure and organisation of the education system, curriculum, pedagogy and textbooks, teaching modes and methods, teacher training strategies, the school provisions such as science laboratories, textbooks, furniture, classroom supplies and so on are not static rather they show regular changes. These changes may initiate either an improvement or deterioration in the quality of education. There is also the possibility of having no impact on quality. If such is the case, one can say with fair amount of certainty that public outlays on education have been wasted.

> The educational planner [researcher] working within this kind of environment must be able to undertake assessments of the effects of major changes and then provide policy advice that will consolidate and extend the post productive courses of action, and also intercept and terminate existing practices that are shown to be damaging and wasteful. (Postlethwaite 2005)

No one can deny the significant role impartial research plays to improve education standards and the quality of teaching and learning in our education system. Research is an integral part of education studies and teacher development. But it is unfortunate that practitioners, administrators, managers and decision-makers not always respect the concept of research. According to Kit Field (2011),

> [T]oo often research in education is seen as an academic activity carried out by others—to the profession, not with the profession. Since research enables professionals learn new ideas and theories, explore things, analyse information, adapt their attitude and behaviour in accordance to the information obtained, look to improve and adapt to modern demands, all of this constitutes research. Educational research deserves respect.

It is through research that teachers understand what works and why in the classroom, the short- and long-term implications of an educational reform or of a particular policy measure, the justification for decisions and actions. It helps them to deal with the unexpected, identify problems and inform improvement and so forth. Research invents and innovates new ways of learning. Thus, research should be future oriented and designed to benefit learners rather than the researchers themselves. Educational research should improve instruction, students' achievement, teachers' competence, satisfy learning needs of all and train graduates to become responsive to economic development of the country and to compete globally. And this is the duty of academic professional everywhere.

The research methodology is a complex topic. This chapter described basic philosophy underlying educational research and some of the most basic issues commonly encountered by educational researchers. The chapter explained several research methods commonly used in educational research studies and some of the complexities involved in deciding upon an appropriate methodology. Ultimately, the methodology the researcher will choose will be determined by the specific research question and by the resources that are available to him or her.

Chapter 2 explains research design and data collection techniques. This will help you to decide what you are doing and why you are doing research.

Self-test Exercise

Exercise 1.1:

1.1.1. In educational research, which reasons would you use for choosing a paradigm. Why?

1.1.2. Explain how research is designed to gain new knowledge.

1.1.3. How does the research protocol contribute to informing MOE's education practices?

1.1.4. Who is ultimately responsible for the design and conduct of an educational research study?

1.1.5. How do you understand and interpret what the research in education says?

References

Babbie, E.R. 2010. *The Practice of Social Research*. Belmont, CA: Wadsworth.

Chinn, C. 2006. The Microgenetic Method: Current Work and Extensions to Classroom Research. In *Handbook of Complementary Methods in Education Research*, edited by J.L. Green. Delors, Jacques, et al. 1996. *Learning: The Treasure Within.* Report to UNESCO of the International Commission on Education for the Twenty-first Century.

Cmailli, G and P.B. Elmore, 439–456. Mahwah, NJ: Erlbaum.

Cunning, David. 2007. 'Descartes on the Dubitability of the Existence of Self'. *Philosophy and Phenomenological Research* 74 (1): 111–131.

Creswell, J.W. 1994. *Research Design: Qualitative and Quantitative Approaches*. Thousand Oaks, CA: SAGE Publications.

———. 2002. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Upper Saddle River, NJ: Merrill Prentice Hall.

———. 2003. *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. Thousand Oaks, CA: SAGE Publications.

———. 2008. *Educational Research: Planning, Conducting and Evaluating Quantitative and Qualitative Research*, 3rd ed. Upper Saddle River, NJ: Pearson Prentice Hall.

Dash, K.N. 2005. *Module: Selection of Research Paradigm and Methodology*. Indira Gandhi National Open University. Retrieved on 23 January 2018 from: https://www.celt.mmu.ac.uk/researchmethods/Modules/Selection_of_methodology/

Edgar, Faure. 1972. *Learning To Be: The World of Education Today and Tomorrow*. Paris: UNESCO.

Field, Kit. 2011. *The Importance of Research for Education's Future*. Available at http://www.expressandstar.com/education/2011/06/14/the-importance-of-research-for-educations-future/ (accessed on 15 November 2017).

Fraenkel, J.R. and N.E. Wallen. 2006. *How to Design and Evaluate Research in Education*, 6th ed. Boston, MA: McGraw Hill.

Guba, E.G. and Y.S. Lincoln. 1989. *Fourth Generation Evaluation*. Newbury Park, London and New Delhi: SAGE Publications.

Keeves, J. 1988. 'The Unity of Educational Research'. *Interchange* 19 (1): 14–30.

Kothari, C.R. 2004. *Research Methodology: Methods and Techniques*, revised edition. New Delhi: New Age International.

Kuhn, T.S. 1962. *The Structure of Scientific Revolution*. Chicago, IL: University of Chicago Press.

Leedy, P.D. 1996. *Practical Research: Planning and Design*. New York, NY: Macmillan Publishing Co.

Mackenzie, N. and S. Knipe. 2006. 'Research Dilemmas: Paradigms, Methods, and Methodology'. *Issues in Educational Research* 16 (2): 193–205.

McCarthy, Thomas A. 1979. *The Critical Theory of Jurgen Habermas*. Cambridge: M.I.T. Press.

Nwana, O.C. 2005. *Introduction to Educational Research*. Ibadan: Heinemann Educational Books.

Omatseye, J.N. 2004. *Educational Philosophy and the African School*. Benin City: Mindex Publishing.

Peter, J.R.M. 1979. *The Analysis of Knowledge*. Oxford: Lincoln College.

Postlethwaite, N. 2005. *Educational Research: Some Basic Concepts and Terminology Module 1*. Paris: UNESCO International Institute for Educational Planning.

Siniscalco, M.T. and N. Auriat. 2005. *Quantitative Research Methods in Educational Planning: Questionnaire Design*, Module 8. Paris: UNESCO International Institute for Educational Planning.

Taylor, P.C. 2010. *Transformative Educational Research for Culturally Inclusive Teaching*. Khabarovsk, Far East Russia, Keynote address delivered at the 7th International Conference on Intercultural Competence.

Taylor, P.C. and M.N.D. Medina. 2013, 10 June. 'Educational research Paradigms: From Positivism to Multiparadigmatic'. *Meaning-Centred Education* 1 (1). Retrieved on 23 January 2018 from: https://www.researchgate.net/profile/Peter_Taylor11/publication/264196558_Educational_research_paradigms_From_positivism_to_multiparadigmatic/links/53d1c13d0cf228d363e8eccd/Educational-research-paradigms-From-positivism-to-multiparadigmatic.pdf

White, A.R. 1982. *The Nature of Knowledge*. NJ: Rowman and Littlefield.

Willis, J.W. 2007. *Foundations of Qualitative Research: Interpretive and Critical Approaches*. Thousand Oaks, CA: SAGE Publications.

Woody, C. 2011, 23 July. *Research Concepts and Statement of Problem*. Available at https://qualityrg4.wordpress.com/2011/07/23/research-concepts-and-statement-of-problem/ (accessed on 15 November 2017).

2
CHAPTER

Qualitative Research Designs

Introduction

Chapter 1 aimed at providing an opportunity to novice researchers to establish or advance their understanding of research through critical exploration of research language, ethics and approaches. The chapter introduced paradigms of research, the language of research, ethical principles and challenges, and the elements of the research process within quantitative, qualitative and mixed methods approaches. These theoretical underpinnings are highly useful to begin to critically review literature relevant to your field or interests and determine how research findings are useful in forming your understanding of your work, social, local and global environment.

This chapter introduces first the concept of research designs and then the most frequently used designs in qualitative research studies. The chapter attempts to answer:

- Why should we care about research design?
- What makes a strong research proposal?
- How to prepare a good research design?
- How we can use a given research design in qualitative research studies?

A similar attempt has been made in Chapter 3 to describe the research designs the researchers use frequently in quantitative research studies.

Research Designs

A research design 'provides the glue that holds the research project together. A design is used to structure the research, to show how all of the major parts of the research project work together to try to address the central research questions' (Trochim 2005). To make it simple, the research design is like a recipe.

Just as a recipe offers your mother or wife a list of ingredients and the instructions for preparing an exotic dish, the research design provides the components and the plan for successfully carrying out your research study. The research design is the 'backbone' of the research protocol.

Researchers design their studies in particular ways for increasing the chances of collecting the information needed to answer a particular question. The information they collect during research can only be meaningful and useful if the research design is sound and follows the research protocol. A careful selection of your research design increases the chance that the results of your research will be accurate and meaningful to others. The selection of an appropriate research design is also necessary because your study will provide to other researchers chances to reproduce similar results. The more often your results are reproduced, the more likely it is that researchers and the public will accept these findings as true. Additionally, it is vital that your research design must make clear the procedures used to ensure the protection of research subjects, whether human or animal, and to maintain the integrity of the information collected in the study.

Thus, before beginning the research study, a researcher should decide how he plans to design his study. The research design ensures the researcher that the evidence he has collected enables him to effectively address the research problem logically and unambiguously as much as possible. In educational research, obtaining information relevant to the research problem generally entails specifying the type of evidence needed to test theory, to evaluate a programme or to precisely assess meaning related to an observable phenomenon.

With this in mind, a common mistake made by researchers is that they begin their researches/investigations far too early, well before giving a critical thought about what information is required to address the research problem. Without paying due attention to these design issues beforehand, the overall research problem will not be adequately and effectively addressed, and any conclusions drawn will run the risk of being weak and unconvincing. As a result, the overall validity of the study will be undermined and questioned.

The decision about the length and complexity of describing research designs in a research study can vary considerably from one research design to another, but any well-developed design should attempt to achieve the following. It should:

- Identify clearly the research problem and justify its selection, particularly in relation to any valid alternative designs that could have been used
- Review and synthesise previously published literature associated with the research problem
- Clearly and explicitly specify hypotheses (i.e., research questions) central to the problem
- Effectively describe the data, which will be necessary for an adequate testing of the hypotheses and explain how such data will be obtained
- Describe the methods of analysis to be applied to the data in determining whether or not the hypotheses are true or false

This tag of the chapter attempts to describe the key features of a research design and the several types of designs commonly used in qualitative research studies together with their strengths and weaknesses.

Meaning of Research Design

Researchers often use the terms 'research approach' and 'research design' interchangeably in research. But in reality, the two terms do not convey the same meaning. Research design is a broader concept, which refers to 'a plan to conduct a study' whereas the term research approach refers to 'a constituent part of research design which governs it'.

In simple terminology, a research design is the framework or road map (guide) that researchers utilise for planning, implementation and analysis of a research study. It is a systematic plan of 'what' is to be done, 'how' it will be done, 'how' the data will be collected and analysed, 'how' the instruments will be used and the 'intended means' for analysing the data and its presentation. A research design basically comprises the research approaches, dependent and independent variables, sampling design, a plan for data collection, analysis and presentation.

According to Ackoff (1953), 'research design is the process of making decisions before a situation arises in which the decision has to be carried out'. The research design is the conceptual framework within which a researcher carries out his or her investigation. It is an important aspect as it binds the research project together. Its aim is to provide methods and techniques for data collection and analysis of relevant information with minimal efforts, time and money.

In simple terms, a research design is made up of the researcher's decisions pertaining to what, where, when, how much and by what means to carry out a research study. It clearly shows the sequence of conditions for the collection and analysis of data in a manner which aspires to combine relevance to the research purpose. It is an overview of what the researcher will do from writing the hypothesis and its operational significance to the final analysis of data.

The research design provides decisions with respect to the following:

- What is the study about?
- What are the reasons for conducting the study?
- Where the study will be performed?
- Which kind of data are required?
- Where can we find the necessary data?
- What will be the duration (time period) of the study?
- What will be the sample design?
- What methods of data collection will be used?
- How will the data be analysed?
- In what style the report will be organised?

Burns and Grove (2003) divide the entire research design into the following sections:

- Sampling design that relates to the process of selecting items to be observed for the research
- Observational design that pertains to the conditions under which the observations are to be made
- Statistical design that concerns with the question of how many items are to be observed and the way the information and data collected should be analysed
- Operational design that is concerned with the techniques through which the procedures given in the sampling, statistical and observational designs can be executed

In a nutshell, a good research design must contain (a) a clear and precise statement of the research problem; (b) data collection processes and methods; (c) the target population to be examined and (d) techniques to be used in processing, examining and analysing data. The design is utilised to structure the research to indicate that all the major elements of the research have been designed to work together.

Why Should We Care about Research Design?

We have been taught in Chapter 1 that research is the study of materials, sources and data in order to get conclusions. Getting the research design right is the first step towards organised research, which is more likely to be good research.

If you are an undergraduate student required to prepare independently a dissertation on a significant topic, your dissertation will be a major part of your degree, and it is important that it is correctly put together. Your dissertation is important because it will have added value in terms of 'many learning outcomes'. The way that you design and plan your research will have significant implications for the success of your research study.

Research design is of paramount importance as it influences the reliability of the results attained. It, therefore, offers to the researchers a solid base for the whole research. It is needed because it allows for the smooth working of the many research operations.

For instance, for building of a house, we must have a suitable blueprint made by an expert architect. Similarly, we should have a research design or a plan prior to data collection and analysis for our research study. It means advance planning of the techniques to be implemented for collecting the appropriate data and the strategies to be employed in their analysis, keeping in mind the purpose of the research and the resources (staff, time and money) we have at our disposal. For all of this, we require a suitable research design or plan just before data collection and analysis of the research project. The research design should be planned carefully as even a small mistake might mess up the purpose of the entire research. The design will help you to organise your ideas, which in turn will help you to recognise and fix your faults, if any. Research design, actually, has a great significance and impact on the reliability of the results achieved and as such constitutes the firm base of the entire edifice of the research work.

A good research design must be comprehensive and should ensure that all the components have been put together with each other in a coherent way. The theoretical and conceptual framework must be in line with the research goals and purposes. In the same way, the data collection method must exclusively be in line with research purposes, conceptual and theoretical framework and method of data analysis. A research design is like a successful journey which:

- Broadens our mind
- Provides fascinating and exciting experience
- Gives understanding of world around us
- Provides chance to meet people and share their experiences
- Gives fun and reward, but sometimes is very tedious and monotonous too

The research design should be prepared meticulously because any error in it may upset completely the research project. Even then, the necessity for a well-planned design is at times not realised by many researchers. The researchers sometimes do not attach significance which this problem warrants. It is because of this that many research studies do not fulfil and serve the purpose for which they are undertaken. The fact is they often lead to misleading conclusions.

Thoughtlessness in developing the research study is likely to render the research exercise futile. It is, for that reason, crucial that you should prepare an efficient and appropriate design before beginning the research operations.

The importance of research design in research methodology is due to the following reasons. The research design:

- May lead to the preferred kind of study and to meaningful conclusion
- Controls inaccuracy
- Allows optimum efficiency and reliability
- Reduces wastage of time
- Minimises uncertainty and confusion related to any research problem
- Helps greatly in collecting research material and testing of hypothesis
- Gives research the right path
- Controls bias and margin of errors
- Provides tentative estimates of resources needed in terms of money, effort, time and manpower
- Sets boundaries and helps prevent blind search
- Maximises reliability of results
- Prevents misleading conclusions, and thoughtless and useless exercise
- Anticipate flaws and inadequacies (anticipates problems)

Briefly speaking, the research design offers you the structure for your research project. It links all of the elements of the research together and provides the researcher the opportunity to carefully consider the research and to plan the way in which they will approach the research.

Now the question is: How can we prepare a good research design? This has been explained in the following paragraphs.

How to Prepare a Research Design

As a starting point, it is very important that you consider whether your research is realisable. Since research projects vary in size from small to large projects, they can offer us varying amounts of information. 'Different problems are suitable for people to tackle at different stages of their career. Is your project of a suitable difficulty for you to tackle, and does to add enough to our body of knowledge?'(Alon 2009). A good research project is the one that is concise, clear and complete—not asking to do too much.

While preparing a research design, you should take into consideration all necessary precautions, as any error may upset the whole project and defeat its utility and purpose. The reliability of result, which you are looking, is proportional with research design that constitutes a firm foundation of entire body of your research work.

Davies (2013) lists 10 steps for you to consider when designing a research project (Figure 2.1). With these steps, you can prepare a good research design. It is important to remember that the steps of research design process work together and frequently occur simultaneously.

Step 1: Reason for doing this research

In the first step of your research design, you should decide what are you doing and why are you doing this research study. Often researchers are so immersed in their work that they think that everybody knows why they are doing this research. But this is not the case. You should be able to explain clearly the

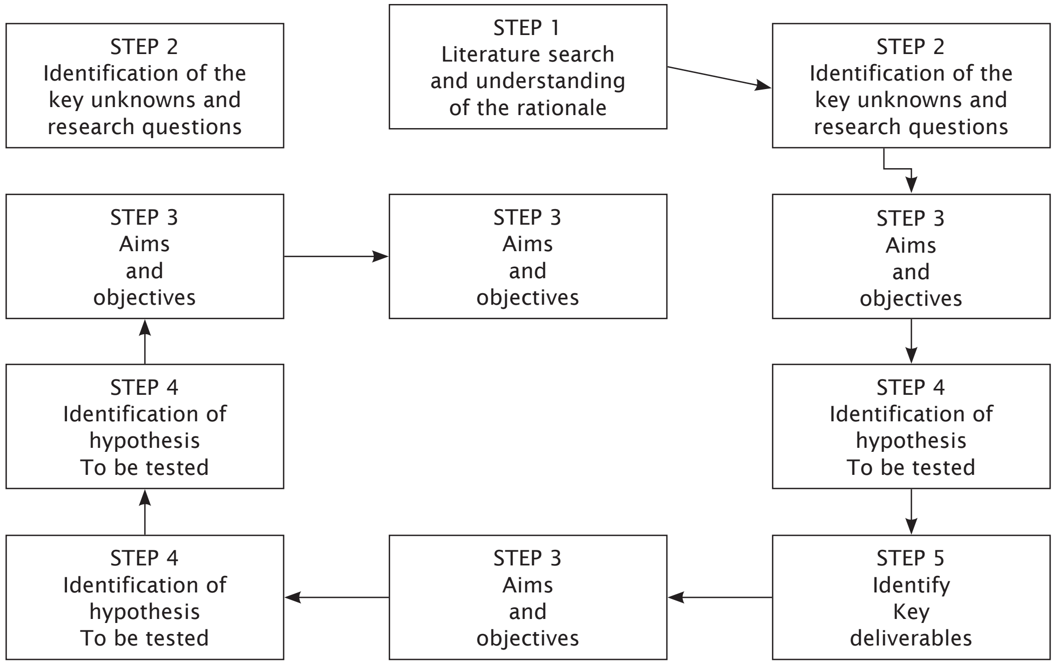


FIGURE 2.1 Basic Steps in Research Design Project

Source: Adapted from Davies (2013).

importance of your work. Here the point that deserves your attention is that you have to begin your research design with a 'rationale'.

Step 2: Identify the important unknowns

Once you have decided Step 1, your next task would be to search and read the earlier work close to your topic of research and identify the facts which you do not know, in particular the key 'research questions'. Then identify and assess those 'gaps', which you think your research can fill. This step is critical, as it will assure you that your research has not already been done.

Step 3: Aims and objectives of your study

These will largely depend on the rationale of your research study. However, it is imperative to decide an 'aim', which in turn should address the key unknowns identified in Step 2 above. The best way for you is to express in one sentence the aim of your study.

Example

> **Aim**: to map patterns of attainment in secondary mathematics and enrolment, retention and attainment in advanced level mathematics, given various measures of students' socio-economic status.

As far as 'objectives' of your study are concerned, they should be written in such ways that they help you to achieve the aim of your research study. The ideal would be to identify not more than five objectives each of which brings you a step closer to achieving your aim. You should make sure that each objective should be linked with research questions so that you are looking to achieve something new and original. For example, for the above aim of the study, the meaningful objectives could be the following:

Example

The following are the specific objectives of this study:

- To access and analyse existing school-level data and department-level data to map patterns within the existing models across the economically backward Western provinces of China
- To generate visual maps of attainment and advanced-level participation using geo-demographic approaches and Geographic Information Systems (GIS) software
- To conduct questionnaire surveys of learners of mathematics in Grade 9 and 10 in four clusters of schools in order to investigate, amongst other things, attitudes to learning mathematics and intentions for post-compulsory study
- To interview local authority subject advisors, heads of mathematics departments and survey mathematics teachers to ascertain how departmental cultures and resourcing relates to learner identities, attainment and advanced level enrolment.

For your aims and objectives to be meaningful, you should be certain that they:

- are specific, achievable and feasible
- are clear in terms of deliverables
- specific, clear and overarching research questions
- realistic about methods and timescale available
- use words such as compare, determine, characterise, explain, quantify, interpret and measure

Step 4: Nature of hypothesis you are testing

Your research study is certainly going to test some 'hypotheses'. You should identify these hypotheses in accordance with your analysis of previous work and key unknowns discussed in Steps 1 and 2 above. You have to formulate one or two hypotheses that you will test. As a good researcher, you should attempt to falsify your hypotheses. They should be grounded on the literature, your identification of the key knowns and unknowns, and should move your enquiry further.

Step 5: Identify the key deliverables

The basic question here that you have to respond is: What are the key 'outcomes' and 'deliverables' of your research going to be? You should use for the deliverables words such as understanding, quantification, conceptual, process, analysis, characterisation and determination.

Example

- An improved understanding of teaching–learning process
- Process-based conceptual models of teachers training
- Quantification of students' learning achievement
- Analysis of trends over time in girls' absenteeism in school

These deliverables will enable you to test your hypotheses and achieve your aim of the study.

Step 6: Identification of resources

For the completion of your research, you should identify and prepare a checklist of all necessary 'resources'. Will your study involve field visits, if so, for how long? Will you need any specific computer resources, assistance from statistician for the analysis and of data?

Step 7: Time frame

How much time will you take to complete each key objective of your study? Your research design should contain somewhat a precise and realistic assessment of the 'time' you will devote to the realisation of each

TABLE 2.1 Time Frame for the Research Study

Objective	*Month 1*	*Month 2*	*Month 3*
Objective 1			
Objective 2			
Objective 1			
Objective 2			
Analysis			
Key Outcomes			

Source: Davies (2013).

objective of your study. This you can do by preparing a Gantt chart that highlights the corresponding time for each objective and the amount of time you have at your disposal. An example of the Gantt chart is shown in Table 2.1.

Step 8: Prepare a workflow model
Upon the completion of all the above steps, your next task is to prepare a coherent 'workflow model'. At the top of the workflow model, write the aim of your research and then the hypotheses beneath it. Below the hypothesis, show your resources or inputs. Then, in each separate box, write each objective and the key deliverables associated with this objective. Finally, at the bottom, present your final result, for example, hypotheses accepted or refuted. This makes it very clear how your research project will fit together, what you will realise and how it relates with your aims and hypotheses.

Step 9: Risks and risk prevention
If your research study involves fieldwork, you should assess 'risks' and clearly identify 'hazards' and how you will control or prevent them. But you should also be aware of more general risks such as:

- Equipment failure
- Not being able to obtain key data sets
- Access to field sites

Step 10: Undertake your research
You are now ready to begin. You are going to investigate something that is important, interesting and something that you will enjoy. You have prepared a good research plan and know that your work is relevant to society and other researchers.

Types of Research Designs

Once you have prepared your research design, your next task would be to select a suitable research design. There are many ways in which you can conduct your research. Researchers doing research in their particular field of study, psychology, sociology, social work and education use each of these designs.

TABLE 2.2 Classification of Research Designs

Qualitative Research Designs	*Quantitative Research Designs*	*Mixed Methods Research Designs*
Action Research	Cross-sectional	Observational design
Case studies	Experimental design	Interviews
Ethnographic studies	Longitudinal design	Traditional surveys
Grounded Theory	Quasi-experimental design	
Historical studies	Surveys	
Phenomenological studies		

Source: Davies (2013).

The most prominent research designs are shown in Table 2.2. Each of these research designs has been discussed in detail in this chapter and Chapters 3 and 4. The data collection techniques (methods) for each of these major research designs are described in Chapters 5 and 6.

Qualitative Research Designs

With qualitative research designs, a researcher garners a complete and good judgement of human behaviour and the underlying reasons that govern such behaviour. Qualitative research designs are by definition exploratory. Often, they are used when the researchers are not able to ascertain their expectations, define the research problem or develop an approach to the problem. These research designs are also used to analyse in-depth the issues of interest and explore differences related to the problem at hand.

The qualitative research process reflects the methodological approach which a team of researchers decides to adopt. Several research designs are described in terms of the type of analysis they imply, as can be seen from Table 2.2. Different research designs also involve different sets of assumptions about what sorts of information (or knowledge) are important. Each of these qualitative research designs is described with examples and is followed by more information so as to enable you understand their basic features. At this stage, the most useful message that you have to keep in mind is that there are many different types of qualitative research designs.

Action Research Design

Action research in the field of education is often used as an interactive method of collecting information, that is, it is used to explore topics of teaching, curriculum development and student behaviour in the classroom.

In simple terms, action research is 'learning by doing'. In this research design, a group of people identifies a problem, takes some action to resolve it, sees how successful it efforts were and, if not satisfied, try again. While this is basically the essence of action research, there are other important attributes of action research that differentiate it from common problem-solving activities that we all engage in everyday life. Two succinct definitions of Gilmore, Krantz and Ramirez (1986) and Kemmis and Mctaggart (2014) define action research as follows.

> 'Action research ... aims to contribute both to the practical concerns of people in an immediate problematic situation and to further the goals of social science simultaneously. Thus, there is a dual commitment in action research to study a system and concurrently to collaborate with members of the system in changing it in what is together regarded as a desirable direction. Accomplishing this twin goal requires the active collaboration of researcher and client, and thus it stresses the importance of co-learning as a primary aspect of the research process' (Gilmore, Krantz and Ramirez 1986).

> 'Action research is a form of collective self-reflective enquiry undertaken by participants in social situations in order to improve the rationality and justice of their own social or educational practices, as well as their understanding of these practices and the situations in which these practices are carried out. The approach is only action research when it is collaborative' (Kemmis and Mctaggart 2014).

Action research aims at finding a solution for an immediate problem. Action research is something that we can use. In other words, we research the best way to study and apply our findings—this is action research. It researches to identify social, economic or political trends that may affect a particular institution.

The main objectives of undertaking action research in education are to:

- Professionalise teaching
- Enhance the motivation and efficacy of a worn-out faculty
- Meet the needs of an increasingly diverse student body
- Achieve success with 'standard-based' reforms

Types of Action Research Designs

Differing in purpose, emphasis and results, three types of action research permit educational researchers to investigate problem areas and overcome the challenges within their classrooms and schools. These action research designs are (a) individual action research, (b) collaborative action research and (c) school-wide action research.

Individual action research: In this research, an individual works independently on a project, such as an elementary school teacher conducting his/her own, in-class research project with his/her students.

The main purpose of this type of research is its focus on changes in a single classroom. The process for conducting the research is simple. A problem of interest in classroom management, institutional strategies or materials, or students' cognitive or social behaviour is identified and defined by a teacher. He then looks for solution(s) for the problem. This research is basically inspired by university courses.

Some teachers use quantitative data, develop measures, and formulate and test hypotheses. They experiment with different actions fashioned to address the problem, study and record the effects of those actions and keep, modify or discard ways of acting based on their findings. Some teachers also use qualitative data in similar processes. A few teachers, operating more like phenomenologists, prefer to let the hypotheses emerge from the process (Carr and Kemmis 1983).

Collaborative action research: This design involves a group of teachers or researchers working together to explore a problem that might be present beyond a single classroom, perhaps at the departmental level or an entire grade level.

Depending on the numbers of teachers involved, collaborative action research can focus on problems and changes in a single classroom or on a problem occurring in several classrooms. The data utilised by collaborative action researchers may be qualitative or quantitative.

School-wide action research: In this design, teams of staff members conduct the research jointly to focus on issues present throughout an entire school or across the district.

In school-wide action research, the teachers of the school select an area or problem of collective interest, then collect, organise and interpret on-site data. Data from other schools, districts or the professional literature are fed into the collective decision-making process of the faculty, who then determines the actions to be taken. The process is cyclic and can serve as a formative evaluation of the effects of the actions taken.

The following example clarifies the three types of action research designs.

Example

Four middle-school teachers—Mohan and Vincent from School A and Tania and Chao from School B—experiment with mnemonic key words in their science classes. The purpose is to help the students remember and understand key science concepts and terms. They consult and discuss regularly with an expert of the district education department and a professor from the nearby university, both of whom are experimenting with the same method.

The Secondary School X staff wants to increase student achievement. To realise this goal, all faculty members add a new instructional intervention (e.g., inductive thinking strategies). All of them observe and note down student responses to the change in instruction and discuss their findings. A leadership team organises each fortnight a meeting to seek technical assistance with the Provincial Council of Educational Research of the MOE.

The above scenarios all explain and describe action research. The first, conducted by a single teacher is 'individual action research', the second, carried out by a volunteer group with a university professor and staff development officer, is 'collaborative action research'. The third, involving the entire faculty in conjunction with a provincial council of educational research, is school-wide action research. It is, thus, clear that teachers and administrators do an action research.

You can conduct action research by using many methods. The important ones include:

- Observation (individuals or groups)
- Audio and video tape recording
- Structured or semi-structured interviews
- Field notes
- Photography
- Surveys or questionnaires

You can also use one method or more than one of the methods above for collecting rich and meaningful data for your action research. These techniques have been further discussed in detail in the chapter dealing with the methods of data collection in educational research. However, your selection of any method will largely depend on your research question(s) and hypothesis.

The action research design follows a typical cycle whereby initially an exploratory stance is adopted. It means it starts with an understanding of a problem and accordingly prepares plans for some form of strategic interventions. Then the researcher carries out the interventions (called 'action' in action research) during which he or she collects pertinent observations in various forms. He or she carries out the new interventions, and he or she continues repeating this cyclic process until a sufficient understanding of (or a valid implementation solution for) the problem is realised. The procedure is iterative or cyclical in nature. It is intended to foster deeper understanding of a given situation, starting with conceptualising and particularising the problem and moving through several interventions and evaluations.

No matter what the scenario is, action research involves the following seven distinct phases. These phases are critically important for the inquiring teacher and are shown in Figure 2.2.

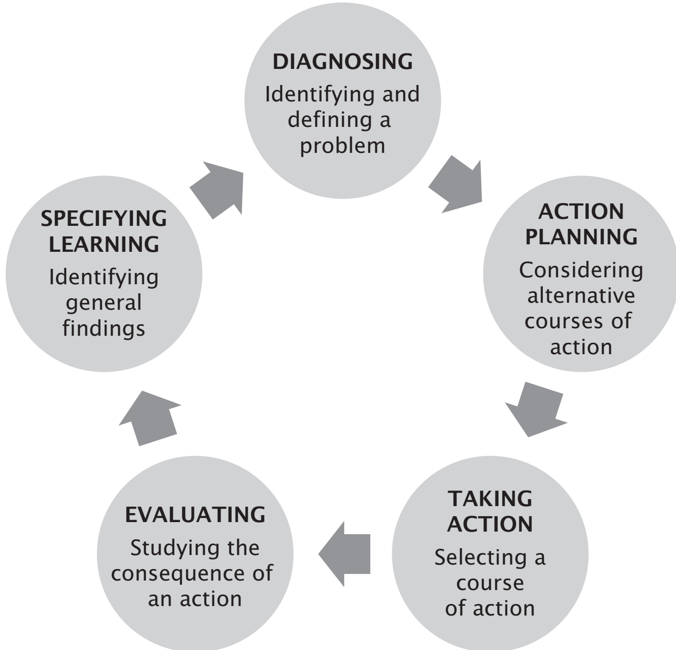


FIGURE 2.2 Action Research Process

Source: Adapted from Susman (1983).

Phase 1: Selecting an area or focus

The first phase starts with serious reflection on selecting an area or focus worthy to be investigated. Knowing the spectacular demands and expectations from today's classroom teachers, it will not be worth doing an activity unless it ensures to make a teacher's work more successful and satisfying. Thus, the selection of an area or focus is critically important. While doing so, the teacher researcher or the team of action researchers should:

- Identify an area of interest
- Focus on students
- Look at both immediate and cumulative effects

The selection of the study area or the focus of the research should be guided by the element(s) of our practice or by the aspect of student learning that we wish to investigate and examine.

Phase 2: Reviewing and clarifying theories

This should then be followed by a critical review of literature, the existing theories in relation to the area or focus of the study. For instance, if teachers are concerned about increasing responsible classroom behaviour, that is, behaviours aimed at developing the whole student, the teacher must be clear from the very beginning by clarifying the approach—using rewards and punishment, allowing student to be ready for the natural consequences of their behaviour or any other relevant strategy—which they consider will be most suitable.

The researcher must identify the following:

- Values
- Beliefs
- Attitudes
- Theoretical perspectives

Phase 3: Identifying research questions
The next step is to develop a set of meaningful research questions to guide the inquiry. Before he or she finalises the research question(s), he or she should be sure that he or she:

- Brainstormed possible research questions
- Did some preliminary research
- Narrowed down the ideas to a maximum of top two or three
- Discussed these questions with all his or her advisors, others faculty peers and so on

Phase 4: Collecting data
In any research, data validity and reliability are two key concerns. To ensure an acceptable validity and reliability, action research should not rely on any single source of data. The researcher should use the process of 'triangulation'—using multiple independent sources of data to answer one's question. The method helps the researcher compare and contrast what is being seen through a variety of lenses.

Phase 5: Analysing data
Here the researcher should illustrate what he or she has learned as related to his or her critical question. He should use his or her data to tell the story of his or her research and support his or her conclusions and emerging theories. This section is the heart and soul of the action research. It is in this process that the teacher researchers will methodically sort, rank and examine their data to answer two generic questions of his story:

- What is the story told by these data?
- Why did the story play itself out this way?

The answers to these two questions can help the teacher researcher to acquire a better understanding of the phenomenon under consideration to improve the situation.

Phase 6: Reporting results
For reporting the results of your research study, simple statistical analyses of quantitative data, such as simple *t*-tests and correlations, are usually sufficient. Tables or graphs are often very helpful. Qualitative data can be analysed for recurring themes, citing supporting evidence. Practical significance, rather than statistical significance, is the goal.

Phase 7: Taking informed action
In this phase of the research study, the researcher should use his or her findings to make decisions about his or her strategies to address critical problems. While suggesting his or her preferred actions, it is equally important that the researcher shares his or her findings with peers in many ways. He or she may submit his or her report to the Faculty of Education or to a special department of the MOE responsible

for reviewing the research reports. He or she might also share his or her work at conferences such as the International Conference for Teacher–Researchers or at regional conferences for his or her discipline.

Scope of Action Research Design in Education

The scope of action research design in education is impressive. It can be used in almost any setting where a problem involving people, tasks and procedures cries out for solution, or where some change of feature results in a more desirable outcome. It can be undertaken by the individual teacher, a group of teachers working cooperatively within one school or a teacher or teachers working alongside a researcher or researchers in a sustained relationship, possibly with other interested parties such as advisers, university departments and sponsors on the periphery (Holly and Whitehead 1986).

The design is widely used in the field of education because there is always room for improvement when it comes to teaching and educating others. We know that teachers use all types of methods of teaching in the classroom; action research works very well because the repetitive cyclic process offers opportunity for continued reflection to improve processes. The research design is also highly useful in areas of teaching practice that need to be explored or settings in which continued improvement is the focus.

Advantages and Limitations of Action Research Design

Advantages

Action research promises progress in professionalisation for teachers, principals and district office personnel. The design permits them to experience problem-solving and to model it for their students. With the data they carefully collect, they use it to diagnose problems, search for solutions, take action on promising possibilities and monitor whether and how well the action worked. The cycle can repeat itself many times, focusing on the same problem or on another. Action research design can help develop a professional problem-solving ethos (Corey 1953; Joyce 1991; Schaefer 1967; Sirotnik 1987).

There are a number of reasons why you might choose to do action research, including for thesis research:

- Action research can stimulate the entire learning community. It can help teachers in changing or reflecting on their classroom practices. It can support initiatives taken by individual teachers, schools, schools working with communities and districts (Box 2.1). Further, researcher can conduct more than one type of action research for a given setting at the same time.
- When practitioners use action research, it has the potential to increase the amount they learn consciously from their experience. The action research cycle can also be regarded as a learning cycle (see Kolb 1984). The educator (Schön 1983) argues strongly that systematic reflection is an effective way for practitioners to learn.
- Action research is usually participative. This implies a partnership between you and your clients. You may find this more ethically satisfying. For some purposes, it may also be more occupationally relevant.

Limitations

So why does not everyone use it? An ignorance of action research is not a reason to avoid it. There are good reasons, however, why you may decide to stay within mainstream research. For example, here are some of the costs of choosing action research as your research paradigm.

BOX 2.1: Community Learning Centres (CLCs) and the Process of Community Ownership: Issues and Challenges

Let us take an imaginary country Ardania. Since 1995, the country has been implementing its Community Learning Centres (CLCs) Project and has established 15,000 CLCs for addressing the demand for adult education. From the available documentation and evidences, the country finds out that the emerging issues for CLCs in recent times have been the issue of sustainability and that of the interaction between external intervention and internal autonomy. The project has fostered to a large extent the 'participation and involvement' of community people, but not so much the process of 'community ownership' in the planning, implementation and assessment of activities.

For this particular case, the possible research questions that you may formulate could be:

- How can the interaction of community ownership and external intervention be balanced to make CLCs effective and sustainable?
- To what extent has the CLC succeeded in mobilising and involving community members in its design and operations? Does the community that a CLC serves really feel a sense of ownership?
- What are the challenges and difficulties encountered in CLC ownership?
- What is the impact of the involvement of external agencies on sustainability of CLCs?
- Is there a balance between local community ownership and the involvement of external agencies?
- Does the relationship contribute to enhancing the sustainability of the CLCs? And so on.

These are complex questions demanding empirical investigation into the actual experience of establishing and running CLCs.

Depending on the nature and complexities of the above research questions, the objectives of your research could be as follows:

- To analyse the nature and means of community participation and identify good practices or models of CLCs which demonstrate community ownership;
- To study the state of balance between community involvement and external support and intervention and the implications for the sustainability of CLCs and
- To identify key issues in community participation and ownership, and explore effective strategies for future action at national and CLC levels

The methodology that you can follow for this particular study is the 'action research' involving documentary analysis, structured and unstructured interviews and observation. This triangulation of research will provide you with firm validity and reliability bases for your study.

As the total number of CLCs in Ardania amounts to 15,000, you may find sampling of CLCs a little bit of a difficult exercise. In terms of documentary evidence, you may also find the lack of records and a weak system or no system for collection and collation of data.

For this particular case, your sampling design will necessarily be small given the usual restraints of time and costs. However, the sites and number of CLCs that would be selected by you for the research should be representative of the range of CLCs, providing valid information through interviews and observation (for detailed explanation of sampling design, data collection and analysis, survey design and questionnaire design, refer to Chapters 2 and 3).

In this particular case, the framework of your research, for example, could be as follows:

(Continued)

(Continued)

Section 1: Rationale for the Study

Section 2: Research Methodology for the Study
- 2.1: Sampling Design
- 2.2: Survey Techniques
- 2.3: Scope and Limitations of the Study

Section 3: Review of Literature of the Research Topic

Section 4: Focus of the Community Learning Centres
- 4.1: Government-supported CLCs
- 4.2: Community CLCs
- 4.3: Number of CLCs and their Regional Distribution

Section 5: Community Learning Centres: A Situational Survey
- 5.1: Functions of CLCs
- 5.2: Programmes of CLCs

Section 6: Community Participation and Ownership
- 6.1: Advantages
- 6.2: Future Directions for Sustainability

Section 7: External Interventions
- 7.1: Advantages
- 7.2: Future Directions for Sustainability

Section 8: Recommendations on Sustainability
- 8.1: The Relationship Between Leaders and Members
- 8.2: Freedom in Expressing Their Ideas
- 8.3: Characteristics of Learning Activities
- 8.4: CLC Management
- 8.5: CLC Leaders

Section 9: Conclusions

Source: Adapted from UNESCO-PROAP.

- It is harder to do than conventional research. You take on responsibilities for change as well as for research. In addition, as with other field research, it involves you in more work to set it up, and you do not get any credit for that.
- It does not accord with the expectations of some examiners. Deliberately and for good reason, it ignores some requirements which have become part of the ideology of some conventional research. In that sense, it is countercultural. Because of this, some examiners find it hard to judge it on its merits. They do not recognise that it has a different tradition and is based on a methodological perspective and principles different to their own.
- You probably do not know much about action research. Action research methodology is something that you probably have to learn almost from scratch.

- The library work for action research is more demanding. In conventional research, you know ahead of time what literature is relevant. In most forms of action research, the relevant literature is defined by the data you collect and your interpretation of it. That means that you begin collecting data first, and then go to the literature to challenge your findings. This is also true of some other forms of field research, though certainly not all.
- Action research is much harder to report, at least for thesis purposes. If you stay close to the research mainstream, you do not have to take the same pains to justify what you do. For action research, you have to justify your overall approach. You have to do this well enough that even if examiners do not agree with your approach; they have to acknowledge that you have provided an adequate rationale (this may be true for other methodologies outside the research mainstream too).
- All else being equal, an action research thesis is likely to be longer than a conventional thesis. As already mentioned, you have to provide a more compelling justification for what you do. In effect, you have to write two theses. One reports your method, results and interpretation. The other explains why these were appropriate for the research situation. In addition, if you use qualitative data (and you probably will), that also tends to take more space to report.

This is particularly relevant for those of you doing a thesis where page limits or word limits are imposed. If there is such a limit, you have to write very succinctly, yet do so without undermining your thesis or your justification.[1]

For most of you and other people, these disadvantages outweigh the advantages. Above all, if you are choosing action research because you think it may be an easier option, you are clearly mistaken. In reality, it is more demanding and more difficult.

In short, action research in education is a valuable exercise for teacher researchers, school administrative staff and other stakeholders to undertake and understand the teaching–learning environment and to improve the quality of the educative process. It allows teachers to follow a systematic, collaborative and participatory process of inquiry that actively seeks to address areas of concern or redress. Further, action research enhances teacher researchers' technical skills and specialised knowledge essential to effect positive change(s) within classrooms, schools and communities. The approach is systematic and participatory in nature. It offers multiple, beneficial opportunities for teachers wishing to develop 'customised' action research projects of their own. Action research improves instruction for students and empowers teachers, since it is a tool that allows them to judge their own efforts (self-evaluation) and gauge the outcomes of their practices.

Action research is designed to apply its findings to a particular situation. It provides a host of interventions to influence change in any given situation and to monitor and evaluate the results. Working with a client, the researcher identifies a particular objective such as ways of improving class attendance in school. He then suggests strategic ways this might be done. He enters into the situation, for example, by introducing new techniques and monitors the results. Action research, thus, requires active cooperation between researcher and client, and a continual process of adjustment to the intervention in the light of new information and responses to it from respondents. Box 2.1 exemplifies a systematic approach for conducting an action research in education.

[1] See http://www.aral.com.au/ (accessed on 10 June 2017).

Case Study

The term case study refers to both a method of analysis and a specific research design for examining a problem. Both the terms are used in most circumstances to generalise across populations. In this section, we focus on the latter—how to design and organise a research study in education that analyses a specific case.

A case study explores and analyses the life or functioning of a social or economic unit, such as a person, a family, a place, an event, a phenomenon, a community, an institution, a firm or an industry or other type of subject of analysis in order to extrapolate key themes and results that help predict future trends and illuminate previously hidden issues that can be applied to practise. It helps researchers understand an important research problem with greater clarity.

Case study research usually examines intensely an individual (single subject of analysis), but you can also organise a case study involving a small group of participants and draw conclusions only about that participant or group and only in that particular context. The purpose of your case study in this case would be a comparative investigation that shows relationships between two or more than two subjects.

In the case study method, the aim of the researcher is neither to discover a universal, generalisable truth, nor to study the cause-effect relationships; instead, prominence is placed on exploration and description. It is important to mention here that the methods used to study a case can rest within a qualitative, quantitative or mixed method investigative paradigm.

> 'A case study refers to the collection and presentation of detailed information about a particular participant or small group, frequently including the accounts of subjects themselves' (Robson 2002).
>
> 'A case study is documented study of a specific real-life situation or imagined scenario, used as a training tool in business schools and firms. Students or trainees are required to analyse the prescribed cases and present their interpretations or solutions, supported by the line of reasoning employed and assumptions made' (*Business Dictionary*).

A case study is a specific research method that is frequently designed to illustrate a more general principle. It is 'the study of an instance in action' (Adelman and Kemmis 1980). The single instance is of a bounded system, for example, a child, a class, a school or a community. For Nisbet and Watt (1984), 'a case study provides a unique example of real people in real situations, enabling readers to understand ideas more clearly than simply by presenting them with abstract theories or principles'. This shows that a case study can enable you understand how ideas and abstract principles can fit together. For Robson (2002), 'case studies can penetrate situations in ways that are not always susceptible to numerical analysis. They opt for analytic rather than statistical generalization; that is, they develop a theory which can help researchers to understand other similar cases, phenomena or situations'.

The purpose of case study is not to represent the world, but to represent the case. Wallace (1998) proposes the following four key aims of the case study research:

- Solving specific problems
- Applying theories into practice
- Generating hypotheses
- Providing illustrations

In the case study method, researchers generally use two distinct research approaches. In the first approach, they prepare an in-depth study of a particular student, classroom or school with the aim of producing a nuanced description of the pervading cultural setting that affects education and an account of the interactions that take place between students and other relevant persons. A good example could be an in-depth exploration of the patterns of friendship between students in a single class (Postlethwaite 2005). The second approach to case study research involves the application of quantitative research methods to non-probability samples, which provide results that are not necessarily designed to be generalisable to wider populations, for example, a survey of the reading achievements of the students in one rural region of a particular country.

Researchers use case studies to establish cause and effect. An important strength of case studies is that the researcher discerns effects in real contexts, admitting that context is a powerful determinant of both causes and effects. Case studies are highly pertinent when the researcher is required not only to investigate but also to report the interaction of interwoven events, human relationships and other factors in typical situations. Case studies are significantly important when the researcher has very little or no control over events.

Case studies can be used to formulate theories or be:

- Exploratory (e.g., as pilot to other studies or research questions)
- Descriptive (e.g., providing narrative accounts of events or current practice)
- Illustrative (e.g., illustrating new practices adopted by an organisation)
- Experimental (e.g., examining difficulties in adopting new practices or procedures)
- Explanatory (e.g., testing theories)

In a case study, an investigator surveys a single case or several cases over time. In-depth data by using diverse sources of information such as observations, interviews, audiovisual material and documents are collected and finally reported on a case (Box 2.2). A case study is basically a methodology or strategy of inquiry. It enables the researcher to study an issue through specific cases. In case studies, emphasis is placed on exploration and description.

BOX 2.2: 'A Case Study of Teaching Students with Attention Deficit'

Six years old Tania moved into a Municipal Elementary School in January of her grade 1 year. Her mother met Miss Aruna, the grade 1 teacher. She informed the teacher that since attending the kindergarten, Tania needed some special attention. She and her husband were looking for any suggestions the school could provide in managing Tania at home as well. Miss Aruna indicated that she would review Tania's file and asked Tania's mother if she and her husband could come in to meet with her and the school-based team next week to discuss Tania's programme.

During the first week, Miss Aruna made the following observations:

- Tania is cheerful and friendly. She seems keenly interested in pleasing the teacher and her classmates.
- Tania appears to have a strong understanding of verbally presented information, recognises colours and can count to 100.
- Tania's literacy skills are at the emergent stage—she cannot recall letter names and does not appear to have any sight vocabulary.

(Continued)

(Continued)

- Maintaining one-to-one correspondence with objects while counting is difficult for Tania.
- Tania completes 2 out of 20 questions when not medicated; she completes entire sheet of 20 questions when she has taken her medication.
- During both individual and group instruction, Tania frequently interrupts to ask unrelated questions and change topics.
- When interacting with peers, Tania constantly changes topics and commonly leaves an activity or game while others continue to play.

Information from Tania's file indicated that she had received a psychological assessment and had been identified as having attention deficit (AD)/hyperactive disorder (HD) as well as learning disabilities. She had been placed on a wait list for a special class placement.

Miss Aruna brought Tania's case forward to the school-based team meeting so that planning could take place immediately. Her parents were invited and were able to meet with the team of the second week Tania had been enrolled at the new school.

At the meeting, the team agreed that acquiring literacy skills and helping Tania to focus on the topic at hand were the most important goals to begin with. The following plan was developed.

Accommodation/Support Plan

Name: Tania	Grade: 1
Date: 21 January 2016	School: Elementary
Completed by: Mr Rehman (local resident).	Review Date: March 2016

The following factors were considered to study Tania's behaviour and learning skills.

1. Indicate the student's areas of strength:

Academics

_____ Reading decoding
_____ Reading comprehension
_____ Written expression
_____ Mathematics—recall of basic facts 'can count to 100'
_____ Mathematics—conceptual understanding
_____ Rich oral vocabulary
_____ Ability to understand complex concepts
_____ Spelling
_____ Responds well to praise
_____ Other, please specify
_____ Other, please specify ____________________

Personal Skills

_____ Computers/technology
_____ Interacting with peers
_____ Memory
_____ Leadership Skills
_____ Interacting with adults
_____ Requests help when needed
_____ Sense of humour
_____ Demonstrates enthusiasm

Areas of Interest

_____ Computers/video games
_____ Television/movies
_____ Sports (specify)
_____ Visual arts
_____ Music (specify)
_____ Volunteer activities (specify)
_____ Other accomplishments

Comments:
____
____
____
____
____
____

2. Indicate areas of concern that are significantly affecting the student's ability to learn and interact with others at school:

Academics

_____ Memory
_____ Understanding and following instructions (Describe) ___________________________
_____ Reading decoding
_____ Reading comprehension
_____ Written expression
_____ Mathematics—recall of basic facts
_____ Mathematics—conceptual understanding
_____ One-to-one correspondence
_____ Limited oral vocabulary
_____ Ability to understand complex concepts
_____ Spelling
_____ Other, please specify _____________

Behaviour

_____ Interacting with adults
_____ Motor activity detrimental to learning
_____ Interrupting, blurting out, inappropriate verbalisations talks constantly, changes topics, leaves activities
_____ Interacting with peers in class
_____ Interacting with peers at lunch and recess breaks
_____ Complying with staff requests
_____ Transitions between activities or classes
_____ Behaviour during loosely structured activities (assemblies, field trips and so on)
_____ Attendance
_____ Other, please specify _____________

Organisation

_____ Handing in assignments
_____ Keeping track of necessary materials
_____ Time management
_____ Completing tasks
_____ Getting started on assigned work
_____ Comments

Source: *Adapted from* http://www.bced.gov.bc.ca/specialed/adhd/case.htm (accessed on 15 May 2017).

How to Approach Writing a Case Study

You should realise that identification of your research problem for investigation is not a quick or easy task. The moment you finally decide to do the research, you should start thinking about it since then. Generally, your university asks you to write about your research problem in either of three ways: (a) your research supervisor (professor) suggests a general topic from which you choose and study a particular aspect, or (b) he or she gives a list of possible topics close to your interest to study or (c) the selection of topic is exclusively on you and you only have to obtain permission from the university dean to write about it before the commencement of your study. Once it is decided, the following paragraphs provide you critical steps to guide your choice of the topic of your investigation (Box 2.1).

Identification of a Research Topic

For any research endeavour, the first task that you will be required to do is to select a meaningful topic for your investigation. How can you do that? If you follow the following steps, you will certainly be in a position to find your topic.

Case 1: Your research advisor has given you the topic

Step 1: The first thing you have to do is to identify concepts and terms that make-up the topic statement of your study. For example, your professor wants the class to focus on the following research problem: 'Is United Nations Educational, Scientific and Cultural Organisation (UNESCO) a credible organisation with the capacity to help provide technical assistance to the member states to realise the Six Education for All (EFA) goals globally?' The key concepts in this particular problem are: UNESCO, EFA goals, credibility and technical assistance (hint: focus on identifying proper nouns, nouns or noun phrases and action verbs in the assignment description).

Step 2: Start reviewing relevant literature close to your topic to help refine the means you will use to investigate your topic as well as a means you will deploy to analyse the topic. You can do this by means of any of the following: reading reference and other materials prescribed in your university syllabus, finding out from the library a recent book, newspapers, specialised educational journals, references cited by authors in footnotes, bibliography and so on. This will help you refine and frame the scope of your research problem. You have to do this quite often before you finalise how to approach writing about the topic.

Case 2: Your research advisor provided a list of possible topics

Step 1: From the list, choose a topic which you consider interesting, engaging and motivating in some way.

Once you have done this, follow steps 1–4 discussed above (Case 1).

Case 3: Your research advisor leaves it up to you to choose

Step 1: Here you have to ask yourself the question: 'What do I want to know?' Consider an open-ended assignment as an opportunity to learn about something that is innovative or exiting to you.

Step 2: Follow steps 1–4 listed above (Case 1).

Now you are fully ready to take off. Go ahead.

Writing a Case Study Research Paper

A research problem is the key principle that will guide the analysis of your research study. The problem that you intend to investigate will offer you an occasion for writing and a focus that governs what you want to say. It represents the core subject matter of an effective and scholarly communication and the means by which you can arrive at other topics of conversations and the discovery of new knowledge and understanding.

Identifying a particular case for investigation in a case study involves more than choosing the research problem. You have seen above that a case study surrounds a problem contextualised around the application of in-depth analysis, interpretation and discussion often resulting in specific recommendations for action or for ameliorating existing conditions. You have taken into account practical realities into consideration such as time and access to information, which eventually can influence your case selection. But these issues should be your sole selection criterion in describing the methodological justification for identifying a particular case study. Considering this, your selection choice of selecting a case should be based on the following consideration:

- Is your case study a reflection of an unusual example of a research problem that requires more in-depth analysis? Cases often represent a topic that rests on the borders of earlier investigations. Let us consider the following example.

Example

Suppose you want to study and identify strategies that may help to improve policies that support girls' access to secondary education in predominantly Muslim countries, you could consider using Uzbekistan as a case study rather than selecting Saudi Arabia, Qatar or any other Middle Eastern nation. Such selection may reveal eye-opening insight into recommending how governments in other predominantly Muslim countries can identify and formulate policies that support improved access to girls' education.

- Is your case study highlighting a previously hidden problem? Your case study should be based on the hypothesis that it will reveal trends or issues that have not been uncovered in prior studies. Or it will reveal new and important implications for practice.

Example

Suppose, anecdotal evidence may suggest that drug use among poor homeless parents of primary schoolchildren in cities and towns in developing countries is related to their travel throughout the day. Assuming that prior studies have not glanced the individual travel choices to study access to illicit drug use, your case study observes that a homeless parent veteran could reveal how issues of personal mobility choices facilitate regular access to illicit drugs. However, you should note that, for this topic of your case study, you have conducted a thorough literature review to ensure that your assumption about the need to reveal new insights is valid and based on evidence.

- Is your case study challenging and offers a counterpoint to prevailing assumptions? Perhaps your case study may fall into trap of developing assumption based on outdated studies, which were not thoroughly tested in practice. Thus, your case study should be such that it offers an opportunity to gather evidence that challenges prevailing assumptions about a research problem and provide new insight and new recommendations that can be applied to practice that have been not tested and validated earlier.

Example

Suppose, perhaps researchers have been applying for a long time a particular theory in explaining the relationship between two subjects of analysis. The topic of your case could challenge this assumption by applying an innovative theoretical framework to study a case in order to explore whether this approach offers new ways of understanding the research problem. Taking a contrary instance is one of the most important ways that new knowledge and understanding develops from existing literature.

- Is your case study providing an opportunity to pursue action leading to the resolution of a problem? Another possible way of identifying a topic for your case study is to reflect how the results of your study from investigating a particular problem may result in findings that demonstrate ways in which to reveal an existing or emerging problem.

Example

Suppose, you want to study the case of an unforeseen incident, such as a fire in the school building, to disclose hidden issues that could be applied to preventative measures that contribute to reducing the chance of fire accidents in the future. In this case, you would investigate an accidental fire that could lead to a better understanding of the sensitive places in the school where additional firefighters can be strategically located. Consequently, you can better warn the school staff of possible reasons of catching fire in the school building, particularly when the school kitchen is operating, people are smoking and fire-prone electricity fittings are present (leading to the possibility of short-circuits, etc.).

- Is your case offering a new direction for future research? Make sure that your topic is exploratory in nature, that is, it demands a further examination of the research problem in future. You may come across a situation where various research studies have predicted an outcome on your topic of study. How best can you handle a situation like this?

  Example

  Suppose after conducting a thorough literature review, you discover that little research exists showing the ways in which mothers contribute to preventing girls to enrol in primary schools in rural communities of Tanzania. A case study of how mothers contribute to allow girls enrol in schools in a particular village can lay the foundation for understanding the need for more thorough research that documents how mothers in their roles as head of the families think about value of girls education as a valuable resource within their community throughout rural regions of East Africa (for instance, Uganda, Kenya, Burundi and Rwanda). The case could also point to the need for scholars to apply feminist theories of work and family to the issue of girls' education.

Structure and Writing Style

Your case study report should follow a standard structure and writing style in accordance to the requirements of your faculty of the university/college. The following tag presents the key elements that will help you to organise and write your case study research report.

Introduction

The introductory chapter of your report should contain the following vital information:

- What I am studying? This should include a description of both your research problem and the subject of analysis you have taken to address the problem. Further, you should also establish the link between them and describe the elements you have included to help to expand the knowledge and understanding of the problem.
- Why this topic is important? Explain the importance and significance of your topic in relation to present status of the problem. Also, explain how your research topic is appropriate in addressing the given problem.
- What do we know about this topic before you undertake this study? This should be directed to an intensive literature review and finally justify why it fails to adequately address the research problem. Explain why your study will be meaningful. If from your literature review you find that your topic is a new topic, then give reasons of your choice of this particular topic.
- How will your study advance new knowledge? Give argument(s) why your study will be fitting in helping to expand and understand about your research problem.

Your introduction comprising each of the above questions should not exceed more than two typed pages. Only exception to this is when you are explaining a complex topic which demands more in-depth background information.[2]

[2] See http://libguides.usc.edu/writingguide (accessed on 10 June 2017).

Literature Review

While describing the literature review, do the following:

- Place relevant works in the context of their contribution to understand the case study being investigated.
- Describe the relationship each work has to the others under consideration that informs the reader why this case is applicable.
- Identify new ways to interpret prior research using the case study.
- Resolve conflicts amongst seemingly contradictory previous studies.
- Point the way in fulfilling a need for additional research.
- Expose any gaps that exist in the literature that the case study could help to fill.
- Summarise any literature that not only shows how your subject of analysis contributes to understanding the research problem but how your case contributes to a new way of understanding the problem that prior researches has failed to do.
- Locate your own research within the context of existing literature. This is very important.

Method

In this section of your case study report, you should explain very precisely the reasons for selecting this particular research topic. You should also explain the strategy you will use to identify and finally decide that your case is appropriate in addressing the research problem. However, the manner you explain and describe the methods you will be using may vary depending on the type of subject of your analysis that will frame your study. You should keep the following key points in mind when describing the method:

- Is your subject of analysis an incident or event?
- Is your subject of analysis a person, a place or a phenomenon?

Evidence that supports the method by which you identified and chose your subject of analysis should be linked to the findings from the literature review. Be sure to cite any prior studies that helped you determine that the case you chose was appropriate for investigating the research problem.

Discussion

In your dissertation around a case study, it is better that you combine a description of findings with discussion about their implications. The objective of your discussion section should contain the following:

- Reiterate the research problem and then state the major findings.
- Explain the meaning of the findings and why they are important.
- Relate the findings to similar studies.
- Consider alternative explanations of the findings.
- Acknowledge the study’s limitations.
- Suggest areas for further research.

Conclusion

The following points will help you ensure that your conclusion is appropriate (Creswell 2012):

- If the argument or purpose of your paper is complex, you may need to summarise these points for your reader.
- If prior to your conclusion, you have not yet explained the significance of your findings or if you are proceeding inductively, use the conclusion of your paper to describe your main points and explain their significance.
- Move from a detailed to a general level of consideration of the case study's findings that returns the topic to the context provided by the introduction or within a new context that emerges from your case study findings.

It is important to mention here that, depending on the discipline you are writing in and your professor's preferences, the concluding paragraph may contain your final reflections on the evidence presented applied to practice or on the dissertation's central research problem. However, the nature of being introspective about the subject of analysis you have investigated will depend on whether you are explicitly asked to express your observations in this way.

Finally, while writing your report, there are certain things that you should avoid as near as possible. They include:[3]

> **Overgeneralisation:** One of the goals of a case study is to lay a foundation for understanding broader trends and issues applied to similar circumstances. However, be careful when drawing conclusions from your case study. They must be evidence-based and grounded in the results of the study; otherwise, it is merely speculation.
> **Failure to Document Limitations**: No case is going to reveal all that needs to be understood about a research problem. Therefore, just as you have to clearly state the limitations of a general research study, you must describe the specific limitations inherent in the subject of analysis.
> **Failure to Extrapolate All Possible Implications**: Just as you do not want to overgeneralise from your case study findings, you also have to be thorough in the consideration of all possible outcomes or recommendations derived from your findings. If you do not, your reader may question the validity of your entire analysis, particularly if you failed to document an obvious outcome from your case study research.

When designing your case study, be sure you have thoroughly addressed all aspects of the problem and do not leave gaps in your analysis.

Strengths of Case Study Research Design

Strengths

Case studies help us generate new ideas that researchers can test by other methods. In educational research, case studies are used as means of illustrating theories or reforms. At the same time, they are

[3] See http://libguides.usc.edu/writingguide (accessed on 10 June 2017).

also used to describe how different aspects of a person's life are related to each other. For those researchers who adopt a holistic point of view, case study method is quite useful research method.

Case studies provide detailed and rich qualitative information. Because the outcomes and results of case studies are mostly written in simple, non-professional and non-technical language, they are easily comprehended and grasped by a wide audience including the non-professionals. As the results speak for themselves, they are immediately intelligible.

Case studies, by providing insights into other similar situations and cases, can assist interpretation of other similar cases. In this way, case studies provide insight for further research. Since case studies are in-depth, multi-sided studies, they help researchers capture unique features that may otherwise be lost in large-scale data. These unique features of case studies help understand a given case or situation. Case studies, thus, are strong on reality. With case studies, researchers can ascertain unethical aspects of human thinking and behaviour that would be unethical or impractical to study in other ways.

In brief, a case study research design:

- Brings us to an understanding of a complex issue through detailed contextual analysis of a limited number of events or conditions and their relationships
- Applies a variety of methodologies and relies on a variety of sources to investigate a research problem
- Extends experience or adds strength to what is already known through previous research
- Is used, in particular, to make wide use of this research design to examine contemporary real-life situations and provide the basis for the application of concepts and theories and the extension of methodologies
- Provides detailed descriptions of specific and rare cases

Limitations

Case studies are not free from weaknesses. They cannot generalise the results to the wider population. Case studies are greatly influenced by researchers' biases and subjective behaviours and feelings. A case study is difficult to replicate as well as time-consuming.

Since a case study relates to only one person, one event or one group, a researcher can never be sure of the possibility of application of conclusions of this particular case elsewhere. Generally, the results of a case study cannot be generalised on a wider body of similar instances. In other words, it is difficult to say whether the case study is representative.

Because conclusions are drawn from the analysis of qualitative (i.e., descriptive) data, a lot depends on how the researcher interprets the data he has gathered for the case study. His interpretation may be influenced by biased attitude and his subjective opinion which in turn may lead to a wrong assessment of what the data really mean.

Thus, we can say that a case study research design:

- Offers little basis for establishing reliability or to generalise the findings to a wider population of people, places or things because of a single or small number of cases
- May bias a researcher's interpretation of the findings due to intense exposure to the study of a case
- Does not facilitate assessment of cause-and-effect relationships
- May miss vital information thus making the case hard to interpret

- May not be representative or typical of the larger problem being investigated
- Represents a very unusual or unique phenomenon or problem for study, so your interpretation of the findings can only apply to that particular case

In a nutshell, case studies are generally based on multiple sources of data. They produce large amount of data and thus are complex. This method of research is widely used in almost all disciplines to build upon theory and claim, to discover new theory, to argue or challenge theory, to explain a situation, to provide a ground or basis to apply solutions to situations, to explore or to describe an object or phenomenon. The case study method can effectively be used to study and describe real-life, contemporary, human situations. The written reports produced are readily accessible for use by the general public. Case study results report directly to the common reader's everyday experience and facilitate an understanding of complex real-life situations.

Ethnographic Research Design

The intensive investigation of a single site or context has been the hallmark for most of the twentieth century. Social anthropologists initially developed these designs. Historical developments of ethnographic research design can be traced from the writings of Bogdan and Biklen (1998), and LeCompte and Preissle (1993). The method was first applied in educational research in the 1950s.

The ethnographic approach to educational research uses a variety of names: the humanistic approach (Walker and Evers 1997), constructivist research (Magoon 1977), naturalistic inquiry (Lincoln and Guba 1985), the symbolic approach (Popkewitz 1984), ethnographic research (Dobbert 1982; Taft 1997; Wolcott 1988), interpretive research (Carr and Kemmis 1986; Erickson 1986) and qualitative research (Bogdan and Biklen 1998; Eisner 1991; Glaser and Strauss 1967; Martella, Nelson and Marchand (1998)).

Definition of Ethnographic Research

Ethnographic research design aims primarily at describing what is happening in a particular setting together with the participant's perspectives on these events. The design focuses on all of the events which occur in a particular setting (rather than on just one or two types of events). It usually provides a holistic picture of how a particular social group (such as a classroom) operates and accomplishes this by means of direct observation and interviews with key participants. Since ethnographic research attempts to explain people's behaviour in terms of the beliefs which people hold about their behaviour, the design is commonly referred to as interpretive research. The design is the most common kind of ethnographic research currently being undertaken in classrooms.

> 'Ethnography literally means "a portrait of a people". An ethnography is a written description of a particular culture—the customs, beliefs, and behaviour—based on information collected through fieldwork' (Harris and Johnson 2000).

Ethnographic research method refers to fieldwork (alternatively participant observation) conducted by a single investigator who 'lives with and lives like' those who are studied (Van Maanen 1996). In this method, the researcher usually observes target users in their natural, real-world setting, and not in an artificial environment of a lab or focus group. The method provides an accurate and deep intuitive understanding about the way people live, what they do, how they use things or what they need in their everyday or professional lives. The researcher lives with the people and becomes a part of their culture.

Ethnographic research is highly useful to single out new or currently unmet user needs. The approach is most valuable at the beginning of a project when the researcher does not have the slightest idea of real end user needs or to perceive the limitations of using a new product or service by people considered in the research study.

Green, Skukauskaite and Baker (2011) suggest that in education, ethnographers enter a classroom, school, family group or community setting to identify inside knowledge by asking questions such as:

- What is happening in the classroom?
- What is being realised, by and with whom, how, in what ways, when and where, under what conditions, for what purposes with what outcomes or consequences for individuals and the group?
- Do individual members have access, orient and hold each other accountable?
- What factors make someone an insider or outsider of particular groups (e.g., class, group within a class, peer group or social network)?
- What criteria do they use for determining disciplinary knowledge (i.e., knowledge of mathematics, science, social science or art) in this particular group or classroom?
- What roles and relationships, norms and expectations, and rights and obligations are practised?
- How does previously constructed cultural knowledge support or constrain participation or create clashes with local knowledge developed in a particular event (or social group)?
- How do decisions beyond the group support constrain ways of knowing members?

They inform that questions such as these have been used to guide ethnographic research in education. Walford (2008) argues that by asking such questions the 'ethnographer tries to make sense of what people are doing and hopes gradually to come to an understanding of the way we do things around here'.

Educational Issues, Topics and Directions

To make visible a range of topics, issues and directions that have been studied ethnographically in education, Green, Skukauskaite and Baker (2011) present the following sketch map of programmes of research across national contexts:

- *Cross-national comparative studies of education and policy–practice relationships* (Alexander 2001; Anderson-Levitt 2002; Castanheira 2004; Kalman and Street 2010; Rockwell 2002; Street 2005; Tobin et al. 2009);
- *Community-based studies of cultural processes and practices* (Brayboy and Deyhle 2000; Delamont 2002; Heath 1983; Philips 1983; Spindler and Hammond 2006);
- *Impact of changing policies on opportunities for learning and teaching* (Carspecken and Walford 2001; Levinson et al. 2002; Mcneil and Coppola 2006; Smith et al. 1987; Stevick and Levinson 2007; Troman et al. 2006);

- *Linguistic and cultural differences between home and school* (Cazden et al. 1972; Gonzalez et al. 2005; Vine 2003);
- *Literacy and discourse practices in homes, schools and communities* (Barton and Tusting 2005; Bloome et al. 2005; Jennings et al. 2010; Martin- Jones et al. 2008; Orellana 1996);
- *Peer culture and social development in school and community contexts* (Corsaro 2003);
- *Learning and teaching relationships as social constructions in classrooms and other educational settings* (Edwards and Mercer 1987; Green and Wallat 1981; Jeffrey and Woods 2003; Mehan 1979; Rex 2006; Santa Barbara Classroom Discourse Group 1992a, 1992b);
- *Disciplinary knowledge in science*
- *Literacy and social constructions in educational contexts* (Bloome et al. 2005; Cochran-Smith 1984). among other subject matter, as;
- *Ways that access to technology in schools is shaped by policy decisions and instructional processes inside and outside of classrooms* (Kitson et al. 2007).

These studies demonstrate the breadth of ethnographic research in education and the range of questions of global, national and local concern arising in the complex social, cultural, linguistic, economic and political contexts in which education is conducted.

Ethnographers interview the most knowledgeable people about the culture. These people are called key informants. They generally collect data through participant observation and interviews. Researchers bracket, or make explicit, their own personal biases and beliefs, set them aside and then try to understand the daily lives of individuals as they live them. Data collection and analysis occur simultaneously. As understanding of the data occurs, new questions emerge. Basically, ethnography is the development of cultural theories (Box 2.3).

BOX 2.3: Action Research

Peine (2003) conducted a grounded theory research to investigate gifted students' experience of sitting and waiting in a regular classroom. He used purposive sampling to select 16 participants. The data were collected through a combination of methods, including semi-structured interviews, field notes from classroom observations and conversations with teachers and maps that were drawn by the participants. Peine further claimed that maps were used as a data source because they allowed the participants to represent their learning setting in an alternative way. In open coding, 3 main categories and 15 sub-categories were formed. The main categories were 'School and Classroom Structure Waiting', 'Instructional Waiting' and 'Assignment Waiting'. Then in axial coding stage, the researcher used a model to illustrate the relationships that emerged among the main categories and sub-categories. The researcher used six components to construct the model, namely, (a) the phenomenon, (b) causal conditions, (c) the context, (d) intervening conditions, (e) action/interaction strategies and (f) consequences. Finally, in the selective coding stage, the core category was expressed as a grounded theory statement, namely, 'Waiting is boring; sometimes waiting is fair'. The researcher explicated the grounded theory by using three propositions: (a) 'already knowing'; (b) 'adjusted doing' and (c) 'being fair'. Based upon Peine, explication of the theory through propositions statements allows the voices of the participants providing concrete evidence to support the generated theory.

Source: Adapted from Peine (2003).

How Can You Collect Data in Ethnography?

Observations, video diaries, photographs, contextual interviews and analysis of artefacts are the most commonly used techniques in ethnographic research. In this method, you can observe the participant at home, at work or in leisure environments. You can study them with their family, on their own, with working colleagues or as part of a group of friends. Often a single participant is selected for the study, but you can also include several more family members or friends and organise interviews with them.

Data are collected by means of either contextual interview of the duration of 4–5 hours, or following participant for several days or a longitudinal study lasting for several weeks or months to investigate. As mentioned above, the process can take the form of an in-depth interview in a person's home or it might involve a person simply maintaining his own video diary over a period of time. Depending on the study needs and the approach, 6–8 weeks from briefing to results can provide rich insight.

There are three kinds of data collection mechanisms in a typical ethnographic research: interviews, observation and documents. Each of these yields three kinds of data: quotations, descriptions and excerpts of documents resulting in one final product: the narrative description. This narrative contains charts, diagrams and additional artefacts that help to tell 'the story'. Ethnographic methods can give shape to new constructs or paradigms and new variables for further empirical testing in the field or through traditional, quantitative social science methods (Hammersley and Gomm 2000).

Ethnographic studies are basically quite expensive and time-consuming, but this depends on the needs of a particular project. However, the accrued benefits can be extremely valuable.

What Should You Do Before You Get Started?

Before you start the study, you have to complete certain general activities. The most important is the identification of your research question. You should do this by asking the question you want to ask and not that of someone else. It is better that you talk to others about your research study and refer to different sources for ideas. It is also necessary that you assess yourself your depth of knowledge about the subject of your research study. For getting access to the setting, you should approach those people (e.g., gatekeeper) who will sanction you permission. It may take sometime or turn out to be difficult if you are unknown to the setting. Finally, you should know your resources and time that you can afford for your study, that is: Do you have the time to conduct the study? What about your other obligations and commitments? What types of resources do you have? What types of equipment will you need? Who will fund the study? These are important questions you have to consider and seek answers before beginning your study.

Stages in Conducting Ethnography

Singleton and Straits (2005) identified the following stages in field research:

- Problem formulation: Defining the main focus of the study by formulating the problem about which you wish to learn more.
- Selecting a research setting: The first question is to know and decide where to start. You must ensure that the setting permit clear observation.
- Gaining access: How do you get into a group that you wish to study? You may need to seek formal permission which can be facilitated if you have a friend who can vouch for you.

- Presenting oneself: You need to decide how you will present yourself to those in the field. What roles will you need to adopt and relate to others? How actively will you be participating in other people's lives?
- Gathering and recording information: Sometimes it can be difficult to record and gather data at the same time. What are the types of information that should be recorded or taken as field notes? If you cannot fully record your observations while you are in the field, what should you do? You should always carry a notepad for brief jottings.
  - Running description: This is the record of the day's observations. The objective is to record accurately what you observe.
  - Forgotten episodes: These are accounts of previous episodes that you have forgotten but are remembering again while you are in the field.
  - Personal impressions and feelings: They may provide you clues to biases which might be clouding your observations.
  - Methodological notes: This refers to any ideas related to the techniques you used to conduct research, for example, any difficulties you have in collecting data, any biases that might be introduced by the data collection techniques or any changes in how you can make and record observations.

Advantages and Limitations of Ethnographic Design

Advantages

What is the justification for choosing ethnographic design over other designs? Wolcott (1999) lists the following advantages of conducting ethnography, most culled from his list of advantages:

- Ethnography can be conducted entirely by one individual.
- It is longitudinal in nature, allowing you to observe and record changes over time.
- It can be carried out almost at any place.
- It focuses on working with others rather than treating them as objects.
- It provides you with a detailed and rich database for further investigation and writing.
- You can make the research not only interesting but also adventurous.
- It requires no expensive or elaborate tools or equipment.
- It may present you with an opportunity to learn and use another language.
- It draws upon your personal skills and strengths to advantage.
- You often have exclusive domain or sole responsibility in the chosen setting or site.
- Your role is recognised.
- It offers you an opportunity to integrate professional and personal life.
- It allows you to get an insider's view of reality.
- It can provide deep insightful data.
- It can be used to study marginalised groups of people closed to other forms of research.
- It allows you to collect data in a realistic or naturalistic setting in which people act naturally, focusing on both verbal and non-verbal behaviours.

Ethnographic research is highly useful to single out new or currently unmet user needs. The approach is most valuable at the beginning of a project when the researcher does not have the slightest idea of real end user needs or to perceive the limitations of using a new product or service by people considered in the research study.

Ethnographic research is used extensively to study international comparative education scenario. In educational settings, the method is used to study particularly schooling. While particularly suited to exploratory research, ethnography draws on a wide range of both qualitative and quantitative methodologies, moving from 'learning' to 'testing'.

Limitations

Although there are many advantages associated with ethnographic research, unfortunately there are also many disadvantages.

- Time requirement: You cannot undertake a good ethnographic research without a huge investment in your time, because some studies can go on for years, and you will always be required as part and parcel of the culture or group for all that time.
- Presentation of results: Since the results of the ethnographic study are so diverse due to multiple perceptions with which you have to deal with, you may end up in serious difficulties in presenting your result. This problem is further aggravated by the length of time and large number of participants and thus collating all the data and results into a coherent presentation or paper.
- Reliability: In an ethnographic study, a researcher works alone and there is nobody who can check the findings for reliability.

Briefly speaking, ethnography is a social science research method. It portrays social groups and situations in their real-life situations. The research relies heavily on personal experience and possible participation and not just observation. Ethnographers often work in multidisciplinary teams. Their focal point includes: (a) intensive language and culture learning, (b) intensive study of a single field or domain and (c) a mixture of historical, observational and interview methods.

Grounded Theory Research Design

Grounded theory has been widely used in a wide range of research primarily in educational research. The theory is one of the forms of qualitative research designs. Glaser and Strauss first propounded the grounded theory in their 1967 book *The Discovery of Grounded Theory*. The theory is developed inductively from a corpus of data. The main focus is to invent theories with respect to social phenomena. In other words, the aim is to develop higher level understanding that is 'grounded' in or derived from a systematic analysis of data (Glaser and Strauss 1967). The theory consists of a set of steps whose careful execution is thought to 'guarantee' a good theory as the outcome.

With this approach, a researcher starts his study with a set of either qualitative or quantitative data. On the basis of this information, he then investigates and explores specific patterns and relationships among the data. Based on these patterns, he develops a theory that is 'grounded' in the data itself.

This method of research differs significantly from the traditional approach to science, which begins with a theory and seeks to test it through scientific method. Grounded theory method is basically inductive method or a form of inductive reasoning. It is both scientific and creative. The key objective of the grounded theory is to explain a 'basic social process', focusing on the delineation of the relationship between categories and emerging theory rather than purely describing categories.

Types of Grounded Theory Designs

For producing high-quality grounded theory research, it is essential for you to understand the grounded theory paradigm and the nature of the study. For this to achieve, Mills, Bonner and Francis (2006) suggest that researchers should select a research design paradigm that is parallel with their beliefs about the nature of the phenomenon of interest. Creswell (2012) categorises three dominant grounded designs—systematic design, emerging design and constructive design.

Systematic Design

This design of grounded theory is widely used in educational research. This design consists of three stages of coding: (a) open coding, (b) axial coding and (c) selective coding.

In open coding, the researcher constructs the initial categories of information about the studied subject by segmenting the collected data. For this, he identifies the important words and phrases, and labels them by using suitable term (Birks and Mills 2011). He then develops corresponding categories of the data that he has collected. However, the data should be coded in multiple possible ways by using memos to construct an emergent concept or theory during data analysis. This coding, in fact, is the first stage of forming emergent theory of conceptualisation.

In axial coding, the researcher chooses an open category and relates it to other categories on the basis of their causal conditions of the studied phenomenon, the strategies that used to resolve the studied phenomenon, environmental variables that influence the adoption of strategies and outcomes. This coding stage helps the researcher establish hypothetical relationships between the major categories and their corresponding sub-categories.

Finally, in selective coding, the researcher interprets the interrelationships that emerge among categories formed in axial coding. This coding category keeps only relevant variables to the core variables in order to produce an explicit theory (Glaser and Holton 2004).

Emerging Design

Emerging design is especially suitable to be applied to study 'real world' which seems relatively complex, poorly controlled and messy.

Constructive Design

A constructive design aims at explaining participants' meanings towards a process in reality. Researchers use constructive design because they have considerable experience in the studied phenomena and worked with the participants.

Key Features of Grounded Theory Research Design

Creswell (2012) has listed out six major characteristics of typical grounded theory research, which are widely utilised by grounded theorists. These characteristics are as follows:

- Process approach: The research process in grounded theory research is a series of interactions and outcomes among a group of people regarding the studied phenomenon. In educational research, some examples of the said phenomenon include AIDS prevention, the first-year teaching life of a new teacher or the leadership of a school principal.

- Theoretical sampling: This sampling technique refers to the ongoing process of coding the data, comparing the data and grouping similar data to build categories and core categories. The technique assist researcher systematically chooses the most important data for the studied phenomenon.
- Constant comparative: One of the fundamental features of grounded theory pertains to constant comparative. As implied by the name, constant comparison is the process of comparing like with like to trace out the emerging pattern and theory. All the collected data are compared constantly to find out their commonalities and variations. For instance, it involves comparing events to events, events to codes, codes to codes, codes to categories and categories and categories.
- Core category: The core category (or central category) portrays the main theme of a study. A core category can be viewed as the integration of other major derived categories into a theory that rooted in the collected data.
- Theory generation: The outcome of grounded theory research is to construct a theory that explains a studied phenomenon from the collected data (Creswell 2012). Since the generated theory is close to the data, it does not have an excellent ability for generalisation, and thus it could not be applied widely for many situations and people (Creswell 2012). He further points out that the resultant theory can be presented in three forms: (a) as a visual coding paradigm, (b) as a series or hypothesis or (c) as a narrative story.
- Memos: Grounded theorists create memos about the collected data. Memo writing is a good idea to record emergent concepts or ideas throughout the research process. These types of memos are known as theoretical memos. Documentation of these ideas and thoughts would prevent paralysis in the process of generating theories as memo writing is helpful to direct researchers into data and questions that need further exploration.

Data Collection and Data Analysis in Grounded Theory

The data collection stage is a crucial stage to obtain different kinds of sources as an endeavour to develop explicit theories. In general, data could be collected in forms of interviews, observations, focus group discussion and documents. Among these methods, interview, arguably, is the most frequently reported method.

Steps in Grounded Theory Design Research

Creswell (2012) has outlined the following eight major steps to conduct grounded theory research.

Step 1: Decide whether a grounded theory design suits the research problem.
Step 2: Plan a feasible process to study.
Step 3: Seek approval and access.
Step 4: Theoretical sampling.
Step 5: Code the data.
Step 6: Use selective coding and develop the theory.
Step 7: Validate the emerging theory.
Step 8: Write a grounded theory report.

For a detailed description of this, you are advised to refer to Creswell's 2012 book *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*.

When to Use the Grounded Theory Design

Grounded theory is an appropriate approach for studying social interactions or experiences. The theory is equally appropriate to explain a process but not for verifying an existing theory. Researchers approach the research question with academic interests and have sufficient background knowledge of relevant literature on the subject of research study. But they neither develop nor test hypotheses. Rather, they generate theory through a close and careful analysis of the data (Box 2.3).

Advantages and Limitations of Grounded Theory Design

Advantages

- It takes researchers' perceptions into account in the research process.
- It offers opportunities to the researchers to use their values and understanding in order to generate a new theory for a very complex phenomenon.
- It permits researchers to have a glance at the studied phenomena with new angles and constructs new perspectives without restriction. Thus, grounded theorists are able to understand the studied phenomena holistically.

Limitations

- The theory is complex and difficult to be followed by novice researchers.
- It is difficult to present the findings of the study in an effective way.
- It is not perfect for all research questions.

Historical Studies

Historical research is another type of research in which a researcher uses documents and remains containing and narrating historical facts to study events or ideas of the past. The researcher spends much of his time on studying intensively the philosophy of renowned persons and groups of people pertaining to his topic of research of distinct historical periods.

> 'Historical research "has been defined as the systematic and objective location, evaluation and synthesis of evidence in order to establish facts and draw conclusions about past events It is an act of reconstruction undertaken in a spirit of critical inquiry designed to achieve a faithful representation of a previous age"' (Borg 1963).
>
> 'It is a truthful integrated account of the relationships between persons, events, times and places' (Best and Kahn 2006).

Purpose of Historical Research

Historical research is undertaken to know and understand the past and then try to understand the present in relation to past events and developments. In an educational setting, historical research is highly useful to know how and why theories and practices have been developed which currently prevail in our schools. The main objectives of historical research in the field of education are:

- To study and learn from the past failure or success
- To understand fully the present education practices and policies
- To assess accurately the past and make predictions accordingly
- To study the evidences ascertained by the probability for obtaining a clear snapshot of the present
- To identify reasons to develop educational theories and practices

In the domain of education, historical research examines the importance and significance of education and its interrelationship with school and curriculum. This type of research is also undertaken to prevent 'the development of same ideas or theories' every few years. The areas of research here include primarily an individual, an idea, a movement or an institution. However, none of these are researched in isolation.

Indeed, the difficulty of obtaining adequate data makes historical research one of the most taxing kinds of inquiry to conduct satisfactorily. Another equally important difficulty encountered in doing historical research in education is the hypotheses formulation of historical investigations. They are usually implied rather than explicitly stated. The researcher attempts to obtain all sorts of evidences and carefully examines and evaluates it reliability and trustworthiness. If the evidence is in harmony with the consequences of the hypothesis, the hypothesis is confirmed, if it is not then the hypothesis is not confirmed. This is how historical generalisations are established.

The following examples suggest hypothesis formulation for a historical research:

Example

- The educational reforms In India since 1947 were based upon practices that previously have been tried and discarded.
- The evolution of teacher training programmes since the Second World War explains the historical origins of the content and processes of current programmes.

When conducting a historical research, you must be a good judge of sources. You have to keep in mind the following four key steps.

Step 1: Collect as many sources as you can

For data collection, you should look at all primary sources (those written at the time of the event) and secondary sources (those written a reasonable time after the event). These sources may be websites, books, newspaper articles, letters and so on.

Step 2: Assess the reliability of your source

You should always assess and ensure the authenticity of data by asking questions like: 'who is the original author or who wrote the source?' 'In which year it was published?' 'When was it last edited/updated?' and so on. You may notice a bias in primary source of data especially for historical events such as wars and revolutions.

Step 3: Assess the usefulness of your source

A source is good only if it provides the researcher correct and reliable data. It is important to remember that even if a source is not reliable, it could still be useful because a biased source tells us opinions of certain people about the key features of an historical event.

Step 4: Gather notes

Notes should be brief and succinct, and written in your own words.

Advantages and Limitations of Historical Research

Advantages

- Better understanding of many current educational practices, theories and issues within the context of experiences of the past
- Researcher can rationally and objectively analyse events of the past

Limitations

- No scope to control and manipulate any variable
- No way the researcher can affect events of the past
- Researcher cannot collect data by administering instruments of data collection
- The research excessively relies on secondary data

Most historical researches are qualitative in nature because the proper subject matter of historical research consists largely verbal and other illustrative material resulting primarily from a society or a culture's past. The analytical tools that a researcher must possess include collecting, classifying, ordering, synthesising, evaluating and interpreting the information and, certainly, his or her sound personal judgement.

Brickman (1982) provides a number of possible topics of historical research in education and an example for each. Table 2.3 repeats his list.

Obviously, these topics are too broad for a student preparing a master's dissertation. In some cases, topics such as these would probably take most of his career. The processes of delimitation and hypothesis formation are needed to make these topics useful.

Phenomenological Research Design

Phenomenological research can be a bit confusing because (a) phenomenology is a school of philosophical thought that underpins all of qualitative research; (b) phenomenological research is a distinct qualitative method for discovering the underlying structure of shared essences of some social phenomenon; (c) phenomenological research might not be adequately addressed in many general research methods text books and (d) very few qualitative methodologists include a basic or generic category of research; the result is that researchers, faculty and learners tend to incorrectly label any form of qualitative research that is not ethnographic, grounded study or case study as phenomenological research. As a result, much of what is not phenomenological research gets labelled as phenomenological research and the confusion regarding phenomenological research is exacerbated.

Phenomenology is a school of thought that emphasises people's subjective experiences and interpretations of the world. A phenomenologist always wants to understand how the world appears to others.

Phenomenology literally means the study of phenomena. Researchers use it to describe phenomena (events, situations, experiences or concepts) that exist as part of the world in which we live. We all know that we are encompassed by several phenomena, aware of them but we do not understand them completely. This may be due to the fact that these phenomena may exist because the phenomenon has not been overtly described and explained or our understanding of the impact it makes may be unclear. For example, we know that lot of people are thinkers. But what does 'thinking' actually mean and what is it like to be a thinker?

TABLE 2.3 Examples of Topics for Education Historical Study

Type	*Topic of Study*
PERIOD	'Education During the First Half of the Fifteenth Century'
GEOGRAPHICAL LOCATION	'German Education under Frederick the Great'
EDUCATIONAL LEVEL	'The Secondary Schools of Ancient Rome'
INSTITUTION	'Amherst College in the Nineteenth Century'
BIOGRAPHY	'Bronson Alcott as an Educator' (Biographical detail, as such, is of less importance for term report purposes than an exposition of the man's educational ideas, work and influence)
INNOVATIONS	'Three Decades of Audiovisual Education'
PHILOSOPHY	'Changing Concepts of American Higher Education in the Nineteenth Century'
METHODOLOGY	'Herbartianism in American Educational Practice'
CURRICULUM	'The Subject of Rhetoric in Ancient Greece'
PERSONNEL	'The Role of the Teacher During the Renaissance'
CHILDREN	'Changing Attitudes Towards Corporal Punishment of Children in the United States'
LEGISTATION	'Compulsory School Attendance Laws in Prussia During the Eighteenth Century'
MATERIALS	'The Evolution of American School Readers, 1700–1830'
NON-SCHOOL AGENCIES	'The Development of the Library in the Nineteenth-century America'
ORGANISATIONS	'History of the Public School Society of New York'
FINANCE	'Methods of School Taxation In Pennsylvania—1820–1880'
ARCHITECTURE	'The Evolution of School Building in Illinois'
ADMINISTRATION	'The Rise of the State Superintendency of School'
LITERATURE	'A Century of Educational Periodicals in the United States'
INFLUENCE	'The Influence of Rousseau upon Pestalozzi'
REPUTATION	'The Reception of Horace Mann's Educational Ideas in Latin America'
COMPARISON	'A Comparative Study of Renaissance Theories of the Education of the Prince'
TEXTBOOK ANALYSIS	'A Study of the Treatment of Primitive Education in Textbooks in Educational History'

Source: Brickman (1982).

According to Grossmann and Armstrong (Hochberg 2010), 'phenomenology research methods and studies analyse experiences, events and occurrences with disregard or minimum regard for the external and physical reality'. Accordingly, the method helps researchers generate ideas using the large amount of data by way of induction and human interests. The method also takes into account the stakeholder's perspective and their reflection on the study. Good examples of this method include studies that attempt to assess the impact of leadership style on employee motivation in an organisation through conducting in-depth interviews with employees.

Another example could be back pain. Correlation studies may inform us about the types of people who experience back pain and the apparent causes associated with this pain. But how people actually live with back pain? What are its effects on peoples' lives? What problems does it cause? A phenomenological study might help the researcher to explore, for example, the effect that back pain has on sufferers' relationships with other people by describing its effect on children of having a disabled parent.

Phenomenological research acknowledges that there is a gap in our understanding and its clarification or illumination is extremely beneficial. This research design not only provides definitive explanations but it also uplifts awareness and increases insight.

A phenomenological study explicates the meaning, structure and essence of the lived experiences of a person or group of people around a specific phenomenon. The phenomenologist attempts to understand human behaviour through the eyes of the participants in the study. A phenomenologist believes that all perceptions and constructions are ultimately grounded in a particular perspective in time and space. Phenomenology does not begin with a theory but instead begins with phenomenon under consideration (Christensen, Johnson and Turner 2010).

'The objective of phenomenology is the direct investigation and description of phenomena as consciously experienced, without theories about their causal explanations or their objective reality. It therefore seeks to understand how people construct meaning' (Creswell 2012).

Michael Patton (1990) has a clear detailed explanation of the aim of phenomenological research since he bases this method of research on:

> the assumption that *there is an essence or essences to shared experience*. These essences are the core meanings mutually understood through a phenomenon commonly experienced. The experiences of different people are bracketed, analysed, and compared to the identity of the essences of the phenomenon, for example, the essences of loneliness, the essence of being a mother, the essence of being a participant in a particular program. *The assumption of essence, like the ethnographer's assumption that culture exists and is important, becomes the defining characteristic of a purely phenomenological study*.

In general, 'a phenomenological research is well suited for studying affective, emotional, and often intense human experiences' (Merriam 2009).

The methodology that the researcher will choose will depend on the nature of the research question and the specific research design. A phenomenological study should have a strong central phenomenological question. You are required to note the 'affective, emotional and intense human experience' conveyed in each example phenomenological research question below and how the phenomenon is clearly identified.

Here are examples of phenomenological research questions:

- What is the experience of motherhood for female soldiers deployed to Afghanistan who have children between the ages of 1 and 3 at home?
  The phenomenon in the question above is motherhood.

- What is self-forgiveness for convicted murders?
  The phenomenon in the question above is self-forgiveness.
- How do high school teachers use intuition in making classroom management decisions during high-risk incidents?
  The phenomenon in the question above is the use of intuition.
- How do female high school teachers who have been physically assaulted by students overcome their fears so they can effectively teach?
  The phenomenon in the question above is the recovery process.
- What role does spirituality play in the remission of cancer in patients?
  The phenomenon in the question above is spirituality.

From these questions, you can see how the central research question identifies phenomenon being examined. Phenomenology attempts to get below the surface of simply perceptions to discover and identify how the phenomenon was experienced and the shared essence of that experience. If a learner or mentee cannot readily identify the phenomenon that will be studied, then a phenomenological design is most likely not the appropriate design.

Main Characteristics of Phenomenology

There are several main characteristics that help to define what exactly phenomenology is.

Methodology

A phenomenological study often involves the four steps of the following:

- Bracketing: method of demonstrating the validity after initiating a phenomenological study.
- Intuiting: ability to acquire knowledge without proof, evidence or conscious reasoning or without understanding how the knowledge was acquired.
- Analysing: daily human behaviour providing one with a greater understanding of nature.
- Describing: personal experiences to provide the researcher with clues for orienting to the phenomenon and thus to all the other stages of phenomenological research.

Sampling

- Small samples of no more than 10 participants are most suitable for this type of research.
- Large samples can become unwieldy.
- Samples or participants in phenomenological research are generally chosen according to what is known as ‘purposive sampling’.

Data Collection Methods

Any way the participants can describe their lived phenomenal experience that can be used to gather data in a phenomenological study. They can use an interview to gather the participants’ descriptions of their experience, or the participants’ written or oral self-report or even their aesthetic expressions (e.g., art, narratives or poetry).

The data collection tools that are most often used are:

- Interviews/speech
- Diaries/written
- Drawings/non-verbal
- Observation/visual

To sum up, the goal of qualitative phenomenological research is to describe a 'lived experience' of a phenomenon. As this is a qualitative analysis of narrative data, methods to analyse its data must be quite different from more traditional or quantitative methods of research.

Summary

Qualitative research designs are best-suited research designs when our aim is to study and discover the underlying motives of human behaviour. Researchers can use these methods for seeking a comprehensive understanding of diverse factors which drive people to behave in a particular manner or which make people like or dislike a particular thing. It is, however, important to state that the application of qualitative research in real-life situations is not as simple as it appears to be because such research demands special guidance from experts. Qualitative research represents a broad framework for conducting educational studies. The main focus of qualitative studies in education is on holistic descriptions of learners and teachers in naturalistic settings.

The domain of qualitative research is to search the 'why' and not the 'how' of the research topic. For doing this, the method makes an extensive use and analysis of unstructured information. The most commonly used instruments include things such as interview transcripts, open-ended survey responses, e-mails, notes, feedback forms, photos and videos. It does not just rely on statistics or numbers, which are the domain of quantitative researchers.

Qualitative research is highly appropriate where the aim of our research is to acquire a clear understanding of people's attitudes, behaviours, value systems, concerns, motivations, aspirations, culture or lifestyles. The approach is used in making informed decisions, policy formation, communication and behavioural research. Focus groups, in-depth interviews, content analysis, ethnography and evaluation are among the many formal approaches that are used in qualitative research.

Chapter 3 describes research designs commonly used in quantitative research studies.

Self-test Exercise

Exercise 2.1:

2.1.1. How would you describe the steps of the research process and key components of designing a study?

2.1.2. What type of strategy in action research would you like to recommend for:

- Professionalising teaching
- Enhancing the motivation and efficacy of a weary staff of your school

- Meeting the needs of an increasing diverse student body
- Achieving success with 'standard-based' reforms

2.1.3. What measures would you suggest for individual teachers conducting action research to make continuous progress in developing their strength as reflective practitioners?

2.1.4. Ms Francoise, Claire's Grade 3 teacher, called Claire's parents, Mrs Daniele and Mr Khan, and asked them to come in for a meeting. The teacher was finding that Claire was increasingly difficult to manage both on the playground and in class. At the meeting, Ms Francoise described her concerns about Claire's behaviour in school, and the parents indicated that Claire had similar difficulties at previous schools.
How would you organise and structure a case study to address this particular problem? In addition to assessing Claire's achievement and ability:

2.1.4.1. Prior to the trial of treatment, list the major observations you would make for the Claire's class.

2.1.4.2. Which theories would you use to collect and analyse the data? Why?
- Case study
- Grounded theory
- Phenomenological analysis

2.1.5. How would you recognise that the processes and outcomes of research should be disseminated to appropriate audiences of professionals, policymakers and consumers in order to advance knowledge?

References

Ackoff, R. 1953. *The Design of Social Research*. Chicago: University of Chicago Press.

Adelman, C.D. and S. Kemmis. 1980. 'Rethinking Case Study: Notes from the Second Cambridge Conference'. In *Towards a Science of the Singular*, edited by H. Simons. Norwich: Centre for Applied Research in Education, University of East Anglia.

Alon, U. 2009. 'How to Choose a Good Scientific Problem'. *Molecular Cell,* 35 (6): 726–728.

Best, J. and J. Kahn. 2006. *Research in Education*. Boston, MA: Pearson Education, Inc.

Birks, M., and J. Mills. 2011. *Grounded Theory: A Practical Guide*. London: SAGE Publications.

Bogdan, R.C. and S.K. Biklen. 1998. *Qualitative Research for Education: An Introduction to Theory and Methods*, 3rd ed. Boston, MA: Allyn & Bacon.

Borg, W.R. 1963. *Educational Research: An Introduction*. New York, NY: David McKay Co.

Brickman, W.W. 1982. *Educational Historiography: Tradition, Theory, and Technique*. Cherry Hill, NJ: Emeritus.

Burns, N. and Grove, S.K. 2003. *Understanding nursing research*, 3rd edition. Philadelphia: Saunders Company. Available at http://uir.unisa.ac.za/bitstream/handle/10500/1796/04chapter3.pdf (accessed on 14 December 2017).

Carr, W. and S. Kemmis. 1983. *Becoming Critical: Knowing Through Action Research*. Geelong: Deakin Press.

———. 1986. *Becoming Critical: Education, Knowledge and Action Research*. London: The Falmer Press.

Christensen, L.B., R.B. Johnson and L.A. Turner. 2010. *Research Methods Design and Analysis*, 11th ed. Boston, MA: Allyn & Bacon.

Corey, S.M. 1953. *Action Research to Improve School Practices*. New York, NY: Teachers College Press.

Creswell, J.W. 2012. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*, 4th ed. Upper Saddle River, NJ: Pearson.

Davies, B. 2013. *Research Design.* Retrieved on 23 January 2018 from: http://www.antarcticglaciers.org/students-3/postgraduate-students/research-design/

Dobbert, M.L. 1982. *Ethnographic Research: Theory and Application for Modern Schools and Societies*. New York, NY: Praeger.

Eisner, E.W. 1991. *The Enlightened Eye: Qualitative Inquiry and the Enhancement of Educational Practice*. New York, NY: Macmillan Publishing Co.

Erickson, F. 1986. 'Qualitative Methods in Research on Teaching'. In *Handbook of Research on Teaching*, 3rd ed., edited by M.C. Wittrock. New York, NY: Macmillan.

Gilmore, T., J. Krantz and R. Ramirez. 1986. 'Action Based Modes of Inquiry and the Host-Researcher Relationship'. *Consultation* 5 (3): 160–176.

Glaser, B.G. and J. Holton. 2004. 'Remodelling Grounded Theory'. *Forum Qualitative Social Research* 5 (2), Article 4. Available at http://nbn-resolving.de/urn:nbn: de:0114-fqs040245 (accessed on 5 June 2017).

Glaser, B. and A. Strauss. 1967. *The Discovery of Grounded Theory*. Chicago, IL: Aldine.

Green, J., A. Skukauskaite and W.D. Baker. 2012. 'Ethnography as Epistemology: An Introduction to Educational Ethnography'. In *Research Methods and Methodologies in Education*, 309–321, edited by J. Arthur, M. Waring, R, Coe, L.V. Hedges. Los Angeles: SAGE Publications.

Hammersley, M. and R. Gomm. 2000. 'Bias in Social Research'. In *Taking Sides in Social Research: Essays in Partisanship and Bias,* Chapter 6, Hammersley, M. (Ed.). London: Routledge.

Harris, M. and O. Johnson. 2000. *Cultural Anthropology*, 5th ed. Needham Heights, MA: Allyn & Bacon.

Hochberg, H. 2010. 'Phenomenological Realism Versus Scientific Realism: Reinhardt Grossmann—David M. Armstrong Metaphysical Correspondence'. *Dialectica* 64 (3): 447–451.

Holly, P. and D. Whitehead. 1986. *Collaborative Action Research*. Cambridge: Cambridge Institute of Education.

Joyce, B.R. 1991. 'Doors to School Improvement'. *Educational Leadership* 48 (8): 59–62.

Kemmis, S., Mctaggart, R., and Nixon, R. 2014. *The Action Research Planner: Doing Critical Participatory Action Research.* Singapore: Springer.

Kolb, D. 1984. *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs, NJ: Prentice Hall.

LeCompte, M.D. and J. Preissle. 1993. *Ethnography and Qualitative Design in Educational Research*, 2nd ed. San Diego, CA: Academic Press.

Lincoln, Y.S. and E.G. Guba. 1985. *Naturalistic Inquiry*. Beverly Hills, CA: SAGE Publications.

Magoon, A.J. 1977. 'Constructivist Approaches in Educational Research'. *Review of Educational Research* 47 (4): 651–693.

Merriam, S.B. 2009. *Qualitative Research: A Guide to Design and Implementation*. San Francisco, CA: Jossey-Bass.

Mills, J., A. Bonner and K. Francis. 2006. 'The Development of Constructivist Grounded Theory'. *International Journal of Qualitative Methods* 5 (1): 1–10.

Nisbet, J. and J. Watt. 1984. 'Case Study', Chapter 5. In *Conducting Small-Scale Investigations in Educational Management*, edited by Judith Bell, Tony Bush, Alan Fox, Jane Goodey, Sandy Goulding. London: Harper and Row.

Patton, M.Q. 1990. *Qualitative Evaluation Methods*, 2nd ed. Thousand Oaks, CA: SAGE Publications.

Peine, M.E. 2003. 'Doing Grounded Theory Research with Gifted Students'. *Journal for the Education of the Gifted* 26 (3): 184–200.

Popkewitz, T. 1984. *Paradigm and Ideology in Educational Research*. London: Falmer Press.

Postlethwaite, N. 2005. *Educational Research: Some Basic Concepts and Terminology Module 1*. Paris: UNESCO International Institute for Educational Planning.

Ronald C. Martella, Ronald Nelson and Nancy E. Marchand-Martella. (1998). *Research Methods: Learning to Become a Critical Research Consumer.* Boston: Allyn and Bacon.

Robson, C. 2002. *Real World Research: A Resource for Social Scientists and Practitioner-researchers*.

Schaefer, R.J. 1967. *The School as a Center of Inquiry*. New York, NY: Harper and Row.

Schön, Donald A. 1983. *The Reflective Practitioner: How Professionals Think in Action*. New York, NY: Basic Books.

Singleton, R.A. and B.C. Straits. 2005. *Approaches to Social Research*, 4th ed. New York, NY: Oxford University Press.

Sirotnik, K.A. 1987. 'Evaluation in the Ecology of Schooling'. In *The Ecology of School Renewal: The Eighty-Sixth Yearbook of the National Society for the Study of Education*, edited by J.I. Goodlad, 41–62. Chicago, IL: The University of Chicago Press.

Susman, G.I. 1983. *Action Research: A Sociotechnical Systems Perspective*, edited by G. Morgan. London: SAGE Publications.

Taft, R. 1997. 'Ethnographic Research Methods'. In *Educational Research, Methodology, and Measurement: An International Handbook*, 2nd ed., edited by J.P. Keeves. Oxford: Pergamon/Elsevier Science Inc.

Trochim, W.M.K. 2005. *Research Methods: The Concise Knowledge Base*. Cincinnati, OH: Atomic Dog Publishing.

UNECO-PROAP. 2011. *Sustainability of Community Learning Centres: Community Ownership and Support.* Bangkok.

Van Maanen, J. 1996. 'Ethnography'. In *The Social Science Encyclopedia*, 2nd ed., edited by A. Kuper and J. Kuper, 263–265. London: Routledge.

Walford, G. 2008. *How to do Educational Ethnography*.

Walker, J.C. and C.W. Evers. 1997. 'Research in Education. Epistemological Issues'. In *Educational Research, Methodology, and Measurement: An International Handbook*, 2nd ed., edited by J.P. Keeves. 22–31. Oxford: Pergamon/Elsevier Science Inc.

Wallace, M.J. 1998. *Action Research for Language Teachers*. Cambridge: Cambridge University Press.

Wolcott, H.F. 1988. 'Ethnographic Research in Education'. In *Complementary Methods for Research in Education*, edited by R.M. Jaeger. 26–35. Washington, DC: American Educational Research Association.

———. 1999. *Ethnography: A Way of Seeing*. Walnut Creek, CA: Altamira Press.

3 CHAPTER

Quantitative Research Designs

Introduction

You have learnt in Chapter 2 how and when educational researchers utilise qualitative research designs to investigate and examine the effects of major changes on several aspects of the education system and then suggest policy advice that will help integrate and enlarge the post-productive courses of action. These investigations are undertaken to seize and terminate existing practices that have proven to be damaging and wasteful. Another key research design widely used in educational research to investigate such changes and their impact on several aspects of the education process is quantitative design.

We have also learnt in Chapter 2 that a research design serves many purposes throughout the design process. It helps us identify and prove or disprove our assumptions, find commonalities across our target audience members and recognise their needs, goals and mental models. Overall, research informs our work, improves our understanding and validates our decisions.

In this chapter, like Chapter 2, we will look at the many elements of those research designs researchers commonly use in quantitative research studies. The purpose here is to suggest a head start on how to use these designs in your research studies and improve your experiences in conducting the research.

What is a Quantitative Research Design?

Most scientific disciplines use quantitative research designs as the standard experimental designs in research. Sometimes referred to as true science, we use them as traditional mathematical and statistical means to measure results conclusively.

> 'Quantitative methods emphasize objective measurements and the statistical, mathematical, or numerical analysis of data collected through polls, questionnaires, and surveys, or by manipulating pre-existing statistical data using computational techniques. Quantitative research focuses on gathering numerical data and generalizing it across groups of people or to explain a particular phenomenon' (Babbie 2010).

The emphasis of these designs is on objective measurements and on the statistical, mathematical or numerical analysis of data collected by means of polls, questionnaires and surveys or by manipulating pre-existing statistical data using computational techniques. The design relies primarily on gathering numerical data and generalising it across groups of people or to explain a particular phenomenon.

Quantitative research designs establish the relationship between an independent variable and a dependent or outcome variable within a population. They are either 'descriptive' (subjects usually measured once) or 'experimental' (subjects measured before and after a treatment). A descriptive study determines only 'associations' between variables; an experimental study establishes 'causality'.

'Numbers', 'logic' and an 'objective' stance are the three vital areas that quantitative research deals and claims to address. The prime focus is on numeric and unchanging data and detailed, convergent reasoning rather than divergent reasoning. Consequently, these designs are most commonly used by physical scientists, although social sciences, education and economics have been known to use this type of research. It is the opposite of qualitative research.

Quantitative designs generate a hypothesis to be proved or disproved. This hypothesis is provable by mathematical and statistical means and is the basis around which the whole experiment is designed.

It is also important that in quantitative research designs, randomisation of any study group is essential and it must include a control group, wherever possible. A sound quantitative design should only manipulate one variable at a time; else, statistical analysis becomes tedious and complex and thus open to question.

Ideally, while conducting your research using quantitative research designs, you should make sure that your research is constructed in a manner that allows others to repeat the experiment and obtain similar results. Box 3.1 shows an example of a quantitative research design and its essential elements.

BOX 3.1: 'Impact of Training and Professional Development of Primary School Teachers on Learning Outcomes'

Background of the Study

There are a number of factors that contribute to the success of any school. These factors include capital, equipment, manpower (teaching and non-teaching staff) and so on. All these factors are important but the most significant factor is the teachers. Since it is the teacher who will put the other resources to work, it should be viewed as such by educational management by giving him or her due attention in order to achieve its organisational goals and objectives.

The main objective of schooling is to provide quality education and to achieve this goal; adequate teacher training and teachers' development programmes should be put in place to enhance teachers' performance.

(Continued)

(Continued)

Statement of the Problem

Teachers are considered the most critical to any school survival of a truism that adequate supply of material and financial resources utilise these available resources to bring about the desired goals.

It is the opinion of educational planners and policymakers that the poor performance of schools and staff follows from their inability to keep abreast with the new technological current as a result of the absence of appropriate and sufficient staff training. It is against this backdrop that this quantitative research considers the impact of teacher training and teachers' development programmes on learning outcomes of primary school pupils.

Research Questions

1. Are the teachers of the school satisfied with government policy on quality of the selection and interview procedure?
2. Are the teachers satisfied with the placement and promotion procedure?
3. Are there training programmes for the teachers?
4. How adequate in terms of content and relevance are these training programmes relevant?
5. Has the promotion process in the school improved teachers' performance?
6. Do the teachers of the school utilise skills that they learnt after their training?

Objective of the Study

- To conduct an empirical investigation through a review of the teachers' training provisions and the development policy of the government
- To access the teachers' recruitment, selection and training programmes and, from it, establish some relationship between these programmes and problems enumerated
- To highlight the need for teachers' training and development at the primary level of education.
- To identify some techniques of teachers' training and development in the ministry and their relevance to the learners' learning needs and the ministry industry at large
- To establish the relationship between personnel training and development, and staff performance
- To correct the belief that in this age of computerisation and technological development, all that schools need to survive is the acquisition of up-to-date capital equipment compared to the power management of its teaching manpower
- To suggest solutions to the identified problems

Research Hypotheses

The hypotheses of this research study are as follows:

HO: There is no direct relationship between teacher training and pupils' learning outcomes.
H1: There is direct relationship between teacher training and pupils' learning outcomes.
HO: Lack of adequate teacher training and development is not directly responsible for high teacher turnover.
H2: Lack of adequate teacher training and development is directly responsible for high teacher turnover.
HO: Training does not improve teachers' and school's productivity and effectiveness.
H3: Training improves the productivity and effectiveness of teachers and the school.

Research Design

This research study is based mainly on the effect of teachers' training and development in human resources management with particular reference to primary level of education.

However, for the purpose of complete analysis, references will be made to other neighbouring countries' teachers training systems only when and where substantial evidence relating to the research study is offered.

The period covered in the study will extend from 2010 to 2013. The research study which is designed to be current has both descriptive and empirical values. The descriptive aspect of it reinforces the behavioural pattern of teachers. The behavioural pattern also includes the contribution from the trained teachers to the education system's goals and the general impact of the training received in the organisation.

The empirical analysis entails the various analyses of the data collected for the purpose of testing and consequently accepting or rejecting the stated hypotheses.

The hypotheses of the study are stated in both null and alternate styles. The null hypotheses (HO) assume that there is no significant difference between the observed frequencies and the expected frequencies, while the alternate hypotheses assume that there is a significance difference. The acceptance of the HO automatically means the rejection of the alternate hypotheses (H1, H2, and H3) and vice versa.

In conclusion, the study not only provides the basic information on the magnitude of training the personnel received but it also focuses on the various problems of training and development.

Source of Data

The data collected for the purpose of analysis and consequently testing of the formulated hypotheses will be secondary data. The primary data will be collected using oral personal interviews and simple questionnaires administered to the primary school teachers.

Sampling Techniques

The sampling technique employed in the data collection will be mainly random sampling. This will be adopted with the view to reducing the degree of bias and sidedness of the respondent's opinion on the topic during the personal oral interview and the distribution of the questionnaires.

Second, to ensure that the views of all types of primary school teachers—senior and junior teachers, male and female teachers, teachers in rural and urban primary schools—are well and adequately represented, cluster sampling techniques will be used to complement the random sampling.

Method of Data Analysis

The method of data analysis adopted in this study will include the sample percentage and tabular presentation. This is because various alternative explanatory variables that are not easily quantified were used. To reduce the problem of the quantification of the variables, attitudinal type measurement using Likert scale will be used for the opinions supplied by the respondents.

To further test and accept or reject the formulated hypotheses, the chi-square distribution will be employed. In this test, if the calculated values of the chi-square is greater than the table values at a given level of significance, the HO would be rejected and automatically the alternative hypotheses accepted and vice versa.

To ensure a high level of confidence, test significance level of 0.0 and 0.05 will be used.

A simple regression analysis will be portrayed to determine the relationship between the amount spent on training and development using the number of primary graduates as the index of schools' effectiveness

(Continued)

(Continued)

and growth. In other words, the number of primary graduates being produced annually by a given primary school will be used to determine the contribution of teachers' training.

Questionnaire

This questionnaire is suggestive and is made up of two sections: Section A and Section B. Section A is the personal data relating to the respondents, while Section B is the main body relating to the major variables in the research topic. You are required to tick 'X' in the appropriate box provided and write briefly your opinion(s) in the space provided.

SECTION A: PERSONAL DATA

1. Sex of respondents:
   Male ☐
   Female ☐
2. Age:
   21–30 years ☐
   31–40 years ☐
   41–50 years ☐
   51 and above ☐
3. Marital Status:
   Single ☐
   Married ☐
   Divorced ☐
   Widowed ☐
4. Length of service with organisation:
   1–5 years ☐
   6–10 years ☐
   11–20 years ☐
   21–25 years ☐
   30 and above ☐
5. Educational background:
   High School ☐
   Senior Secondary ☐
   Bachelor Degree ☐
   Master's Degree ☐
   Doctoral Degree ☐
6. Professional qualification:
   Teacher Training Certificate ☐
   Teacher Training Diploma ☐
   Bachelor in Education ☐
7. Current position in the organisation:
   Junior Teacher ☐
   Senior Teacher ☐
   Part-time Teacher ☐
   Full-time Teacher ☐

SECTION B: TRAINING AND DEVELOPMENT PROGRAMME

8. What are some of the services provided by your school?

____

____

9. Does the school management encourage any self-development programmes?
   Yes ☐
   No ☐
10. What form of development does the school management recommend for its teachers?
   On-the-job training ☐
   Seminars ☐
   Simulations ☐
   Conferences ☐
   Induction Courses ☐
11. Does the school inspectorate review the training programmes?
   Yes ☐
   No ☐
12. Have you attended any training/development programme since the assumption of duty in this school?
   Yes ☐
   No ☐
13. Has the training development programme attended by you affected your performance in your job?
   Yes ☐
   No ☐
   If yes, state briefly how the training programme has affected your performance.

____

____

14. What benefits do you think the school management has derived from any development programme sponsored?
   Retention ☐
   Leadership sustenance ☐
   Increased number of graduates arising from you being efficient ☐
   Quick and efficient response to the learning needs of children ☐
   Rapid expansion ☐
   Low teacher turnover ☐
15. How is the selection and recruitment being determined in the school? Please specify.
   A: ________________________________
   B: ________________________________
   C: ________________________________
16. Has training influenced your desire to stay in the school?
   Yes ☐
   No ☐
17. Are there any social and personal benefits resulting from training?
   Yes ☐
   No ☐
   If yes, briefly state the social and personal benefits.

____

____

(Continued)

(Continued)

18. In your opinion, which of these element(s) is/are lacking in the organisation?
    Good working condition ☐
    Better salaries/wages ☐
    Opportunity for promotion ☐
    Job recognition and achievement ☐
    Opportunity for further training ☐
    None of the above ☐
19. Are you encouraged to develop by yourself?
    Yes ☐
    No ☐
    State why and how, if yes or no.
    ____
    ____
    ____
    ____

Source: Adapted from OECD (2016).

All the elements of the above quantitative research design are described and discussed in detail below.

Features of Quantitative Research Design

The main characteristics of quantitative research designs are as follows:

- The data are usually collected using structured research instruments.
- Samples are large and thus outcomes are representative of the population.
- Given its high reliability, research studies can easily be replicated or repeated.
- Research questions are clearly and precisely formulated and defined to which objective answers are sought.
- Prior to the collection of data, all aspects of the study are carefully designed.
- Data are in the form of numbers and statistics, often arranged in tables, charts, figures or other non-textual forms.
- Project can be used to generalise concepts more widely, predict future results or investigate causal relationships.
- The researcher uses tools, such as questionnaires to collect data and computer software, to analyse it.

The overarching aim of a quantitative research study is to classify features, count them and construct statistical models in an attempt to explain what is observed.

Types of Research Questions in Quantitative Research Designs

A quantitative research design should attempt to answer at least one quantitative research question. In some research studies, these quantitative questions are followed by either research hypotheses or HO. However, in this chapter, our sole focus is on quantitative research designs and hence research questions. Furthermore, since there is more than one type of quantitative question, that you can attempt to answer in your dissertation (descriptive research question, comparative research questions and relationship-based research question), we discuss each of these below.

Descriptive research questions: their aim is simply to describe the variable you are measuring. These questions start basically with words such as 'how much?' 'how often?', 'what percentage?' and 'what proportion?', but sometimes these also start with words such as 'what is?' and 'what are?'

Example

Question: **How many calories do Indians consume per day?**
Variable: Daily calorific intake
Group: Indians
Question: **What percentage of American men and women exceed their daily calorific allowance?**
Variable: Daily calorific intake
Group: 1. American men
2. American women
Question: **What proportion of Oxford university male and female students use the top 5 social network?**
Variable: Use of top 5 social networks (Facebook, Myspace, Twitter, LinkedIn and Classmates)
Group: 1. Male, Oxford university students
2. Female, Oxford university students
Question: **What are the most important factors that influence the career choices of Canadian university students?** (e.g., salary and benefits, career prospects, physical working condition and so on. We ranked each of them on a scale 1 to 10, 1 = least important and 10 = most important and use frequencies and measures of central tendency and measures of dispersion)
Variable: Factors influencing career choices
Group: Canadian university students

Comparative research questions: These questions examine the differences between two or more groups or one or more dependent variables. These questions start by asking 'what is the difference in?' a dependent variable between two or more groups.

Example

Question: **What are the differences in perceptions towards Internet banking security between adolescents and pensioners?**
Dependent variable: Perceptions towards Internet banking security
Groups: 1. Adolescents
2. Pensioners

Relationship research questions: These questions simply highlight associations or trends between variable and not causal relationship which is best explained by experimental designs discussed in the next tag.

Example

Question: **What is the relationship amongst career prospects, salary and benefits, and physical working conditions on job satisfaction between managers and non-managers?**

Dependent variable:	Job satisfaction
Independent variable:	1. Career prospects 2. Salary and benefits 3. Physical working conditions
Group:	1. Managers 2. Non-managers

Basic Methodology for a Quantitative Research Design

Basically, your overall structure in a quantitative design should be based on a scientific method and your research design must use deductive reasoning, where you need to form an hypothesis, collect data about the research problem and use the data from your investigation. This should be followed by drawing the conclusions of your research study with a view to reject or accept the hypotheses. The quantitative research design should follow, as much as possible, the underneath procedural steps.

Step 1: Make your observations about something that is unknown, unexplained or new. Investigate and review current theory surrounding your problem or issue.
Step 2: Hypothesise an explanation for those observations.
Step 3: Build a prediction of outcomes based on your hypotheses. Design a plan to test your prediction.
Step 4: Collect and process your data. If your prediction is correct, go to step 5. If not, the hypothesis has been proven false. Return to step 2 to form a new hypothesis based on your new knowledge.
Step 5: Verify your findings. Make your final conclusions. Present your findings in an appropriate form for your audience.

The quantitative research design that we select subsequently determines whether we look for relationships, associations, trends or interactions.

When Can You Use the Quantitative Research Design?

When do we use quantitative methods? To be realistic and pragmatic, the main question that we need to answer is: what kinds of questions are best answered by using quantitative research? Box 3.1 gives a comprehensive example of conducting systematically a research study using quantitative research methods.

If you take a pragmatic approach to research methods, first of all you need to find out what kinds of questions are best answered using quantitative as opposed to qualitative designs. There are four main types of research questions that quantitative research designs are particularly suited to find an answer to.

In an educational setting, Postlethwaite (2005) suggests four main types of research questions that quantitative research is particularly suited to finding an answer:

1. In the first category of questions, he includes only those questions demanding quantitative answers. Examples are: ‘How many students choose to study education?’ or ‘How many maths

TABLE 3.1 Scope of Quantitative Research Design

	Quantitative research designs
General Framework	• Seek to confirm hypothesis about phenomena • Instruments use more rigid style of eliciting and categorising responses to questions • Use highly structured methods such as questionnaires, surveys and structured observation
Analytical Objectives	• To quantify variation • To predict causal relationships • To describe characteristics of a population
Question Format	• Close-ended
Data Format	• Numerical (obtained by assigning numerical values to responses)
Flexibility in Study Design	• Study design is stable from beginning to end • Participants' responses do not influence or determine how and which questions researchers ask next • Study design is subject to statistical assumptions and conditions

teachers do we need and how many are there in our school district?' To answer these kinds of questions, obviously the researcher is required to use quantitative research. Qualitative research designs cannot provide the researcher with the (numerical) answer he is looking for.

2. In the second category of questions, he suggests that numerical change can likewise accurately be studied only by using quantitative methods. For example: Are the numbers of primary school teachers in our district rising or falling? Is student learning achievement going up or down? For answering these types of questions, quantitative studies are best suited.
3. In the third category of questions, he includes those that attempt to find out about the state of something or other, we often want to explain phenomena. Examples are: What factors predict the recruitment of maths teachers? What factors are related to changes in students' achievement over time? Researchers can also successfully study this type of question by using quantitative methods. Statisticians have developed many statistical techniques that permit us to predict scores on one factor, or 'variable' (e.g., teacher recruitment) from scores on one or more other factors or 'variables' (e.g., unemployment rates, pay, conditions).
4. Finally, he suggests that quantitative research is especially suited for those questions involving the testing of 'hypotheses'. We might want to explain something—for example, whether there is a relationship between pupils' achievement and their self-esteem and social background. The researcher could look at the theory and come up with the hypothesis that lower social class background results in low self-esteem. This can be further related to low achievement. We can test this kind of hypothesis by using quantitative research.

Out of the four categories listed above, categories 1 and 2 are called 'descriptive' because here we are merely trying to describe a situation. Categories 3 and 4 are basically 'inferential' where we are attempting to explain something rather than just describe it.

In a nutshell, the main purpose of quantitative research is the quantification of data and its analysis by using several statistical methods. In principle, quantitative research highlights precisely what, where, when, how often and how long social phenomena occur. For example, imagine you wanted to research the impact of school feeding programme on pupils' attendance at primary level of education in your district. If you were to conduct quantitative research, you could count the number of pupils' who participated in this programme.

The main techniques used in quantitative research include highly structured, rigid techniques such as online questionnaires or telephone interviews. Unlike qualitative research, which seeks a deep, contextual understanding grounded in the participants' context, quantitative research designs rely primarily on responses to pre-formulated questions. Quantitative research is predominantly deductive. It means researchers start with a hypothesis, collect variables to test the hypothesis and draw conclusions about what the relationships between variable means.

When You Should Not Use Quantitative Methods?

As mentioned above, while quantitative designs are good at answering the above listed four types of questions, there are, however, other types of question that are not well suited to quantitative research designs:

1. The first situation where quantitative research will fail is when we want to explore a problem in depth. Quantitative research is good at providing information in breadth from a large number of units. But when we want to explore a problem or concept in depth, quantitative methods are too shallow. To get really under the skin of a phenomenon, we need to go for ethnographic methods, interviews, in-depth case studies and other qualitative techniques.
2. As mentioned earlier, quantitative research is well suited for the testing of theories and hypotheses. What quantitative methods cannot do very well is to develop hypotheses and theories. The hypotheses to be tested may come from a review of the literature or theory, but can also be developed using exploratory qualitative research.
3. If issues to be studied are particularly complex, an in-depth qualitative study (a case study, for example) is more likely to pick up on this than a quantitative study. This is partly because there is a limit to how many variables can be looked at in any one quantitative study, and partly because in quantitative research it is the researcher who defines the variables to be studied. In qualitative research, unexpected variables may emerge.
4. Finally, while quantitative methods are better to looking at cause and effect ('causality', as it is known), qualitative methods are more suited to looking at the meaning of particular events or circumstances.

What then do we do if we want to look at both breadth and depth, or at both causality and meaning? In these situations, it is best to use a so-called a 'mixed method design' in which we use both quantitative (for example, a questionnaire) and qualitative (for example, a number of case studies) methods. Mixed method research is a flexible approach where the research design is determined by what we want to find out rather than by any predetermined epistemological position. In mixed method research, qualitative or quantitative components can predominate or both can have equal status.

When conducting quantitative research, a researcher should have a clear understanding about the concepts of 'research units and variables' (discussed in succeeding chapters). When we collect data for quantitative research, we have to collect them from someone or something. The people or things (e.g., schools) we collect data on or from are known as research units, units or cases. If we have to draw the data from the samples of population, the samples are also called sampling units. They all are the same.

With quantitative research designs, you can generalise the findings of your research study to the population from which you draw its sampling units. The data that you collect from these units are known as 'variables'. Variables are any characteristic of the unit you are interested in and want to collect (e.g., gender, age, self-esteem). The term variable refers to the fact that this data will differ between units. For example, achievement will differ between students and schools; gender will differ between students and so on. Be careful; if there are no differences at all between units you want to study, you probably are not going to be able to do any interesting research. For example, studying whether students are human would not yield interesting findings.

How Can You Prepare a Research Design for Quantitative Studies?

If you are going to be conducting quantitative research, you might want to know how to prepare a basic research design. You will find this information in this tag which will help you understand how to do that, as well as how to interpret the research of others and arrange it into a cohesive paper.

Introduction

The introduction of your quantitative study should usually be written in the present tense and from the third person point of view. It should contain the following important information:

- Identify the research problem: describe emphatically, clearly and concisely the research problem you are planning to investigate.
- Review the literature: review pertinent and relevant materials on the topic, synthesise key themes and, if necessary, note down studies that have used similar methods of inquiry and analysis. Also, note down the areas where key gaps exist and how your study would help to fill these gaps or clarify existing knowledge.
- Describe the theoretical framework: provide an outline of the theory or the hypothesis you plan to test. Define unfamiliar or complex terms, concepts or ideas and provide the appropriate background information to place the research problem in proper context, if necessary (e.g., historical, cultural, economic and so on).

Methodology

In this section of your study, you should describe how each objective of your study would be achieved. Provide enough detail to enable the reader to make an informed assessment of the methods being used to obtain results associated with the research problem. Use past tense for writing this section of your study.

- Study population and sampling: identify and assess where did the data come from; how robust is it; where gaps exist or what was excluded. List the procedures used for their selection.

- Data collection: describe the data collection tools and methods used to gather information. Identify the variables being measured and note if the data were pre-existing (i.e., government data) or you gathered it yourself. If you gathered the information yourself, highlight and describe what type of instrument you used and why. Remember that no data set is perfect—describe limitations, if any, in methods of gathering data.
- Data analysis: describe all procedures you will be using for processing and analysing the data. Describe the typical instruments of analysis you used to study each research objective, including mathematical techniques and the type of computer software.

Results

- Findings: make sure that the finding of your study should be written objectively and in a succinct and precise format. Use graphs, tables, charts and other non-textual elements to help the reader understand the data. Ensure that non-textual elements do not stand in isolation from the text but are being used to supplement the overall description of the results and to help clarify key points being made.
- Statistical analysis: explain the techniques and methods of data analysis, that is, how did you analyse the data? What were the key findings from the data? Present the findings in a logical, sequential order. Highlight but do not interpret these trends or negative results; save that for the discussion section. Write the results in the past tense.
- Discussion: develop an analytic, logical and comprehensive discussion. Make sure that your discussion melds together your findings in relation to those identified in the literature review and is placed within the context of the theoretical framework underpinning the study. Write the discussion in the present tense.
- Interpretation of results: recapitulate the research problem being investigated, and compare and contrast the findings with the research questions underlying the study. Did they affirm predicted outcomes or did the data refute it?
- Description of trends, comparison of groups or relationships among variables: describe any trends that emerged from your analysis and explain all unanticipated and statistical insignificant findings.
- Discussion of implications: explain what the meaning of your results is. Underline key findings based on the overall results and note findings that you believe are important and innovative. How have the results helped fill gaps in understanding the research problem?
- Limitations: describe any limitations or unavoidable bias in your study and, if necessary, note why these limitations did not inhibit effective interpretation of the results of your study.

Conclusion

Conclude your study by summarising the topic and provide a final comment and assessment of the study on the following:

- Summary of findings: synthesise the answers to your research questions. Do not report any statistical data here; just provide a narrative summary of the key findings and describe what was learned that you did not know before conducting the study.

- Recommendations: if appropriate to the aim of the assignment, categorise key findings with policy recommendations or actions to be taken in practice.
- Future research: Describe the need for the follow-up of your research study linked to your study's limitations or to any remaining gaps in the literature that were not addressed/explored in your study.[1]

Strengths and Limitations of Quantitative Research Designs

Quantitative researchers try to recognise and isolate specific variables contained within the study framework, seek correlation, relationships and causality, and attempt to control the environment in which the data are collected to avoid the risk of variables, other than the one being studied, accounting for the relationships identified.

Among the specific strengths and limitations of using quantitative methods to study education and social science, Nenty (2009) identifies the following:

Strengths

- Allow for a broader study, involving a greater number of subjects and enhancing the generalisation of the results;
- Allow for greater objectivity and accuracy of results. Generally, quantitative methods are designed to provide summaries of data that support generalisations about the phenomenon under study. In order to accomplish this, quantitative research usually involves few variables and many cases and employs prescribed procedures to ensure validity and reliability;
- Apply well-established standard so that the research can be replicated and then analysed and compared with similar studies;
- Summarise vast sources of information and make comparisons across categories and over time and
- Avoid personal bias by keeping a 'distance' from participating subjects and using accepted computational techniques

Limitations

Quantitative methods presume to have an objective approach to studying research problems, where data are controlled and measured, to address the accumulation of facts and to determine the causes of behaviour. As a consequence, the results of quantitative research may be statistically significant but are often humanly insignificant (Nenty 2009).

Some specific limitations associated with using quantitative methods to study research problems in the social sciences and education, as stated by Nenty, include:

- These methods provide more efficient data that can be used to test hypotheses but may miss contextual detail.
- These methods are rigid and static and thus use an inflexible process of discovery.

---

[1] For a comprehensive account of basic research design for quantitative studies, see Nenty (2009).

- The development of standard questions by researchers can lead to 'structural bias' and false representation, where the data actually reflect the view of the researcher instead of the participating subjects.
- These methods provide very little, less detailed information on behaviour, attitudes and motivation of the participating subjects.
- There is a possibility of obtaining a much narrower and sometimes superficial data set.
- The results provide mostly numerical descriptions rather than a detailed narrative and generally provide less elaborate accounts of human perception.
- The environmental setting for carrying out research is often unnatural and artificial, so that a level of control can be applied to the exercise. This level of control might not normally be in place in the real world, thus yielding 'laboratory results' as opposed to 'real-world results'.
- Preset answers will not necessarily reflect how people really feel about a subject that, in some cases, might just be the closest match to the preconceived hypothesis.

Types of Quantitative Research Designs

We have listed the most widely used quantitative research designs in Chapter 2. In addition to these, there are some other designs used by researchers for some specific research studies. They have also been introduced in this section with a view to providing you a complete set of the quantitative research designs.

Experimental Research Design

In education, researchers use experimental method to manipulate those variables that define and explain one or more 'causes' in order to discern 'effects' on other variables. For instance, he wants to know how far two new textbooks are effective. The simplest way to investigate this is to use random assignment of teachers and students to three groups—two groups for each of the 'new textbooks', and one group as a 'control' group to use the 'existing textbook'. Let us describe and understand the basic elements of the experimental research design.

What is an Experimental Research Design?

An experimental research study is a study in which a treatment, procedure or programme is deliberately introduced to observe a result or an outcome. *The American Heritage Dictionary of the English Language* defines an experiment as 'A test under controlled conditions that is made to demonstrate a known truth, to examine the validity of a hypothesis, or to determine the efficacy of something previously untried'. According to Mayer (2005), 'an experimental research is a study in which a treatment, procedure, or program is intentionally introduced and a result or outcome is observed'.

The following example will enable you understand the concept of an experimental design in the field of education. Suppose you want to study the effect of a new curriculum designed to teach adolescent girls about HIV and pregnancy prevention, you introduce the curriculum into the classroom setting and do an initial presentation of first few chapters. You collect data and on the basis of these data, you might make some alterations to the next chapters. You can continue this process until the curriculum is substantially improved.

'In its simplest form, an experiment is a study in which a treatment, procedure or programme is intentionally introduced and a result or outcome is observed'.

Features of Experimental Research Design

True experiments have four elements: 'manipulation', 'control', 'random assignment' and 'random selection'. The most important of these elements are manipulation and control. 'Manipulation' means that something is purposefully changed by the researcher in the environment. 'Control' is used to prevent outside factors from influencing the study outcome. When something is manipulated and controlled and then the outcome happens, it makes us more confident that the manipulation 'caused' the outcome. In addition, experiments involve highly controlled and systematic procedures in an effort to minimise error and bias, which also increases our confidence that the manipulation 'caused' the outcome.

Mayer (2005) suggests the following five overarching features of design experiments.

- The purpose of design experiments is to develop theories about learning (including how learning is supported).
- Design experiments involve an intervention or the introduction of a new instructional technique.
- In design experiments, researchers attempt to develop new theoretical perspectives, but they also must test and refine their theories along the way.
- Design experiments have iterative designs; as theories change during the study, the design of the study must be revised and altered accordingly.
- The theories that are developed in design experiments should affect future instruction.

You should, however, note that it is extremely difficult to conduct the experiments in actual education settings, that is, in schools. The main reason is that the students can rarely be randomly assigned to classrooms in school settings (Box 3.2).

BOX 3.2: 'Video Games in Children's Learning of Mathematics in Malaysia'

Rationale

Past literature indicates that game and play are some of the best approaches for learning. However, contemporary society and educational discourse regard human learning only to be achieved through non-playful process as the public has associated gaining knowledge with hard labour. In contrast to this dominant belief that learning is through great effort and persistence, play and enjoyment can and should be considered as an integral part of a learning process. Although extensive studies have been done on educational computer games in Malaysia, the country is still behind other countries, and most of the studies and researches carried out here focus more on students' and teachers' perception and attitude towards accepting computer games in education and computer games as motivational tool for learning. A wide gap still exists in studies focusing on the effectiveness of computer or video games in children's learning of certain subjects in schools. Hence, this research aims to investigate the effectiveness of computer-based video game in facilitating children's learning of multiplication facts in Mathematics.

(Continued)

(Continued)

Objectives

There are two objectives which the study seeks to fulfil:

- To determine the relationship between the use of video games and learning
- To determine the effectiveness of computer-based video games in children's learning of multiplication facts in Mathematics

In order to achieve the above objectives, this study attempts to answer the research questions given below:

Research questions

- What is the difference in learning achievement of multiplication facts between students who used computer-based video game in learning and those who did not?
- What is the difference in learning achievement of multiplication facts between male and female students who used computer-based video game in learning and those who did not?
- What is the difference in learning achievement of multiplication facts between ethnic groups of students who used computer-based video game in learning and those who did not?
- What is the difference in learning achievement of multiplication facts between urban and rural students who used computer-based video game in learning and those who did not?

Methodology

The study investigates the effectiveness of a readily available computer game originally designed to teach multiplication facts to primary level students. The video game used in this study, Timez Attack, is a free downloadable game produced by Big Brainz, Inc., California. Timez Attack is a commercial game available off the shelf which is specifically designed to teach multiplication facts to children. This game is proposed as a supplementary classroom activity in schools or educational activity at home.

The study aims to generate a variety of data sets, allowing a comparison between students who participated in the game and students who did not. Data sets included gender, race and locality (urban students versus rural students). The various data groups were tested using identical testing situations and materials to allow a quantitative comparison of the scores of students participating in the game and students not participating in the game. Students have the same learning material and class content apart from the video game to reinforce the credibility of results generated from students' participation in the video game. As the data were expected to be approximately normally distributed, normal curve goodness of fit testing was used to test these assumptions. These tests were based on different pairs of sample data as laid out in the research questions and accompanying hypothesis. In addition, a standard six-step hypothesis test for each research question was used to determine whether to reject the null hypothesis.

Experimental design

Approximately 160 primary-level schoolchildren of two randomly selected schools in Perak (from one urban area school and one rural area school) will participate in the study. All students will be given a pretest on multiplication facts. Marks accrued from the test individually will be used to divide the students into two groups which should be equal in numbers and in students' performance. This is partly to ensure validity and reliability of the results of the experiment. The marks will also be compared with post-test results later; analysis of results using comparisons of means, *t*-tests, ANOVA and chi-squared tests will be used to investigate the efficiency of computer games as an added learning activity. The control group will be taught multiplication facts and skills conventionally. The experimental group would also undergo the same lessons but with the added activity: a computer-based video game on multiplication. The equal distribution of students of high and low achievers for both control and experimental groups are carefully administered. The experiment will be conducted for three months. Then all participants will be given a timed post-test to gauge students' accurate recall of multiplication facts in a given time accurately. Knowledge improvement from pre to post would facilitate the outcome measures for the study.

Source: Adapted from Video Games in Children's Learning of Mathematics, *International Journal of Basic and Applied Sciences IJBAS-IJENS* Vol. 11 No: 02.

What are the Aims of Experimental Research?

Experiments are conducted to be able to predict phenomenon. Typically, an experiment is constructed to be able to explain some kind of causation. Experimental research is important to society—it helps us to improve our everyday lives.

To simplify things, let us consider this example. A teacher in a primary school may be interested in knowing the impact of video game violence on children's aggression in his or her class. He or she randomly assigns some children to play violent video game for one hour and other children to play a non-violent video game also for one hour. Then he or she observes the children socialise afterwards to determine if the children in the 'violent video game' condition behave more aggressively than children in the 'non-violent video game' condition. In this example, the independent variable is video game group. The independent variable has two levels: violent video games and non-violent video games. The dependent variable is the thing that he or she wants to measure—in this case the aggressive behaviour.

From the above example, you must understand that the purpose is to investigate the relationship between two variables to test a hypothesis. By using the scientific method,[2] the teacher can plan and design an experiment that will answer the research question.

When Can You Use an Experimental Research Design?

You can use experimental research relatively easily when you are sure that:

- There is time priority in a causal relationship (cause precedes effect)
- There is consistency in a causal relationship (a cause will always lead to the same effect)
- The magnitude of the correlation is great

It may, however, be wise that you first conduct a 'pilot study' or two before you do the real experiment. This will ensure you that the experiment measures what it should, and that everything is set up right.

[2] The scientific method is the process by which new scientific knowledge is gained and verified. First, you must identify a question and, after some preliminary research, form a hypothesis to answer that question. After designing an experiment to test the hypothesis and collecting data from the experiment, a scientist will draw a conclusion. The conclusion will either support the hypothesis or refute it. The scientist will then either reformulate the hypothesis or build upon the original hypothesis. The scientific method cannot prove a hypothesis, only support or refute it.

You may notice minor errors, which could potentially destroy your experiment, are often found during this process. The pilot study can help you get information about errors and problems, and improve your design, before you put a lot of effort into the real experiment.

If the experiments involve humans, a common strategy that you should follow is to first have a pilot study with someone involved in the research, but not too closely and then arrange a pilot with a person who resembles the subject(s). Those two different pilots are likely to give you good and relevant information about any problems in the experiment.

How Can You Identify the Research Problem?

After deciding your topic of interest, the next task for you is to try to define the research problem. This will help you to focus on a more narrow research area to be able to study it appropriately. Defining the research problem will help you to formulate a research hypothesis, which is tested against the null hypothesis.

The research problem is often operationalised to define how to measure the research problem. The results will depend on the exact measurements that you will choose and may be operationalised differently in another study to test the main conclusions of the study.

An ad hoc analysis is a hypothesis invented after testing is done to try to explain why there is contrary evidence. Remember that a poor ad hoc analysis may be seen as your inability to accept that your hypothesis is wrong, while a great ad hoc analysis may lead to more testing and possibly a significant discovery.

How Can You Conduct an Experimental Research Study?

An experiment is typically carried out by manipulating a variable, called the 'independent variable', affecting the experimental group. The effect that the researcher is interested in, the 'dependent variable(s)', is measured.

Identifying and controlling non-experimental factors, which you do not want to influence the effects, is crucial to drawing a valid conclusion. This is often done by controlling variables, if possible, or randomising variables to minimise effects that can be traced back to third variables. You aim is only to measure the effect of the independent variable(s) when conducting an experiment, allowing you to conclude that this was the reason for the effect.

For a systematic conduct of an experimental study, Bogdan and Biklen (2007) suggest the following steps:

- Identify a question and perform preliminary research to determine what is already known.
- Formulate hypotheses and deduce their consequences.
- Identify and define the independent and dependent variables.
- Determine how the independent variable will be manipulated and how the dependent variable will be measured.
- Construct an experimental design that represents all the elements, conditions and relations of the consequences:
  - Select a sample of subjects.
  - Group or pair subjects.
  - Identify and control non-experimental factors.
  - Select or construct and validate instruments to measure outcomes.
  - Conduct a pilot study.
  - Determine the place, time and duration of the experiment.

- Conduct the experiment.
- Compile raw data and reduce them to a usable form.
- Apply an appropriate test of significance.

Mayer (2005) suggests that when researchers conduct design experiments, they should examine the effects of educational interventions in actual classrooms while the interventions are implemented.

Strengths and Limitations of Experimental Research Design

Strengths

One of the main strengths of experimental research is that it often enables the researcher to determine a cause-and-effect relationship between dependent and independent variables. By virtue of a systematic manipulation and isolation of the independent variable, the researcher can determine with confidence the causal effect of an independent variable and causal effect on the dependent variable.

Another strength of experimental research design is the ability to assign participants to different conditions through random assignment. This typical feature of experimental design ensures that each participant is equally likely to be assigned to one condition or another, and that there are no differences between experimental groups.

Limitations

Although experimental research can often answer the causality questions that are left unclear by correlational studies, this is not always the case. Sometimes experiments may not be possible or ethical. Consider the example of studying the correlation between playing violent video games and aggressive behaviour. It would be unethical to assign children to play lot of violent video games over a long period of time to see if it had an impact on their aggression.

Additionally, because experimental research relies on controlled, artificial environments, it can at times be difficult to generalise its outcomes or findings to real-world situations, depending on the experiment's design and sample size. If this is the case, it simply means that the experiment has poor external validity, meaning that the situation the participants were exposed bears little resemblance to any real-life situation.

In a nutshell, experimental research is one of the most appropriate methods for drawing and understanding causal conclusions about instructional interventions, that is, which instructional method is most effective for which type of student under which conditions. Mayer notes that 'experimental methods—which involve random assignment to treatments and control of extraneous variables—have been the gold standard for educational psychology. When properly implemented, they allow for drawing causal conclusions, such as the conclusion that a particular instructional method causes better learning outcomes'. This clearly suggests that if a researcher wants to determine whether a particular instructional intervention causes an improvement in student learning, then he or she should use experimental research methodology.

Experimental research is a study in which a researcher manoeuvres the level of some independent variable and then measures the outcome. The research method is powerful method for assessing cause-and-effect relationships. Many researchers consider experiments the 'gold standard' against which all other research designs should be judged (Mayer 2005). Experiments are conducted both in the laboratory and in real-life situations.

Quasi-experimental Research

In educational research, quasi-experimental designs are used to observe and analyse systematically the effects of a particular treatment on a particular population (through the use of a representative sample). Researchers use the method to infer if a treatment has any effect(s). In order to do this, it is essential for the researcher to first have some knowledge of what would have happened if the treatment had not been administered. They infer the effect(s) of a treatment on control group (the group that has not received the treatment) and experimental group (the group that receives the treatment). The basic aim of conducting a quasi-experiment, then, is to determine whether or not a given treatment caused a given effect.

What is a Quasi-experimental Research Design?

Quasi-experimental research designs, like experimental designs, is another widely used research design in social sciences. Researchers use it to test causal hypotheses. In both experimental and quasi-experimental designs, the researcher views a programme or policy as an 'intervention' in which a treatment—comprising the elements of the programme/policy being evaluated—is tested for how well it achieves its objectives, as measured by a pre-specified set of indicators. By definition, a quasi-experimental design does not have in it random assignment, however. Assignment to conditions (treatment versus no treatment or comparison) is done by means of self-selection (participants choose treatment for themselves) or chosen by administrator (e.g., by officials, teachers, policymakers and so on) or by both (Shadish, Cook and Campbell 2002).

> 'A quasi-experimental research aims to determine whether a programme or intervention has the intended effect on the participants of the research study. It is an empirical study used to estimate the causal impact of an intervention on its target population' (Shadish, Cook and Campbell 2002).

In simple terms, you should remember the following four main points to understand what a quasi-experimental research design is:

- It is a research design, like experimental design, which tests causal hypotheses.
- The design lacks, by definition, a random assignment.
- This design identifies a comparison group that is as similar as possible to the treatment group in terms of baseline (pre-intervention) characteristics.
- This design uses different techniques for creating a valid comparison group such as regression discontinuity design (RDD)[3] and propensity score matching (PSM).[4]

---

[3] An RDD is a quasi-experimental pretest–post-test design that elicits the 'causal effects' of interventions by assigning a cut-off or threshold above or below which an intervention is assigned. By comparing observations lying closely on either side of the threshold, it is possible to estimate the 'average treatment effect' in environments in which 'randomisation' is unfeasible.

[4] A PSM is a 'statistical matching' technique that attempts to 'estimate' the effect of a treatment, policy or other intervention by accounting for the 'covariates' that predict receiving the treatment. PSM attempts to reduce the 'bias' due to 'confounding' variables that could be found in an estimate of the treatment effect obtained from simply comparing outcomes among 'units' that 'received the treatment versus those that did not'.

As an overarching goal, the body of quasi-experimental research attempts to answer questions such as: 'Does a treatment or intervention have an impact?' and 'What is the relationship between programme practices and outcomes?'

There are several ways in which quasi-experimental studies take on forms. But these may best be defined as lacking key components of a true experiment. While a true experiment includes (a) pre- post-test design, (b) a treatment group and a control group and (c) random assignment of study participants, quasi-experimental studies lack one or more of these design elements (Shadish, Cook and Campbell 2002).

Quasi-experimental designs identify a comparison group that is as similar as possible to the treatment group in terms of baseline (pre-intervention) characteristics. The comparison group captures what would have been the outcomes if the programme/policy had not been implemented (i.e., the counterfactual). Hence, the programme or policy can be said to have caused any difference in outcomes between the treatment and comparison groups (Mayer 2005).

Similarities and Difference Between True Experimental and Quasi-experimental Research Designs?

Since true experimental and quasi-experimental research designs belong mainly to the category of experimental research, it is important for you to know the similarities and differences between these two designs so that you can choose the most appropriate one out of the two designs to serve the purpose of your research study.

Similarities between true and quasi-experimental research design include the following:

- Study participants are subjected to some type of treatment or condition
- Some outcome of interest is measured
- The researchers test whether differences in this outcome are related to the treatment

Differences between true experiments and quasi-experiments are as follows:

- In a quasi-experiment, the treatment or the control group are not assigned randomly.
- In a quasi-experiment, the control and treatment groups differ not only in terms of the experimental treatment they receive but also in other, often unknown or unknowable, ways. It is, therefore, necessary for the researcher to try to statistically control for as many of these differences as possible.
- Because control is lacking in quasi-experiments, there may be several 'rival hypotheses' competing with the experimental manipulation as explanations for observed results.

When Can You Use Quasi-experimental Research Designs?

You have seen above that quasi-experimental research designs resemble quantitative and qualitative experiments but lack random allocation of groups or proper controls, so firm statistical analysis can be very difficult. Thus, you can use this design when the aim of your study is:

- To meet the requirements of funding, school administrators and ethics
- To evaluate the effectiveness of an intervention
- To dedicate greater resources to issues of external and construct validity

In other words, quasi-experimental research designs are best suited when the aim of your research is to answer the effects of questions such as 'What particular effect(s) did a particular treatment have on a particular population?'

Example

- How does students' choice of reading material effect primary-aged students' reading motivation?
- What are the effects of inquiry-based teaching on high school students' reading comprehension scores on science texts?
- Which programme, writing workshop or self-regulated strategy development (SRSD) is more effective at increasing elementary students' narrative genre knowledge and writing ability?

How Can You Conduct an Experimental Research Study?

As most quasi-experimental studies embrace a pre- post-test design for both a treatment group and a control group, these studies are often considered as an impact evaluation. A possible danger in this method is that the treatment and control group may differ at the outset. For solving this problem, researchers doing quasi-experimental studies attempt to address this in a number of other ways, for example, by matching treatment groups to like control groups or by controlling for these differences in analyses.

Thus, your first step would be to select a quasi-experimental design which in turn would involve selecting groups upon which you will test a variable, without any random preselection process. For example, if you plan to perform an educational experiment, you divide the class arbitrarily by alphabetical selection or by seating arrangement. This division is quite convenient and in an educational setting will cause to as little disruption as possible.

After this selection, proceed your experiment in a very similar way to any other experiment, with a variable being compared between different groups, or over a period of time.

The two forms of quasi-experimental studies commonly used in education research: (a) a pre- and post-test design study without a control group, and (b) a pre- and post-test design with a control group are shown in Figure 3.1.

Once you have analysed pretest and post-test, your final step would be to interpret the findings of your study that have used a non-equivalent control group design with pretest and post-test cautiously since control and treatment groups may differ due to selection bias (Best and Kahn 2006). This is illustrated in Exercise 1 given in Box 3.3.

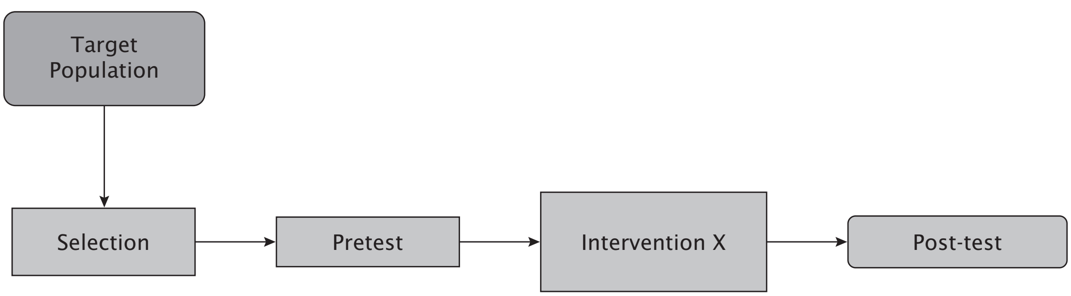


FIGURE 3.1 Quasi-experimental Research Study Design

Source: Authors.

BOX 3.3: 'The Effectiveness of Human Capital Policies for Disadvantaged Groups in the Netherlands'

Research design

This research study presents results on the effectiveness of different human capital policies targeted to disadvantaged groups in the Netherlands. The study concludes that lowering the compulsory school attendance age is the only intervention producing significantly positive effects. For other interventions, substantial positive effects can be ruled out.

Research questions

The study examines and analyses the effectiveness of the following policy measures/interventions on disadvantaged groups:

- Changes in the age at which young children are allowed to attend school
- Reducing class size in primary education
- A scheme that gives extra funding for personnel to primary schools
- A scheme that gives extra funding for computers to primary schools
- Increasing the length of lower vocational programmes from three to four years

Using advanced statistical methods, the study summarises the results from evaluation studies of five different education interventions.

- Effect of extra time in school on early test scores
- Effects of class size reduction on achievement in primary schools
- Effect of extra funding for personnel for schools with minority students
- Effect of extra funding for ICT for schools with disadvantaged students
- Effects of extending lower vocational programmes from three to four years

Source: Adapted from Leuven and Oosterbeek (2010).

While the control group showed a decline in reading scores from a mean of 40.02 to a mean of 36.25, the treatment group showed an increase in reading scores from a mean of 39.67 to a mean of 43.05. Does this mean that the school reform programme significantly impacted students' achievement scores? To make such a determination, researchers must complete a statistical analysis of the data with either an independent *t*-test or analysis of variance (ANOVA; Heiman 1999).

Advantages and Limitations of Quasi-experimental Research Design

Advantages

An important advantage of quasi-experimental studies is that the researcher can attempt to determine a cause-and-effect relationship even when there are limitations as to who and what you could study.

Especially in social sciences, where preselection and randomisation of groups is often difficult, quasi-experimental studies can be very useful in generating results for general trends. For example, if we study the effect of maternal alcohol use when the mother is pregnant, we know that alcohol does harm embryos.

A strict experimental design would conclude that mothers were randomly assigned to drink alcohol. This would be highly illegal because of the possible harm the study might do to the embryos. So what researchers do is to ask people how much alcohol they used in their pregnancy and then assign them to groups.

Quasi-experimental design can often be integrated with individual case studies. This allows the figures and results generated to reinforce the findings in a case study and permit some sort of statistical analysis to take place.

In addition, they do reduce the time and resources required for experimentation without an extensive pre-screening and randomisation.

Limitations

Without proper randomisation, statistical tests can be meaningless. Two limitations are generally associated with quasi-experiments. The first relates with internal validity. Since no random assignment is involved, the problem arises with regard to internal validity. Second, because there are problems with internal validity, the researcher cannot truly determine a cause-and-effect relationship.

A quasi experiment constructed to analyse the effects of different educational programmes on two groups of children, for example, might generate results that show that one programme is more effective than the other. These results cannot be considered rigorous statistical scrutiny because the researcher also needs to control other factors that may have affected the results. It is really very hard for the researcher to do it properly.

In short, quasi-experimental designs employ techniques to compare groups. A serious limitation of the quasi-experimental research design is randomisation. With this design, a control group is compared with an experimental group, but the groups are chosen and assigned out of convenience rather than through randomisation.

Cross-sectional Research Design

We have seen above that choice of selecting a study design depends greatly on the nature of the research question. In other words, you should know the research methodology the kind of information your study should collect. This is the first step in determining how you will carry out your study.

Consider this example. Let us say we want to investigate the relationship between private coaching and students' test score in economics in our secondary school. One of the first things we would have to determine is the type of study that will tell us the most about that relationship. Do we want to compare test score levels in economics among different students populations of taking private coaching and non-coaching at the same point in time? Or do you want to measure test score levels in a single population taking private coaching over an extended period of time?

The first approach is typical of a cross-sectional study. The second requires a longitudinal study. To make our choice, we need to know more about the benefits and purpose of each study type.

Figure 3.2 presents a cross-sectional design. The researcher formulates a research question, identifies three distinct groups and compared them at a given point in time.

What is Cross-sectional Research Design?

A cross-sectional attempts to examine people of different ages at the same time(s). A researcher develops cohorts from the study population, so that he can examine how people of different ages perform, behave or respond to a particular function. For instance, he may ask one type of test to

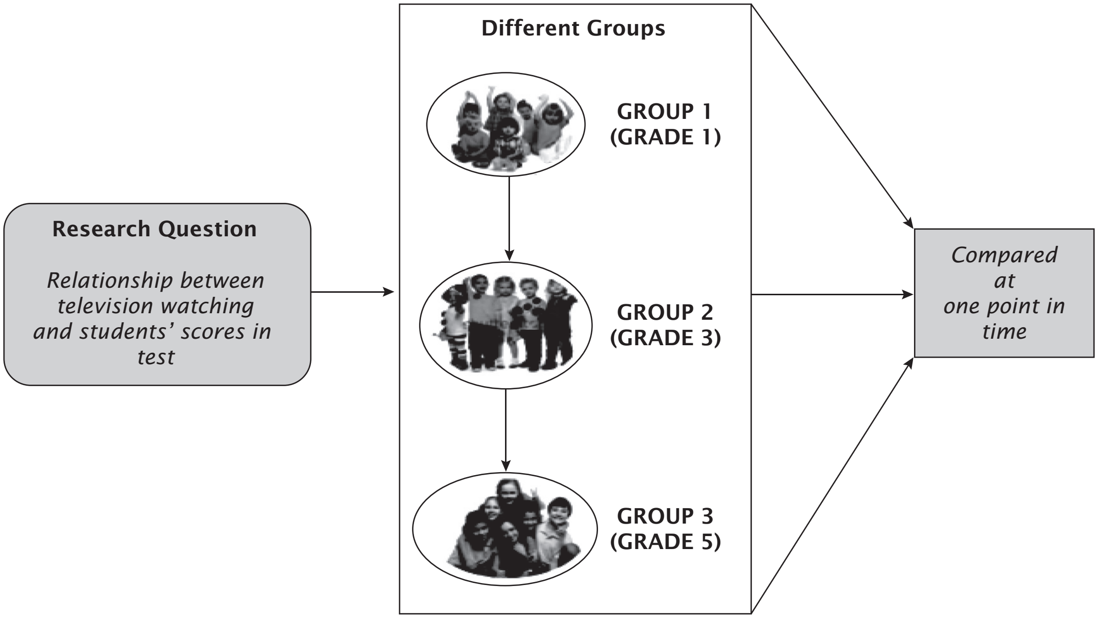


FIGURE 3.2 Cross-sectional Research Design

Source: Authors.

children in the Grade 3 (cohort 1), 4 (cohort 2) and 5 (cohort 3), to examine the differences in performance across these age groups.

> 'A cross-sectional study is one that produces a "snapshot" of a population at a particular point in time. The study involves looking at people who differ on one key characteristic at one specific point in time' (Ruspini 2002).

To simplify the concept further, let us consider another example. With a cross-sectional research study, you can attempt to determine whether there is a relationship between television watching and students' grades because you believe that students who watch lot of television do not have time to do homework and perform poorly in school. In this case, the researcher will design a simple questionnaire about number of hours spent watching television and course grades and then collect data to all of the children in a given school.

The design is used to capture information based on data gathered for a specific point in time. In this design, the investigator gathers data from a pool of participants with varied characteristics and demographics known as variables such as age, gender, income, education, geographical locations and ethnicity. However, the choice of variable or demographics depends on the type of research being conducted and on what the study aims to prove or validate. The research findings assist him to remove assumptions and replace them with actual data on the specific variables studied during the time period under consideration in the cross-sectional study.

A cross-sectional study is primarily an observational study. This means that researchers record information about their subjects without manipulating the study environment. The researchers do not try to modify the behaviour of participants.

Cross-sectional studies are basically descriptive. A purely 'descriptive cross-sectional' study is used to assess the frequency and distribution of a particular trait in a defined population. For example, a random sample of schools across Bangkok may be used to assess the burden or prevalence of asthma among 12–14 year olds.

A national census is the perfect example of the cross-national study. Here, a representative sample of the population comprising individuals of different ages, different occupations, different educational levels and income levels and residing in different parts of the country is interviewed on the same day. Ruspini asserts that in education, cross-sectional studies depend primarily on indirect measures of the nature and rate of changes in the physical and intellectual development of samples of children drawn from representative age levels (Box 3.4).

BOX 3.4: 'Effect of an Educational Intervention on HPV Knowledge and Vaccine Attitudes among Urban Employed Women and Female Undergraduate Students in China: A Cross-sectional Study'

Background

Despite recent strides made in Human Papilloma Virus (HPV) vaccination and education, over half a million women worldwide develop cervical cancer each year. Over 85 per cent of these cases occur in developing countries such as China due to the lack of effective screening and prevention programmes. The incidence of cervical cancer in Mainland China increased from 5.14 persons per 100,000 in 2004 to 6.87 persons per 100,000 in 2008.

Methods

Participants and recruitment

This was a multi-centre, cross-sectional study across the five main geographical regions of Mainland China. One large, economically developed city was selected from each of the five geographic regions of China. Participants were then selected from one comprehensive university and three to four companies in each city for a total of six comprehensive universities and 16 companies.

Informed consent

Prior to enrolment, all participants were fully informed about the objectives of the study, benefits and risks associated with participation and assured of the confidentiality of the information they provide. After informed consent was obtained, a 62-items, multiple-choice questionnaire was administered to participants.

Educational intervention

The educational intervention consisted of a one-hour group lecture followed by the same questionnaire as the one given after informed consent. Questions were written in Mandarin Chinese in simple terms appropriate for the level of education of the participants and previously validated. Participants were encouraged to complete the questionnaire independently, and trained staff members stood by to answer any questions that participants had.

Conclusion

This study demonstrates that incorporation of the proposed educational initiative into a government-sponsored or school-based programme may improve willingness to accept HPV vaccination. The potential implications of the study are far reaching during this pivotal time of HPV vaccine clinical trials in China. It is important to establish the link between HPV infection and the development of cervical cancer and to educate women on the ideal time to vaccinate their children. Studies have shown that maternal attitudes towards vaccination greatly impacts vaccination rates among children. It is also imperative to dispel common beliefs that only women with multiple sexual partners need to be vaccinated.

Future directions for research include the development of a longitudinal, national HPV education programme appropriate for school-based curriculum. Due to our limited participant sampling, our findings may not be generalised to the entire heterogeneous Chinese female population. Widespread implementation among different geographic regions and minority populations across China would shed light on the long-lasting impact of HPV education. Also, it is important to develop evaluation and feedback tools to assess the effectiveness of the programme. At last, follow-up studies must be performed to see if increased HPV and HPV vaccine knowledge translates into higher vaccination rates and vaccine compliance within China.

Source: Adapted from Chang et al. (2013).

What Are the Salient Features of a Cross-sectional Research Design?

The salient feature of a cross-sectional study is that it can compare different population groups at a single point in time. It is just like taking a snapshot. Findings are drawn from whatever fits into the frame. However, you should recognise that cross-sectional studies may not provide definite information about cause-and-effect relationships. The simple reason is that such studies offer a snapshot of a single moment in time; they do not consider what happens before or after the snapshot is taken. Therefore, you cannot know for sure if, for example, your students taking private coaching had low test score levels in economics before taking up their private coaching, or if taking private coaching helped them to obtain higher test score levels that previously were low.

Cross-sectional research studies are based on observations that take place in different groups at one time. This means that there is a complete absence of any kind of experimental procedure. In other words, the researcher cannot use any variable for manipulation. Instead of performing an experiment, he or she would simply record the information that he or she observes in the groups he or she is examining. Consequently, he or she uses a cross-sectional research study for describing the characteristics that exist in a group, but it cannot be used to determine any relationship that may exist. This method is used to gather information only. The information may then be used to develop other methods to investigate the relationship that is observed.

Cross-sectional research studies all have the following characteristics:

- They take place at a single point in time.
- Researchers cannot manipulate variables.
- They provide information only; they do not answer 'why'.

When Can You Use Cross-sectional Research Design?

This type of study is widely used in business, psychology, social science, retail, medicine, education, religion and government. In each of these disciplines, cross-sectional research provides important data that inform all kinds of actions. For business marketing, in particular, this tool is used to learn more about various demographics for the purpose of analysing target markets to sell to or introduce products and services.

You can use a cross-sectional study design when the purpose of your study is to find the prevalence of the outcome of interest, for the population or subgroups within the population at a given time point.

Repeated cross-sectional studies may be carried out to give a pseudolongitudinal study, where the individuals included in the study are either chosen from the same sampling frame or from a different one. An example might be the study of underweight prevalence among children in which you can examine children aged less than 5 years annually and record the prevalence of underweight. The prevalence of underweight for this age group is monitored over time and this information is used in school health policy planning and in the development of targeting strategies.

How Can You Conduct an Experimental Research Study?

The steps that you will have to follow to conduct a cross-sectional study are more or less the same as the ones discussed above for the quantitative research design except some minor adjustment. These are explained below.

Choosing a Representative Sample

A cross-sectional study should be representative of whole population, if generalisations from the findings are to have any validity. For example, a study of the prevalence of diabetes among women aged 40–60 years in Town A should comprise a random sample of all women aged 40–60 years in that town. For your study to be representative, you should always attempt to include hard-to-reach groups, such as people in institutions or the homeless.

Sample Size

The size of your sample should be sufficiently large enough to estimate the prevalence of the conditions of interest with adequate precision. Sample size calculations can be carried out using sample size tables or statistical packages such as statistical package for social sciences (SPSS). The larger the study, the less likely the results are due to chance alone, but this will also have implications for cost.

Data Collection

Since cross-sectional studies collect data simultaneously on both exposures and outcomes, it is necessary that the data collection should establish specific inclusion and exclusion criteria at the design stage, to ensure that those with the outcome are correctly identified. The data collection methods will depend on the exposure, outcome and study setting, but will include questionnaires and interviews. Routine data sources may also be used.

Potential Bias in Cross-sectional Studies

A key problem affecting cross-sectional studies relates to non-response which, in turn, can result in bias of the measures of outcome. This is a typical problem when the characteristics of non-responders differ from responders.

Analysis of Cross-sectional Studies

In a cross-sectional study, all factors (exposure, outcome and confounders) should be measured simultaneously. Prevalence is the main outcome measure obtained from a cross-sectional study.

For continuous variables such as blood pressure or weight, values will fall along a continuum within a given range. Prevalence may, therefore, only be computed when the variable is divided into those values that fall below or above a particular predetermined level. In this typical case, mean or median values should alternatively be calculated for such variables.

Advantages and Limitations of Cross-sectional Research Design

Advantages

An important benefit of a cross-sectional study design is that it allows researchers to compare many different variables at the same point in time. We could, for example, look at age, gender, income and educational level in relation to study the relationship between having school textbooks and test score levels, with little or no additional cost.

Other advantages include:

- Relatively inexpensive and takes up little time to conduct
- Can estimate prevalence of outcome of interest because sample is usually taken from the whole population
- Focus on studying and drawing inferences from existing differences between people, subjects or phenomenon. They help the generation of hypotheses
- Groups identified for the study are purposely selected based upon existing differences in the sample rather than seeking random sampling
- Contains multiple variables at the time of the data snapshot
- Many findings and outcomes can be analysed to create new theories/studies or in-depth research

Limitations

Cross-sectional research design is not free from limitations. Some the important limitations are:

- Finding people, subjects or phenomena to study that are very similar except in one specific variable can be difficult
- Results are static and time bound and, therefore, give no indication of a sequence of events or reveal historical or temporal contexts
- Studies cannot be utilised to establish cause-effect relationships
- This design only provides a snapshot of analysis so there is always the possibility that a study could have differing results if another time frame had been chosen
- There is no follow-up to the findings
- The timing of the snapshot is not guaranteed to be representative
- Findings can be flawed or skewed if there is a conflict of interest with the funding source

In a nutshell, a cross-sectional study involves looking at people who differ on one key characteristic (such as age) at one specific point in time. The data are collected at only one point in time from people who are similar on other characteristics but different on a key factor of interest such as age, income levels and geographic locations. Participants are usually separated into groups known as cohorts. For example, researchers might create cohorts of participants who are in their 20s, 30s and 40s.

This type of study uses different groups of people who differ in the variable of interest but who share other characteristics such as socio-economic status, educational background and ethnicity. Cross-sectional studies are often used in developmental psychology, but this method is also utilised in many other areas including social science and education.

Longitudinal Research Study Design

The world's longest running longitudinal study is the Genetic Studies of Genius. It is referred to now as the Terman Study of the Gifted (Terman 1925). The study originally started in 1921 by psychologist Lewis Terman to investigate how highly intelligent children developed into adulthood.

It is still an ongoing study, although the original sample has understandably grown much smaller. The study originally started with over 1,000 participants, but that number had dwindled to just 200 by 2003. Researchers at the Stanford University plan to continue the study until the last participant either drops out or dies.

What is Longitudinal Research Design?

A longitudinal study (aka longitudinal survey, or panel study) is a research design that involves repeated observations of the same variables (e.g., people) over long periods of time, often many decades (i.e., uses longitudinal data). It is often an observational study, although they can also be structured as longitudinal randomised experiments.

A longitudinal research design is a type of non-experimental developmental research design that can be used in order to study age-related changes in behaviour. Within the design, the same participants are observed (Figure 3.3) recurrently over a period of time. This period of time may be as brief as six months to a year or quite long—sometimes even spanning a lifetime. The researcher may study one specific

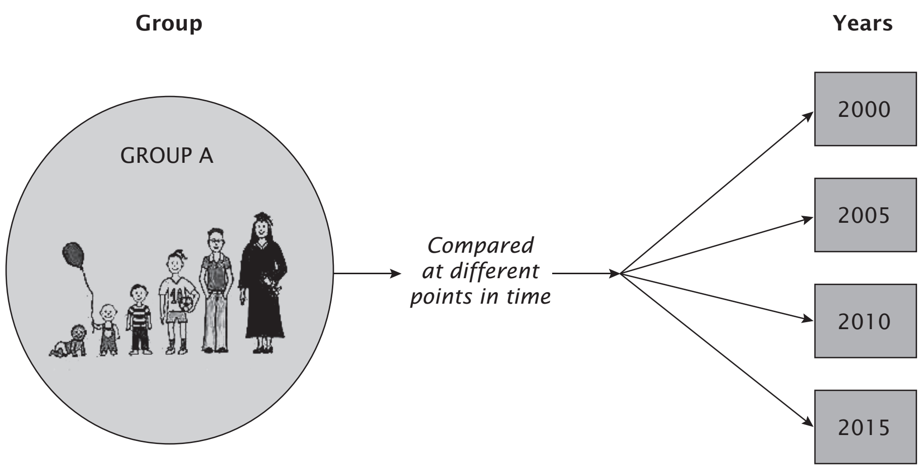


FIGURE 3.3 Longitudinal Research Study

Source: Authors.

BOX 3.5: 'Pre-elementary Education Longitudinal Study (PEELS)'

PEELS is following almost 3,000 children with disabilities as they progress through preschool and into their early elementary years. The study collects data on the preschool and early elementary school experiences of a nationally representative sample of children with disabilities and the outcomes they achieve. It focuses on children's preschool environments and experiences, their transition to kindergarten, their kindergarten and early elementary education experiences and their academic and adaptive skills. The children were 3 to 5 years old at the start of the study in 2003 and were followed through to 2008.

Design

Over several years, information about the children, their families and instructional/service programmes were collected using family interviews, child assessments and multiple questionnaires. The PEELS sample consists of approximately 1,000 children in each age cohort (3, 4 and 5 year olds) receiving special education services at the study onset. The data collected through PEELS are used for descriptive as well as for explanatory research purposes. Direct assessments of children's academic skills and parent interviews have been conducted. In addition, questionnaires were sent to participating children's teachers, school principals, directors of local education agencies and state's 619 coordinators.

Source: Adapted from the Institute of Education Sciences, National Centre for Special Education Research, US Department of Education.

aspect of development (for example, intelligence) or many. The subjects are usually studied in the form of cohort. It means they roughly belong to the same age and have grown up in similar circumstances. Longitudinal studies are often described as a number of observations followed by a period of ageing and development then another set of observations (Box 3.5).

> 'Longitudinal research enables researchers to analyse the duration of social phenomena, highlight similarities, differences and changes over time in respect of one or more variables or participants (within and between participants); identify long-term effects; and explain changes in terms of stable characteristics, e.g. sex, or variable characteristics, such as income' (Ruspini 2002).

Longitudinal studies can be retrospective (looking back in time, thus using existing data such as medical records or claims database) or prospective (requiring the collection of new data). ~~[citation needed]~~

One example of a longitudinal developmental study is that of Howes and Matheson (1992), in which the pretend play of a group of 1 to 2 years old children were repeatedly observed every 6 months over a period of 3 years. Howes and Matheson found that complexity of pretend play increases with age and is also a reliable predictor of children's future social competence with peers.

What Are the Different Types of Longitudinal Research?

There are three major types of longitudinal studies:

- Panel Study: involves sampling a cross-section of individuals. A panel study involves a sample from different groups that still fit the subjects of the study.

- Cohort Study: a group of people who share a defining characteristic, typically who experienced a common event in a selected period, such as birth, geographic location, historical experience or graduation and perform cross-section observations at intervals through time.
- Retrospective Study: involves looking at the past by looking at historical information such as medical records.

Longitudinal studies of the cohort analysis type have an important place in educational investigations. They have considerable potential for yielding rich data that can trace changes over time and with great accuracy (Gorard 2001). On the other hand, there is serious problem of attrition (participants leaving the research over time, a particular problem in panel studies which research the same individuals over time). The studies can be both costly and time-consuming (Ruspini 2002).

Cohort, trend and panel studies are prospective longitudinal methods, in that they are ongoing in their collection of information about individuals or their monitoring of specific events (Ruspini 2002). The cohort method studies separate samples from a single group over time. In a cohort study, a researcher selects a group on the basis of a specific event such as birth, geographic location or historical experience. Trend studies analyse samples from succeeding groups over time whereas in a panel study a single sample (same individuals) from a group is studied over time.

How Does Longitudinal Research Work?

We have seen that longitudinal research is used to discover relationships between variables that are not related to various background variables. This observational research technique involves studying the same group of individuals over an extended period.

At the outset of the study, you should collect data and then do the same repeatedly throughout the length of your study. For example, imagine that you are interested in studying how exercise during middle age might impact cognitive health as people age. You hypothesise that people who are more physically fit in their 40s and 50s will be less likely to experience cognitive declines in their 70s and 80s.

In this particular case, you obtain a group of participants who are in their mid-40s to early 50s. Then collect data related to how physically fit the participants are, how often they work out and how well they do on cognitive performance tests. Periodically over the course of the study, you should collect the same data from the participants to track activity levels and mental performance.

There are many research studies in education that focus on developmental issues—how individuals change overtime? For instance, we all know that the reading methods of young children are not the same as the ones adopted by older children. You can examine these changes over time by taking longitudinal research (Box 3.5). For longitudinal studies, the data collection process is relatively simple. It is collected for the same individual over a number of different time periods. We collect data for the same group of students who successfully complete class assessments at the end of Grade 1, Grade 2, Grade 3 and Grade 4. We then use these data to examine change in student data across those four grades over four years.

Consider another example. Imagine that you want to study how physical exercise during middle age might impact cognitive skills as people age. A possible way to formulate your hypothesis could be:

> People who are more physically fit in their 40s and 50s will be less likely to experience declines in cognitive skills in their 70s and 80s.

For this particular case, you will select a sample of participants who are in their mid-40s to early 50s. You will collect data related to how physically fit the participants are, how often they work out and how well they do on cognitive performance tests. Periodically over the course of the study, you will also collect the same data from the participants to track activity levels and mental performance.

For testing the hypothesis, you have to decide how long the subjects need to be assessed and evaluated, and how often we will be evaluating them. If we are interested in studying the initial effects only, then we might choose to study the group for one year after we return. If our aim is to study long-term changes, we may do the exercise for 20 years or more depending on how much resources and time we have at our disposal.

There are certain key things that you have to keep in mind about longitudinal studies. They are observational studies in nature and are a kind of correlational research. Longitudinal research is often contrasted with cross-sectional research and involves collecting data over an extended period, often years or even decades whereas cross-sectional research involves collecting data at a single point in time (see cross-sectional research above). The identifying feature of longitudinal research is studying one group of individuals over time.

To sum up, a longitudinal study is observational. There is no interference with the respondents. What makes a longitudinal study unique is the timeline. Instead of a researcher collecting data from varying respondents in order to study the same variables, he observed the same respondents several times and often over the course of many years (Apel 2014).

Another important consideration that you have to remember is that you must decide who are the participants to be researched and how many of them to be included in the study (sample). The type of longitudinal study you choose will help you determine the sample size. For instance, if you choose one group of pupils who were admitted together in a given year in Grade 1 of primary education in a given school you would be doing a cohort study (student cohort analysis). If you instead look for pupils who were in different districts and in different socio-economic settings, then we would do a panel study. A panel study involves a sample from different groups that still fit the subjects of the study. But, if you are interested in looking at past and current pupils, you could use historical records. This study is a retrospective study. It is a study that looks back. For this retrospective longitudinal study, you will have to retrieve (or search) the past data/records of research participants.

When Can You Use Longitudinal Research Design?

Longitudinal studies are widely used in psychology to study developmental trends across the lifespan and in sociology to study life events throughout lifetimes or generations. The reason for this is that longitudinal studies track the same people and so the differences observed in those people are less likely to be the result of cultural differences across generations. Longitudinal studies thus observe changes and allow the researcher to make them more accurate. The studies allow social scientists to distinguish short from long-term phenomena, such as poverty. If the poverty rate is 10 per cent at a point in time, this may mean that 10 per cent of the population is always poor or that the whole population experiences poverty for 10 per cent of the time. It is impossible with longitudinal studies to conclude which of these possibilities is the case by using one-off cross-sectional studies.

In an educational setting, longitudinal research deals with a wide range of research questions. They enable you to address topics such as:

- Describing and defining developmental change
- Predicting event occurrence
- Identifying treatment effects

- Assigning causality
- Measuring durations between events
- Describing occurrences of phenomena over different time periods

What Are the Advantages and Limitations of Longitudinal Research Design?

So why might you choose to conduct longitudinal research? For many types of research, longitudinal studies can provide you unique insight that might not be possible with other forms of research designs.

Advantages

- The design allows researchers to assess the stability and continuity of several attributes of a sample by repeatedly observing the same participants (Kagan and Moss 1962).
- The design allows researchers to identify developmental trends by looking for common attributes that the subjects share, for example, points at which most children undergo changes (Newman et al. 1997).
- Longitudinal design avoids effects of different cohorts because only one group of people is examined by the researcher over time, rather than comparing several different groups that represent different ages and generations.
- Researchers can describe how a single individual's behaviour changes with age.
- The design combines both qualitative and quantitative data, creating more in-depth research (Ruspini 2002).
- The researcher can look at changes over time. Thus, longitudinal methods are particularly useful when studying development and lifespan issues.
- Since the genetics of participants share are similar, it is assumed that any differences are due to 'environmental factors'.
- They help researchers to establish a sequence of events when looking at the ageing process.

Limitations

Following are the limitations and the disadvantages associated with using this form of research design.

- Longitudinal research is very time-consuming.
- The design demands high commitment in order to continue and complete the duration of the study. Participants tend to drop out over time.
- So is the case with the researcher. He or she must remain interested in the research whilst he or she waits for years to see the final results.
- The research design is also very expensive to conduct, since the researchers must track people down and persuade them to come back and participate in the study.
- The design requires periodic training of the research team over many years.
- By the end of the research project (after 10 or 20 years), the project outcome may seem trivial.

In conclusion, in the longitudinal developmental research design, the same group of participants are observed and measured at different intervals over a period of time, thus cohort effects are not a problem.

Stability, continuity and normative trends can easily be identified, and quantitative methods can be used together with qualitative methods.

Survey Research

If you have ever been sitting at a train station, in a particular lecturer's classroom or in a public area and a person with a stack of papers in his or her hands comes up to you out of the blue and asks if you have a few minutes to talk, then you have likely been asked to take part in a survey.

What is Survey Research?

A survey research provides a comprehensive, representative summary of specific characteristics, beliefs, attitudes, opinions or behaviour patterns of a population. The researcher can collect survey information by using different methods, such as in person, phone and e-mail questionnaires and interviews.

Surveys, though primarily quantitative in design and implementation, are always conducted in response to a qualitative inquiry (Lodico, Spaulding and Voegtle 2006). Surveys are not experimental, so data are not collected to test a hypothesis, but rather, to describe—both qualitatively and quantitatively, but primarily quantitatively—existing conditions and attitudes.

In education, surveys are used particularly in descriptive research studies. Researchers use them to gather information on test scores in order to identify patterns of low achievement, to see new teachers' attitudes towards teaching and to identify changing trends in student interests. Survey research has also been used to collect information about adult students' reading patterns and about students' traits in order to identify ways to encourage adults to read (Mellard, Patterson and Prewett 2007).

How Can You Implement and Administer a Survey?

There are different ways in which a survey can be administered. For instance, in a 'structured interview', the researcher asks each participant the questions. In a 'questionnaire', the participant completes/fills the questionnaire by himself or herself. There are no specific questions in either of them; in interviews, participants are required to choose their answers between previously determined categories, while in questionnaires the researcher may ask for open-ended responses. Di Iorio (2005) suggests that for gathering information of a personal or sensitive nature, it is better that we use 'open-ended questions'. As regards quantifiable information, he suggests the use of 'multiple-choice questions'.

Survey approach is most suited for gathering descriptive information. Surveys can be both structured and unstructured. In structured surveys, a formal list of questions is for all respondents in the same way. However, unstructured surveys allow the interviewer probe respondents and guide the interview in accordance to their answers.

Survey research may be direct or indirect. In direct approach, we ask direct questions about behaviours and thoughts. For example, 'Why do you not attend the class?' In indirect approach, the researcher might ask: 'What kind of pupils attend the class?' From the response, the researcher may be able to discover why the pupil avoids the class.

To ensure reliability and validity, surveys are generally standardised. Standardisation is also essential for the generalisation of results to the larger population.

How Can You Design a Survey?

The key to a good survey is its design. Thus, it is necessary to choose the survey items carefully to produce the data needed to answer your research questions. You should make sure that the survey items are clear and are not respondent-biased towards particular answers. When the survey is the main data collection instrument in a study, you should include the full content (design) of the survey in an appendix or make it available upon request.

The first step in the design of a survey is the enunciation of the research question, the scope of the study and the targeted population. Next, the research question should be subdivided into themes or subquestions that reflect different dimensions of the problem. Survey items are then formulated and selected using scaling (see the following sections), a method of generating questions and statements that adequately measure—quantitatively—the qualitative aspects of the attitude or belief under investigation.

Once the researcher has selected survey items, and completed the questionnaire or interview format, the survey must undergo pilot testing before it can be used to collect and generalise information about the targeted population.

Which Scaling Techniques Can You Use in Survey Research?

We have seen above a classification system that is often used to describe the measurement of concepts or variables that are used in educational and behavioural research. This classification system categorises the variables as being measured either by a nominal, ordinal, interval or ratio scale. For example, when a question has the response choices of 'always', 'sometimes', 'rarely' and 'never', this is a scale because the answer choices are rank-ordered and have differences in intensity. Another example would be 'strongly agree', 'agree', 'neither agree nor disagree', 'disagree' and 'strongly disagree'.

You should keep in mind that scaling is not simply a process of assigning numerical values to qualitative statements, such as a score of 5 for 'highly agree' or of 1 for 'highly disagree'. Scaling is the intricate, nuanced technique of developing qualitative statements that gauge respondents' beliefs about a particular issue; thus, scaling research always begins with the clear identification of a research question or aim. In this particular case, you should conduct an experimental research to decide appropriate qualitative statements for use in surveys that adequately gauge the desired features in the population of interest.

There are several formalised conceptual and experimental frameworks commonly used in scaling research. The formalised conceptual and experimental frameworks most relevant and commonly used by educational researchers are:

Likert Scale

Likert scale, the most commonly used scale in social science research, typically has the following format:

Scale 0 = Strongly agree
Scale 1 = Agree
Scale 2 = Neither agree nor disagree
Scale 3 = Disagree
Scale 4 = Strongly disagree

You should note that the individual questions that use this format are called Likert items while the Likert scale is a sum of several Likert items. To create the scale, each answer choice should be assigned a score (say 0–4) and the answers for several Likert items (that measure the same concept) then be summed up together for each individual to get an overall Likert score.

Let us say that you are interested in measuring prejudice against women. One possible way to do that would be that you create a series of statements reflecting prejudiced ideas, each with the Likert response categories listed above. One of your items might be 'women should not be allowed to vote', while another might be 'women cannot drive as well as men'. You then assign each response category a score of 0 to 4 (e.g., assign a score of 0 to 'strongly disagree', a 1 to 'disagree', a 2 to 'neither agree nor disagree' and so on). Finally, you sum up the scores for each of the statements for each respondent to create an overall score of prejudice. If we had five statements and a respondent answered 'strongly agree (score 4)' to each item, his or her overall prejudice score would be 20, indicating a very high degree of prejudice against women.

Bogardus Social Distance Scale

The Bogardus social distance scale is a technique for measuring the willingness of people to participate in social relations with other kinds of people.

Let us say we are interested in the extent to which Hindus in India are willing to associate with, say, Schedule Caste (SC) Hindus. We might ask the following questions:

- Are you willing to live in India as SC?
- Are you willing to live in the same community as SC?
- Are you willing to live in the same neighbourhood as SC?
- Are you willing to live next door to a SC?
- Are you willing to let your child marry a SC?

The Bogardus scale exhibits that scales can be important data reduction tools. By knowing how many relationships with SC a given Upper Class Hindu respondent will accept, we know which relationships were accepted. A single number can thus accurately summarise five or six data items without a loss of information.

Thurstone Scale

The Thurstone scale is intended to develop a format for generating groups of indicators of a variable that have an empirical structure among them. For example, if you were studying discrimination, you would put together a list of items, say, 10 items. Then you will ask the respondents to assign scores of 1 to 10 to each item. In fact, you ask them to rank the items in order of which is the weakest indicator of discrimination all the way to which is the strongest indicator.

Once the respondents have scored the items, only then you examine the scores assigned to each item by all the respondents to determine which items the respondents agreed upon the most.

Semantic Differential Scale

In this type of scale, you ask respondents of your research questionnaire to select between two opposite positions using qualifiers to bridge the gap between them. Let us look at an example. Suppose your aim is to obtain respondents' opinions about a new comedy television show. The first thing is that you have to decide the dimensions you wish to measure and then find out two opposite terms that represent those dimensions. For example, 'enjoyable' and 'unenjoyable', 'funny' and 'not funny', 'relatable' and 'not

TABLE 3.2 Dimensions of Responses

	Very Much	*Somewhat*	*Neither*	*Somewhat*	*Very Much*
Enjoyable		X			Unenjoyable
Funny				X	Not Funny
Relatable			X		Unrelatable

Source: Adapted from Babbie (2001).

relatable'. You would then create a rating sheet for each respondent to indicate how he or she feels about the television show in each dimension. Your questionnaire would look something like the one shown in Table 3.2.

Guttman Scaling

Guttman scaling technique is used for determining whether a set of indicators constitute a scale and, if so, how to score them. The goal of the scaling is to derive a single dimension that can be used to position both the questions and the subjects. The position of the questions and subjects on the dimension can then be used to give them a numerical value (Abdi 2010). Guttman scaling is used in social psychology and in education.

Suppose that you want to test a set of children and that you assess their mastery of the following types of mathematical concepts:

1 = counting from 1 to 50
2 = solving addition problems
3 = solving subtraction problems
4 = solving multiplication problems
5 = solving division problems

You will find from your interaction with the children that some children will be unable to master any of these five mathematical skills. Therefore, these children will not provide information about the problems so you will not consider them. You then will notice that some children will master only counting and nothing else, some will master addition and you expect them to have mastered addition but no other concepts; some children will master subtraction and you expect them to have mastered both counting and addition; some children will master multiplication and you expect them to have mastered subtraction, addition and counting. Finally, you will find that some children will master division problem and you expect them to have mastered all: counting, addition, subtraction and multiplication. What you do not expect to find, however, are children, for example, who have mastered division but who have not mastered addition, subtraction or multiplication. So the set of patterns of responses that you expect to find is well structured and is shown in Table 3.3. The pattern of data shown in this table is consistent with the existence of a single dimension of mathematical ability. In this framework, a child has reached a certain level of this mathematical ability (division) and can solve all the problems below this level (addition, subtraction and multiplication) and none of the problems above this level.

TABLE 3.3 Pattern of Responses of a Perfect Guttman Scale*

	Problems				
Children	*Counting*	*Addition*	*Subtraction*	*Multiplication*	*Division*
S_1	1	0	0	0	0
S_2	1	1	0	0	0
S_3	1	1	1	0	0
S_4	1	1	1	1	0
S_5	1	1	1	1	1

Source: Abdi (2010).
*A value of 1 means that the child (row) has mastered the type of problem (column); a value of 0 means that the child does not master the type of problem.

Both rows and the columns of the table can be represented on a single dimension. You will order the operations from the easiest to the hardest and will position a child on the right of the most difficult type of operation solved. So the data from Table 3.3 can be represented by the following order:

Counting S_1 Addition S_2 Subtraction S_3 Multiplication S_4 Division S_5

In this scaling technique, you can transform this order into a set of numerical values by assigning numbers with equal steps between two contiguous points. For example, this set of numbers can represent the numerical values corresponding to Table 3.3:

Counting	S_1	Addition	S_2	Subtraction	S_3	Multiplication	S_4	Division	S_5
1	2	3	4	5	6	7	8	9	10

This scoring scheme implies that the score of an observation (i.e., a row in Table 3.3) is proportional to the number of non-zero variables (i.e., columns in Table 3.3) for this row.

How Can You Collect Survey Data?

You can collect data for your survey research in a number of different ways. Some of the most common ways to implement a survey include:

- Mail
- Telephone
- Online
- At-home interviews (census surveys)

Advantages and Limitations of Using Surveys

Advantages

- Surveys allow collection of a large amount of data in a relatively short period of time.
- They are less expensive.

- They can be easily and quickly created and administered with ease.
- Surveys can be used to collect information on personal facts, attitudes, past behaviours and opinions.

Limitations

- Poor survey construction and administration can jeopardise and undermine otherwise well-designed studies.
- The answer choices contained in a survey may not be a true and accurate reflection of participants' feeling.
- Survey response rates can bias its results.
- Respondent's reluctance to answer questions asked by unknown interviewers.

Correlational Research Design

Correlational research attempts to discover or establish relation with two or more variables and to make predictions based on this relationship. Correlation does not imply 'causation'. It means that if one factor correlates with another, it does not mean that the first factor is the cause of the other or that these are the only two factors involved in the relationship. Only the experimental research can establish cause and effect.

> 'A correlational study is one that attempts to collect data to determine whether and to what degree a relationship exists between two or more variables. The degree of relationship is expressed as a correlation coefficient (*r*). One of the features of correlational research is that cause and effect relations cannot be determined, i.e. which variable causes a change in the other' (Bloomberg and Volpe 2012).

Thus, correlation research serves primarily two aims:

- It determines relationship between two or more variables.
- It makes prediction and does not explain the behaviour.

The attributes of correlations include strength and direction. The strength, or degree, of a correlation ranges from +1 to –1 and therefore can be positive, negative or zero. Direction refers to whether the correlation is positive or negative. For example, two correlations of +0.87 and –0.87 have the exact same strength but differ in their directions (+0.87 is positive and –0.87 is negative). In contrast, two correlations of +0.04 and +0.97 have the same direction (positive) but are very different in their strength. Although +0.04 indicates a relatively weak relationship, +0.97 indicates an extremely strong relationship between two variables. A correlation of 0.00 indicates no relationship between the variables.

In short, a correlational study is a quantitative method of research in which the researcher has two or more quantitative variables from the same group of subjects. He tries to determine if there is a relationship (or covariation) between the two variables (a similarity between them, not a difference between their means). Theoretically, any two quantitative variables can be correlated as long as the researcher has scores on these variables from the same participants; however, it is probably a waste of time to collect and analyse data when there is little reason to think these two variables would be related to each other.

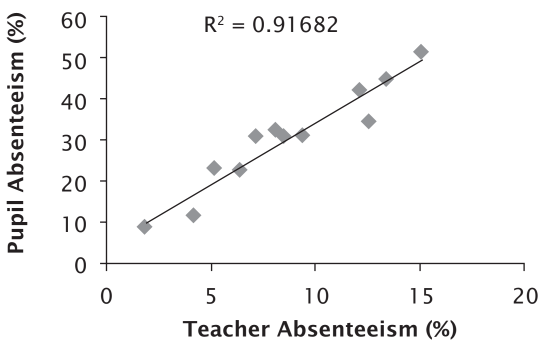
Correlational research is defined and described by a 'scatter plot' (XY plot), which has several points that show the relationship between two sets of data.

Let us consider the following example. Suppose that a headmaster in a rural primary school wants to investigate the relationship between pupils absenteeism versus teacher absenteeism. He or she keeps track of pupils' and teachers' attendance (pupils and teachers who were absent) on a particular day. Following are their figures.

Day	Pupil Absenteeism Versus Teachers Absenteeism in Rural Primary School	
	Teachers Absent (%)	Pupils Absent (%)
1	4.2	11.5
2	6.4	22.5
3	1.9	8.5
4	5.2	23.2
5	8.5	30.6
6	12.1	42.2
7	9.4	31.2
8	15.1	51.4
9	13.4	44.4
10	8.1	32.1
11	12.6	34.5
12	7.2	30.8

Source: Authors.

This data are shown below as scatter plot. It is now easy to see that higher absenteeism of teachers leads to pupils' higher absenteeism, but the relationship is not perfect ($r \neq +1$).



Source: Authors.

It should be noted that a correlation coefficient (*r*) near +1.00 indicates that the variables are highly and positively related. An increase in one variable is associated with an increase in the other.

In correlational research, we cannot manipulate variables; rather we collect data on existing variables and examine relations between those variables. There are several statistical techniques that you can use to analyse correlational data (see chapter on data analysis, Chapter 13). An analysis of statistical relation between high school students' standardised examination scores in mathematics and the students' demographic characteristics (e.g., gender, ethnicity, socio-economic status and so on) is a good example of correlation analysis.

Correlation research should not be confused with 'causal research'. A correlational study only describes or predicts behaviour but does not explain it. Thus, while using this method of research, you should always keep in mind that correlation does not imply causation. The fact that two variables are related does not necessarily imply that one causes the other. To prove any kind of causal relationship, you are required to do further research.

A minimally accepted sample size for a correlational study is 30 participants. In addition, the instruments chosen for the research must be valid and reliable. Therefore, for ensuring higher validity and reliability, the researcher should keep the samples smaller in size.

Let us assume that you want to measure the pupil–teacher ratio in each classroom in a school district and the average student achievement on the state assessment in each of these same classrooms. Next, you want to measure whether the pupil–teacher ratio and student achievement in the school district are connected numerically, that is, when the pupil–teacher ratio changes in value, so does student achievement. You can use the pupil–teacher ratio to predict student achievement by a statistical technique called 'regression analysis'. But for studies having more than one predictor variable, the statistical technique called 'multiple regression analysis' is used.

The explanation of exploratory, descriptive and correlation research methodologies given above suggests that the three research approaches are important elements in fields such as marketing, technological and educational research and the social sciences (Bloomberg and Volpe 2012). Let us remind ourselves that researchers use exploratory research when little is known about the topic and previous theories or ideas do not apply. They use exploratory research when the aim of the study is to clarify problems, collect data and create initial hypothesis and theories about subjects. Finally, they employ correlation methods to search the relationship between variables. These three concepts can be best explained in the following example cited by (Babbie 2010).

The TV shows produced by a popular crime investigation television channel provide a fairly good illustration of the three types of research design. These shows usually begin with a crime that must be investigated (an unplanned change has occurred in the marketplace). According to Babbie, the first step is to search for clues that can help establish what has happened (exploratory research). The clues uncovered in the exploratory phase of the investigation often point towards a particular hypothesis or explanation of the events that occurred, and investigators begin to focus their efforts in this direction, conducting interviews with witnesses and suspects (descriptive research). Finally, a trial is held to determine whether the evidence is sufficient to convict a suspect of the crime (correlation/causal research; Babbie 2010).

Explanatory Research

Explanatory research is a method of research in which the principal objective is to know and understand the trait and mechanisms of the relationship and association between the independent and dependent variables. It is undertaken to explain 'why' and 'how' phenomena occur, that is, the relationship between two or more aspects of a situation or phenomenon and to predict future occurrences. The approach uses correlations to study relationship between dimensions or characteristics of individuals, groups, situations or events. It explains how the parts of a phenomenon are related to each other.

> 'Explanatory research is an attempt to connect ideas to understand cause and effect, meaning researchers want to explain what is going on. Explanatory research looks at how things come together and interact. This research does not occur until there is enough understanding to begin to predict what will come next with some accuracy' (Bogdan and Biklen 2007).

Explanatory studies are characterised by research hypotheses that specify the nature and direction of the relationships between or among variables being studied (Box 3.6). Within this method, we use questionnaires, group discussions, interviews and random sampling techniques to explain behaviour and identify the underlying causes and effects of associated with it.

BOX 3.6: 'Dropping Out of School in Southern Ghana: The Push and Pull Factors'

Background

This research study goes beyond the first-level attributions of dropout to delve under the surface and explore the dynamics of decision-making and actions that lead to dropout, profile the forms it takes and encourage or discourage re-entry into school. This generates a rich tapestry of insight into the life histories of different children who have dropped out. This shows how both factors on the supply and demand side are influential and may interact. It also illustrates how important action at the local level may be and how the actions of 'significant others' may determine if dropout occurs and if it becomes permanent.

Purpose of the study

The purpose of this study is to highlight the causes of dropout (why children drop out) from the accounts of children who experienced dropout. This study highlights the conditions that create the processes that led to dropout.

Organisation of the study

This study is organised into six chapters. Chapter 1 focuses on the background and purpose of the study. Chapter 2 highlights literature on the factors that influence dropout. Chapter 3 presents the context and methods of the study. Chapter 4 is organised into two parts around the research questions, with Part 1 focusing on conditions outside school that pushed and/or pulled children out of school and Part 2 focusing on conditions within the school that led to dropout. Chapter 5 addresses the final research question about the motivations behind children's decision to drop back in school or not. And finally, views from children who return to school about their chances of staying and completing school are explained in Chapter 6.

Data collection

Data were gathered by means of in-depth interviews with children and observation of their activities.

Source: Adapted from CREATE.

Exploratory research is an investigation into a problem or situation. It provides details where a small amount of information exists and an understanding concerning a problem or situation even for cases where the amount of exiting information is too little. In exploratory research, you may use a variety of methods such as trial studies, interviews, group discussions, experiments or other tactics for the purpose of gaining information.

Exploratory Research

Exploratory research deals with the formulation of problems, clarifying concepts and forming hypotheses. The method enables the researcher to explore an unknown area or an area where little is known. The method investigates the possibilities of undertaking a particular research study (feasibility study/pilot study).

> 'Exploratory research is defined as the initial research into a hypothetical or theoretical idea' (Bogdan and Biklen 2007).

With this method, a researcher seeks to understand more about an idea or about something he or she has already observed. An exploratory research study is an attempt to lay the foundation or groundwork that will lead to future studies or to determine if what is being observed might be explained by a currently existing theory.

The exploratory research intends to produce the following insights:

- Close knowledge of and familiarity with basic details, settings and concerns
- Well-founded scenario of the situation
- Generation of new ideas and assumptions
- Development of tentative theories or hypotheses
- Directions for future research and the techniques

The three well-known and widely used research designs—exploratory, descriptive and correlation—are interrelated, that is, one method follows the other. With 'exploratory research', we discover ideas and insights. 'Descriptive research', on the other hand, usually describes a population in relation to key variables. Finally, 'correlation (causal) research' enables the researcher to establish cause-and-effect relationships between variables. The purposes differ and are distinct. We use experimental designs to study causality for determining cause and effect.

Almost all education research studies use exploratory and descriptive research designs. How much of each design is necessary will depend mostly on how much we know about the issue to be studied at the time of initiating our research. We also know that each education system experiences periodic emergence of pressing problems emanating as a result of unplanned changes in the environment. In situation such as this, there is usually a need for exploratory research to better understand what is happening and why it is happening. When confronted with such issues, the focus of the researcher quickly shifts to descriptive research which is geared more towards providing answers than generating initial insights.

Exploration generally starts with a literature search, a focus group discussion or case studies. If our aim is to conduct research for exploratory purposes, there is no need to examine a random sample of a population; rather, we should usually look for individuals who are knowledgeable about a topic or process. Exploratory research typically seeks to create hypotheses rather than test them. Data from exploratory studies tend to be qualitative. Examples include brainstorming sessions, interviews with experts and posting a short survey to a social networking website.

Consider this example. Imagine that you conducted interviews with parents of primary school pupils and noticed that they seemed to be more concerned about the quality of education in their primary school as compared to similar other primary schools in the town. This might lead to the hypothesis that teachers' teaching style and preferences had changed, resulting in lower learning achievements of pupils. You cannot really confirm or reject the hypothesis with exploratory research though. That job is left for descriptive and/or causal research (these are often called 'quantitative research').

Exploratory research (sometimes referred to as 'qualitative research') should not be expected to provide answers to take the decision pertaining to the problem that the researcher is attempting to solve. There are two reasons for this:

- Exploratory research usually entails only a relatively small group of participants.
- These participants are not selected randomly as participants of the study.

Being small in size, we simply cannot afford to devote the bulk of the research budget to exploratory research. And because we often do not know a lot at the beginning of a project, exploratory studies are very flexible with regard to the methods used for gaining insights and developing hypotheses. Although there are a number of common types of exploratory research, some of the more popular methods of exploratory research include literature searches, depth interviews, focus groups and case analyses.

'Literature search' is one of the fastest and least expensive means to discover hypotheses. There is enormous quantity of information available in libraries, via Internet sources, in commercial databases and so on. The literature search may also include newspapers, magazines, academic literature or published statistics from research organisations or governmental agencies' Census Bureau.

Imagine that we want to undertake a research inquiry on the issue: 'Why is the quality of primary level education lower?' This can easily be evaluated with the help of published data which should indicate whether the issue is a 'system problem' or a 'particular school problem'. If we admit the specific situation that our school's quality of education is lower regardless of the national average showing an uptrend, then we have to investigate and evaluate both the endogenous and exogenous school mix variables (Kothari 2004).

In any research inquiry, a good literature search is important. But at the same time, 'depth interviews' with persons who are well informed in the area being investigated should be organised. For this, the researcher does not need a questionnaire. But the depth interviews should be highly unstructured, so as to obtain divergent views of participants. The purpose of depth interviews is to tap the knowledge and experience of individuals with information directly associated with the situation or opportunity at hand. Anybody with related information could be a potential candidate for a depth interview.

Yet another frequently used method in exploratory research is the 'focus group'. In a focus group, only a few people are brought together to study and talk over some theme of interest. A moderator who is in the room with the focus group participants directs the discussion. The group usually is of 6–10 persons. While choosing these individuals, care must be exercised to see that they should have a common background and have comparable experiences. This is certainly needed since there should not be a conflict among the group members on the common problems that are being discussed and talked about. Throughout the discussion, individual attitudes with respect to the issues related to quality education should be collected.

'Case analyses' provide considerable information about a problem by studying carefully selected examples or cases of the phenomenon. Case histories of schools that have gone through an identical problem are generally available in both libraries and the documentation centre of the MOE. These case studies are suitable to undertake exploratory research. A researcher must examine carefully the previously published case studies with regard to variables such as schooling condition, teachers' involvement and commitment to teaching and so on.

To conclude, an exploratory research is a systematic investigation of relationship among two or more variables. The method describes primarily relationship, predicts the effects of one variable on another and tests relationships that are supported by clinical theory. Exploratory research explores the research questions and does not intend to offer final and conclusive solutions to existing problems.

Simulation Research

Literally, simulation means imitation of a situation or process. However in research, 'simulation refers to the process of designing a model of a real system and conducting experiments with this model for the purpose of understanding the behaviour of the system and/or evaluating various strategies for the operation of the system' (Pegden, Shannon and Sadowski 1995).

> 'Simulation is the imitation of the operation of a real-world process or system over time' (Banks et al. 2001).

Researchers undertake simulation for:

- Gaining insight into the operation of a system.
- Developing operating or resource policies to improve system performance.
- Testing new concepts and/or systems before implementation.
- Gaining information without disturbing the actual system.

Simulation's greatest strength is its ability to answer 'what if' question. The models are an integral part of an education plan. Educational planners use simulation models in different planning steps. The simulation models are capable of estimating future enrolment and resource needs for specific quantifiable goals and targets concerning:

- Grade 1 intake rate
- Promotion, repetition and dropout rates
- Transition rates
- Re-integration rates (of previous years' dropouts)
- Class size
- Pupil–teacher ratio or class–teacher ratio and use of classrooms (single or multi-shifts)
- In-service teacher training
- Distribution of textbooks, teacher guides and other teaching–learning materials
- Staff salaries (basic scales or increments)
- Special education development programmes, and initiation and upgrading of facilities (such as libraries, labs)
- Share of public/private education
- Parents' contribution and external assistance to education (loans and grants from IOs and NGOs) and so on.

Simulation models are useful in exploring the consequences of specific decisions on the evolution of the education sector. However, they cannot predict exactly what will happen, but can only show what might happen if specific decisions are made and nothing else changes. At the same time, the outputs of the simulation models are useful in projecting what might be achieved, when and at what cost, if and only if the assumptions, under which the models are developed, remain valid until the end of the projection period.

A simulation exists in a variety of forms, from the mathematical (consisting of abstract numerical expressions) to the physical. Four primary types of simulation models include the iconic, analogue, operational and mathematical. The first two have more to do directly with the physical context ('iconic' = testing of materials or products; 'analogy' = dynamic simulation of a physical system). 'Operational' models emphasise role-play of individuals in a physical system. 'Mathematical' models are systems of numerical coding that capture quantifiable real-world relationships

Postlethwaite (2005) states the following strengths and weaknesses of simulation models:

Strengths

- Simulation research is able to capture complexity without being reduced to a limited number of discrete variables.
- This research strategy provides a variety of ways of understanding future behaviour.
- Because all research strategies involve the 'real world' in some way, simulation tends to be useful to a variety of other strategies.

Weaknesses

- The project of replicating a slice of the real world is necessarily limiting (never 'complete'). Particular limitations include lack of spontaneity in role-playing or the challenges associated with coding aspects of human behaviour into computer equivalents.
- Simulation research can become very expensive very fast. Often, computer experiments are used to study simulation models. Simulation is also used with scientific modelling of natural systems or human systems to gain insight into their functioning (encyclopaedia of computer science). Simulation can be used to show the eventual real effects of alternative conditions and courses of action. Simulation is also used when the real system cannot be engaged, because it may not be accessible, or it may be dangerous or unacceptable to engage or it is being designed but not yet built or it may simply not exist (Sokolowski and Banks 2009).

Summary

Quantitative methods are most commonly used method for examining the relationship between variables. Since the data collected are numeric, allowing for collection of data from a large sample size, the statistical analysis used in this method permits for greater objectivity when reviewing results and, therefore, results are independent of the researcher.

Numerical results are generally displayed in graphs, charts, tables and other formats that allow for better interpretation.

Data analysis is less time-consuming. Statistical computations can be done using user-friendly statistical software. Results can be generalised if the data are based on random samples and the sample size is sufficient large enough. Data collection methods are relatively quick and less expensive depending on the type of data being collected. Numerical quantitative data are more credible and reliable especially to policymakers, decision-makers and administrators.

There are a variety of quantitative methods and sampling techniques that have been cited in other chapters.

The collection of numerical data through quantitative research methods lends itself well to large variety of research questions. The forthcoming chapters will explore when to choose quantitative methods, how to write a good research question, types of quantitative methods, data analysis, ethics and many other topics that will lead to better understanding of quantitative research.

Self-test Exercises

Exercise 3.1:

3.1.1. Testing the relationship between the scores on an intelligence test and scores on a personality test. Which type of research is it?

- ☐ Qualitative research
- ☐ Quantitative research
- ☐ Both

3.1.2. A researcher wants to investigate the impact of a math intervention on adding and subtracting fraction. Two groups of students are randomly assigned either to a group that will be taught using the math intervention or to a control group that does not receive the math intervention.

- 3.1.2.1. How would you measure student achievement in adding and subtracting fractions?
- 3.1.2.2. Which statistical methods would you use to detect causal relationships?
- 3.1.2.3. Based upon your theoretical or conceptual framework, what types of prediction you can make tentatively and use them to support or reject these predications?

3.1.3. A psychology professor is interested in whether implementing weekly quizzes improves student learning. He or she decides to use the weekly quizzes in one section of his or her introductory psychology class and not to use them in another section of the same.

- 3.1.3.1. Which type of quasi-experimental design do you recommend for this study?
- 3.1.3.2. Identify some possible confounds in studies you outlined in answers to Q 3.1.
- 3.1.3.3. Give three reasons a researcher might choose to use a single-case design.

3.1.4. A teacher may wonder how students benefit from the use of reading comprehension strategies to recall information from a science textbook.

Design a cross-sectional research outline for answering the following questions:

- 3.1.4.1. How would you test the hypothesis that ‘there are age differences in children’s ability to benefit from using reading comprehension strategies that can be traced using a cross-sectional design’?
- 3.1.4.2. After formulating the research question, how would you identify groups of children to participate in the study and to determine what, if any, experimental groups you want to include?

3.1.4.3. How would you test special groups of children, such as those with learning disabilities, slow readers or those from less-advantaged homes and compare them to typically developing children?

3.1.5. A researcher wants to look how identical twins reared together versus those reared apart differ on a variety of variables and how growing up in a different environment influences things such as personality and achievement. What method would you use to look at the following?

3.1.5.1. What the participants have in common versus where they differ?
3.1.5.2. Which characteristics are more strongly influenced by either genetics or experience?
3.1.5.3. What are the noticeable personality and behavioural changes overtime?
3.1.5.4. How would you establish a sequence of events when looking at the ageing process?

3.1.6. In an educational research programme, two groups of students are taught a topic in different ways. An experimental group uses a spreadsheet to explore the topic and a control group uses a more traditional pen and paper activity. Each group contains 20 students. At the end of the class, the teacher tests the two groups on their understanding of the topic and obtains the following data:

Experimental	5	11	25	33	35	40	45	46	52	55
	56	56	57	59	69	74	75	89	92	97
Control	33	39	44	45	45	46	47	48	49	49
	53	54	54	55	58	60	61	63	65	69

Computed data

	Mean	Standard Deviation
Experimental	53.6	25.0
Control	51.8	9.1

3.1.6.1. How would you interpret these findings?

3.1.7. A school inspector would like to know what relationships exist between different domains of Criterion-Referenced Competency Test (CRCT) scores. He or she has provided your principal with the included excel file, 'Correlation Data', which is a random sample of Grade 7 students. The students were chosen randomly in such a way that all Grade 7 students had an equal chance of being selected. Since the principal knows you are taking a research course, he or she has asked you to complete the assignment.

3.1.7.1. How can you interpret the coefficient of correlations given in the following table?

Seventh Grade CRCT Score Correlations

Reading	English	Math	Science	Social Science
1.00				
0.87	1.00			
0.71	0.76	1.00		
0.74	0.67	0.71	1.00	
0.80	0.65	0.98	0.65	1.00

References

Abdi, H. 2010. 'Guttman Scaling'. In *Encyclopaedia of Research Design*, edited by Neil Salkind, 558–560. Thousand Oaks, CA: SAGE Publications.

Apel, B. 2014. 'What is a longitudinal study?' *Survey Gizmo*. Available at https://www.surveygizmo.com/survey-blog/longitudinal-vs-cross-sectional-studies-whats-the-difference/ (accessed on 16 November 2017).

Babbie, E. 2001. *The Practice of Social Research*, 9th ed. Belmont, CA: Wadsworth Thomson.

———. 2010. *The Practice of Social Research*, 12th ed. Belmont, CA: Wadsworth Cengage.

Banks, J., J. Carson, B. Nelson and D. Nicol. 2001. *Discrete-Event System Simulation*, 3. Upper Saddle River, NJ: Prentice Hall.

Best, J. and J. Kahn. 2006. *Research in Education*. Boston, MA: Pearson Education, Inc.

Bloomberg, L.D. and M. Volpe. 2012. *Completing Your Qualitative Dissertation: A Road Map from Beginning to End*. Thousand Oaks, CA: SAGE Publications.

Bogdan, R.C. and S.K. Biklen. 2007. *Qualitative Research for Education: An Introduction to Theories and Methods*. Boston, MA: Pearson.

Chang, I.J., Rong Huang, Wei He, Shao-Kai Zhang, Shao-Ming Wang, Fang-Hui Zhao, Jennifer S. Smith and You-Lin Qiao. 2013. *BMC Public Health 2013*. Available at http://www.biomedcentral.com/1471-2458/13/916 (accessed on 28 July 2017).

Di Iorio, C.K. 2005. *Measurement in Health Behavior: Methods for Research and Evaluation*. San Francisco, CA: Jossey-Bass.

Gorard, S. 2001. *Quantitative Methods in Educational Research: The Role of Numbers Made Easy*. London: Continuum.

Heiman, G. 1999. *Research Methods in Psychology*. Boston, MA: Houghton Mifflin Company.

Howes, C. and C.C. Matheson. 1992. 'Sequences in the Development of Competent Play with Peers: Social and Social Pretend Play'. *Developmental Psychology* 28 (5): 961–974.

Kagan, J. and H.A. Moss. 1962. *Birth to Maturity*. New York, NY: Wiley.

Kothari, C.R. 2004. *Research Methodology: Methods and Techniques*, revised edition. New Age International.

Leuven, E. and H. Oosterbeck. 2010. *The Effectiveness of Human Capital Policies for Disadvantaged Groups in the Netherlands*. Available at: https://www.researchgate.net/publication/237622322_The_effectiveness_of_human_capital_policies_for_disadvantaged_groups_in_the_Netherlands1 (accessed on 29 January 2018).

Lodico, M.G., D.T. Spaulding and K.H. Voegtle. 2006. *Methods in Educational Research: From Theory to Practice*. San Francisco, CA: Jossey-Bass.

Mayer, R.E. 2005. *The Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press.

Mellard, D., M.B. Patterson and S. Prewett. 2007, April/May/June. 'Reading Practices among Adult Education Participants'. *Reading Research Quarterly* 42 (2): 188–213.

Nenty, H.J. 2009. 'Writing a Quantitative Research Thesis'. *International Journal of Educational Science* 1 (1): 19–32.

Newman, D.L., A. Caspi, T.E. Moffitt and P.A. Silva. 1997. 'Antecedents of Adult Interpersonal Functioning: Effects of Individual Differences in Age 3 Temperament'. *Developmental Psychology* 33 (2): 206–217.

OECD. 2016. *Teaching and Learning International Survey (TALIS)*. Paris: Organisation for Economic Co-operation and Development.

Pegden, C.D., R.E. Shannon and R.P. Sadowski. 1995. *Introduction to Simulation Using SIMAN*, 2nd ed. Singapore: McGraw Hill.

Postlethwaite, N. 2005. *Educational Research: Some Basic Concepts and Terminology Module 1*. Paris: UNESCO International Institute for Educational Planning.

Ruspini, E. 2002. *Introduction to Longitudinal Research*. London: Routledge.

Shadish, W., T. Cook and D. Campbell. 2002. *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin Company.

Sokolowski, J.A. and C.M. Banks. 2009. *Principles of Modelling and Simulation*. Hoboken, NJ: Wiley.

Terman, Lewis M. 1925. *Mental and Physical Traits of a Thousand Gifted Children*, Genetic Studies of Genius, Vol. 1. Stanford, CA: Stanford University Press.

4 CHAPTER

Other Research Methods

Introduction

We described in Chapters 2 and 3 the most commonly and widely used qualitative and quantitative research designs and methods in social sciences in general and in educational research in particular. This chapter attempts to develop a basic understanding of other important designs and methods hitherto not been discussed in Chapters 2 and 3 so as to provide you an understanding of the key principles and practices of research designs necessary to get your research project up and running.

We have noticed that the range of research issues and research designs the researchers use in their investigation is vast. But, if we look at the research studies published in professional journals, we might observe that these studies are largely quantitative in nature. It does not mean that other types of research designs are not important for gathering and understanding of information. There is likelihood for the non-quantitative studies to contain very little quantitative data, but these studies are very important and highly useful as they often help you provide diverse snapshots, colours, understanding and personalisation to information that might otherwise be missed. This chapter attempts to describe these methods.

Scientific Research Methods

Experts often divide scientific research methods into two broad classes: 'conceptual research method' and 'empirical research method'. These are the two most commonly and distinct ways of doing research.

The central theme in scientific research methods is that all evidence must be empirical which means it is based on evidence. In scientific methods, the word 'empirical' refers to the use of working hypothesis that can be tested using observation and experiment. The empirical data are produced by experiment and observation. Thus, any scientific research process should serve the following objectives. It should:

- capture contextual data and complexity
- learn from the collective experience of others from the field
- identify, explore, confirm and advance the theoretical concepts
- further improve educational design

We explain both these methods in some details in the following paragraphs.

Conceptual Research Method

Conceptual research is basically considered as analytical studies. They relate to some abstract idea(s) or theory. Its users are primarily the philosophers and thinkers who employ the method to develop new concepts or to redefine and reinterpret the existing ones.

> 'Conceptual research focuses on developing a theory to explain specific phenomena or behaviours. It tests the predictions of conceptual research by focusing on real people and real situations. Then concepts or theories may need to be adjusted to explain empirical findings' (Glaeser and Dickson 2013).

The conceptual research is a highly popular research method in the fields of humanities and philosophy (Box 4.1). Here, the researcher addresses and tackles the problem part by part. For better understanding the concept and its scientific application, the researcher breaks down the concept into smaller simpler parts. This method of analysis has gained immense popularity. In spite of this popularity, the method is widely criticised. Most researchers will agree that conceptual research is a very accurate and useful method for analysis. However, the method should be used along with other research analyses and methods to get results that are more accurate.

BOX 4.1: 'Didactics of ICT in Secondary Education'

Rationale of the study

The ICT subject curriculum from 2006 is intended to improve the content of the subject, on the one hand, and to enhance the teaching and learning of ICT, on the other hand. Unfortunately, the curriculum is restricted to what is to be taught (content) and makes few suggestions as to what didactical skills teachers need to acquire in order to teach ICT as a school subject. Hence, despite a clear specification of content, ICT lacks a disciplined approach to didactics. The lack of a coherent framework for the didactics of ICT makes it difficult to identify which didactical skills teachers need to acquire in secondary schools, which learning strategies are efficient to construct knowledge and which assessment approaches are adequate to assess the students' learning.

Objective

The main objective of this research is to investigate the effectiveness of the framework and the factors influencing its implementation in secondary schools. The research uses design-based research framework for designing, implementing and evaluating the framework through successive cycles of experimentation.

This research work is situated within teacher education. It is aimed at analysing the deficiencies of teaching and learning ICT, designing a solution based on a conceptual framework, and implementing and evaluating the framework in secondary schools. The research involves trainee teachers, school teachers and school students, as well as a number of influencing factors. The author argues that given the complexity of these settings, design-based research is one of the most appropriate paradigms to investigate teaching and learning processes in the field of ICT in secondary schools. The study is organised in four steps:

- It begins with the analysis of the problems and deficiencies of current educational practice in the field of ICT. A critical literature review is conducted, thereby generating the formulation of research questions.
- It continues with the design of a framework for the didactics of ICT, which is used to overcome the deficiencies of current practice. The framework supports the designers' work, forming the foundation for implementation and evaluation.
- Then, an attempt at implementing the suggested framework is performed, using multiple methods for collecting empirical data, for example, formal and informal discussions, project report analysis, observations, field notes and so on.
- Finally, the implementation of the framework is evaluated. The evaluation is concerned with the systematic analysis of the data collected and critical reflection according to the features of the framework and the research questions.

Research questions

The research purpose of the work is threefold:

- To analyse trainee teachers' uses of the framework for the didactics of ICT
- To explore personal and contextual factors that affect the implementation of the framework
- To identify the implications of the framework for the teaching and learning of ICT in secondary schools

Consequently, the research focuses on the following research questions:

- How do trainee teachers use the framework in their teaching of ICT in secondary schools?
- How do trainee teachers' personal perceptions of ICT affect the use of the framework in secondary schools?
- What are the contextual factors that emerge as influential on trainee teachers' uses of the framework?
- What are the implications of the framework for the teaching and learning of ICT in secondary schools?

Methods of data collection

In an attempt to provide a consistent evaluation of the experimentations, this research uses the qualitative paradigm, mostly because the experiments focus on trainee teachers' teaching and learning experiences based on their perceptions, conceptions, motivations, expectations, thoughts, feelings and actions in classroom settings within the school context. Accordingly, particular attention is predominantly devoted to three types of evaluation methods and their combination to a multi-strategy of data collection:

- Trainee teachers' feedback on their experiences with the framework for the didactics of ICT and associated formal discussions during oral exams, informal discussions with the trainee teachers over a three-month time period and associated data that were collected from short conversations

(Continued)

(Continued)

as well as informal discussions with school teachers (the supervisors of trainee teachers in schools) during meetings organised by the teacher education institution.

- Analysis and evaluation of trainee teachers' written project reports and associated educational materials that they produced during the experiments, such as teaching practice reports, instructional plans, information about school context, ICT topics and exercises, computer lab assignments, role of school teachers during the experimentations, school students' assessments of teaching sessions by means of survey questionnaires and so on.
- Field notes and observations of trainee teachers' activities in classrooms during their teaching practice in secondary schools; informal discussions with school students and school teachers, who participated in the teaching sessions performed by trainee teachers.

The collection of data is guided by the conceptual framework and associated research questions. The data were organised and categorised according to four issues:

- The trainee teachers' use of the framework for the didactics of ICT
- Trainee teachers' personal conceptions about ICT and related issues
- Contextual factors that affect the application of the framework and
- Implications of the framework for the teaching and learning of ICT

Participants

The experiments, which were performed over a five-year period, are based on three cohorts of participants: trainee teachers, school teachers and school students between 16 and 18 years as given in the following table.

	Trainee Teachers	School Teachers	School Students
Academic Year 2004–2005	9	11	191
Academic Year 2005–2006	4	4	65
Academic Year 2006–2007	4	4	101
Academic Year 2007–2008	4	4	75
Academic Year 2008–2009	1	1	25

Finally, the study concludes and suggests research findings, recommendations and policy implications.

Source: Adapted from Hadjerrouit.

Conceptual research is based on the theory that describes the phenomenon being studied. For instance, what causes disease? How can we explain and describe the motions of the planets? What are the building blocks of matter? The conceptual researcher sits at his or her desk with pen in hand and tries to solve these problems by reflecting on them. The researcher does not experiment, rather he or she uses enormous amount of data dealing with observations of other researchers to make sense. Conceptual research requires the use of brain not the hands. The method is generally used by philosophers and thinkers to develop new concepts or to reinterpret the existing ones.

Conceptual research is a theoretical structure of assumptions, principles and rules that holds together the ideas comprising a broad concept. It involves speculating about connections that have yet to be confirmed with intervention research or descriptive studies. Conceptual analyses are often developed around a review of the research literature related to the concept under consideration.

Empirical Research

In an empirical research study, the researcher uses his or her experience and direct observation, or he experiments alone without due regard for system and theory. Without these basic principles and properties, a research may not be classified as empirical. Empirical research is data-based research, and researchers can verify its conclusions by observation and experiment. To some extent, it is experimental type of research. In other words, empirical research enables the researcher to answer his or her own questions with corresponding evidence. The evidence (proof) contained in empirical research may be quantitative or qualitative in nature. The data the researchers obtain from the results from observation or experiments may either be interpreted with a quantitative value or qualitative property. Further, empirical research methods help integrating research and practice.

> 'Empirical research is a way of gaining knowledge by means of direct and indirect observation or experience' (Saunders, Lewis and Thornhill 2009).
>
> 'Empirical research is a research based on experimentation or observation (evidence). It is based on observed and measured phenomena and derives knowledge from actual experience rather than from theory or belief' (Johnson 2005).

Empirical research is data-based research. It produces conclusions that can be verified by observation or experiment. An essential aspect of an empirical research process is to obtain the first-hand facts, their sources and to organise certain things to stimulate the production of desired information. It is vital for the researcher to formulate first a working hypothesis for the expected or probable results of the research analysis, that is, the researcher should work consistently and seriously to obtain enough facts (information) with a view to prove or disprove his or her research hypothesis. A key feature of empirical research design is, thus, the experimenter's control over the variables under study and his or her deliberate manipulation of one of these variables to study and analyse its effects. The method is appropriate when the aim of the research study is to prove that certain variables affect other variables in one way or the other. Evidence gathered through experiments or empirical studies is considered to be the most powerful support possible for a given hypothesis (Shavelson and Towne 2002).

Empirical research is conducted to test a hypothesis. By formulating a hypothesis or theory, a given concern (problem) is tested or experimented to come up with an outcome that is fully supported by reliable and accurate data or evidence. But, for scientific research to be accepted as factual or accurate, the activities involved must not rely on basic observation and instead focus on testing a hypothesis through experimentations (Klein 2012).

The advantages of using an empirical research process are numerous. An empirical research process:

- Goes beyond simply reporting observations
- Promotes environment for improved understanding
- Combines extensive research with detailed 'case study'

- Proves relevancy of theory by working in a real-world environment (context)
- Enables the researcher better understand and respond more appropriately to dynamics of situations
- Provides respect to contextual differences
- Helps to build upon what is already known
- Provides opportunity to meet standards of professional research

Empirical research is organised in five distinct stages: observation, induction, deduction, testing and evaluation (Figure 4.1). The observation stage involves collecting and organising empirical facts to form the hypothesis. The induction stage involves the process of forming the hypothesis. A detailed analysis of consequences with newly gained empirical data is carried out at the deduction stage. At the testing stage, the researcher tests the hypothesis with new empirical data. Finally, at the evaluation stage, the researcher performs a systematic evaluation of outcome of testing.

For undertaking an empirical research, it is critical that the researcher should have at least completed a course in statistics and have been exposed to the following concepts.

- Random variables
- Population and samples
- Data tables (rows = sample units and column = variables)
- Summary statistics: mean median, variance, covariance and correlation
- Graphs: box plots, bar charts, histograms, scatter plots
- Inference: standard errors, confidence intervals, hypothesis tests and so on
- Models: bivariate regression, perhaps ANOVA

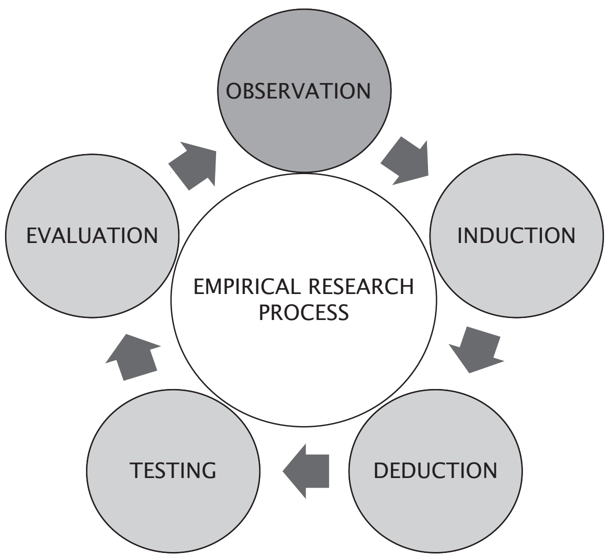


FIGURE 4.1 Process of Empirical Research

Source: Authors.

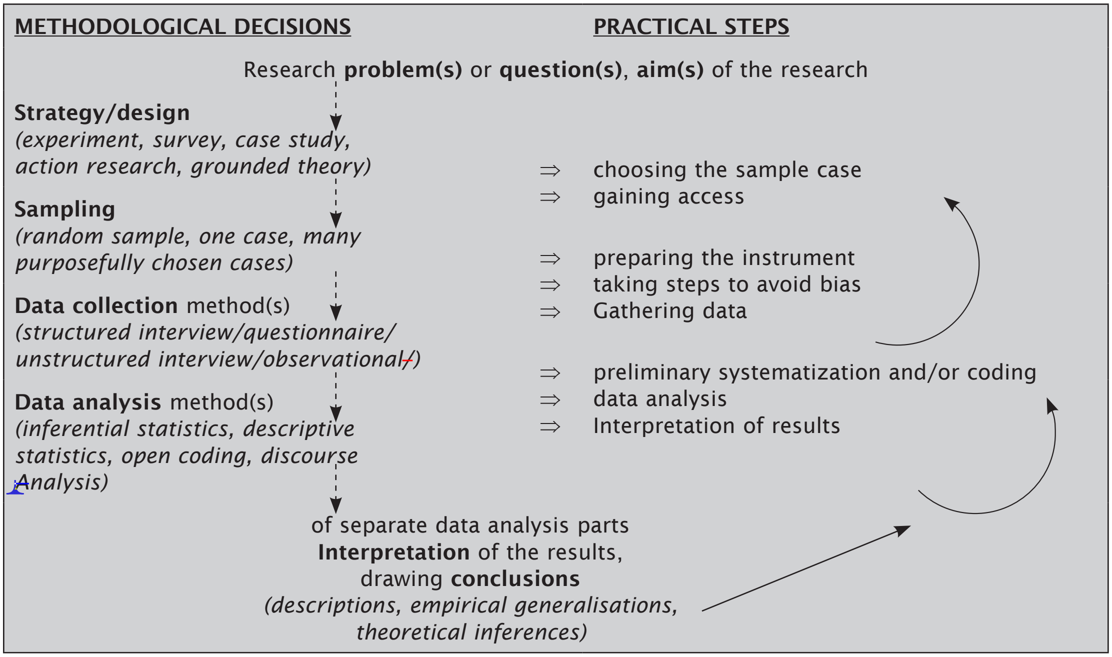


FIGURE 4.2 Methodological Decisions and Practical Steps in an Empirical Research Study

Source: Adapted from Niglas (2001).

Figure 4.2 highlights methodological decisions and practical steps to be followed in an empirical research study.

Researchers make an extensive use of various statistical methods and formulas such as uncertainty coefficient, regression, *t*-test, chi-square and different types of ANOVA (analysis of variance) to form logical and valid conclusion. It should, however, be remembered that any of these statistical formulas do not produce proof and can only 'support' a hypothesis, 'reject' it or do neither.

In the realm of education, decision-makers may want to know if, for instance, pupils of a certain grade at primary level of education will learn faster according to the amount of a teacher's contact hours with pupils. This question may then form part of the hypothesis that a teacher's contact hours with children increase the learning achievement of children. The researcher will test this hypothesis to come up with results. The results of this sample experiment shall then be analysed to come up with a final answer, conclusion or generalisation (Box 4.2).

Stevenson (2014) argues,

> The modern scientific method is really a combination of empirical and conceptual research. Using known experimental data, the researcher formulates a working hypothesis to explain and describe some aspect of nature. He then takes up new experiments designed to test predictions of the theory, to support it or disprove it. Random screening of myriad possibilities is still valuable.

Empirical research is, thus, a research process based on experimentation or observation (evidence). Such research is used to test a hypothesis. In this type of research, most information is gained by

BOX 4.2: 'Learning from Leadership: Investigating the Links to Improved Student Learning'

Purposes

Education is widely held to be crucial for the survival and success of individuals and countries in the emerging global environment. US politicians of all stripes have placed education at the centre of their political platforms, and education has been at the centre of many European and Asian policy agendas. Comparable agreement is also evident about the contributions of leadership to the implementation of virtually all initiatives aimed at improving student learning and the quality of schools. It is, therefore, difficult to imagine a focus for research with greater social justification than research about successful educational leadership. That is the broad focus for this study to identify the nature of successful educational leadership and to better understand how such leadership can improve educational practices and student learning.

More specifically, the study sought to do the following:

- Identify state, district and school leadership practices that directly or indirectly foster the improvement of educational practices and student learning.
- Clarify how successful leadership practices directly and indirectly influence the quality of teaching and learning.
- Determine the extent to which individuals and groups at state, district, school and classroom levels possess the will and skill required to improve student learning and the extent to which their work settings allow and encourage them to act on those capacities and motivations.
- Describe the ways in which, and the success with which, individuals and groups at the state, district, school and classroom levels help others to acquire the will and skill required to improve student learning.
- Identify the leadership and workplace characteristics of districts and schools that encourage the values, capacities and use of practices that improve student learning.

Significant features of the research

The noteworthy features of this study, as against other educational leadership studies, include the size of the database, the use of multiple theoretical and methodological approaches to the research and the comprehensive sources of leadership examined.

Size of the database

Data were collected from a wide range of respondents in 9 states, 43 school districts and 180 elementary, middle and secondary schools. At the state level, the study conducted interviews with legislators, stakeholders and members of state education agencies. In districts, the study interviewed senior district leaders, elected board members, representatives of the media and other informants. The study used survey instruments and interviews with teachers and administrators, and conducted classroom observations with most of the teachers the study interviewed. These efforts yielded, by the end of the research project, survey data from a total of 8,391 teachers and 471 school administrators; interview data from 581 teachers and administrators, 304 district level informants and 124 state personnel and observational data from 312 classrooms. Finally, the study obtained student achievement data for literacy and mathematics in elementary and secondary grades.

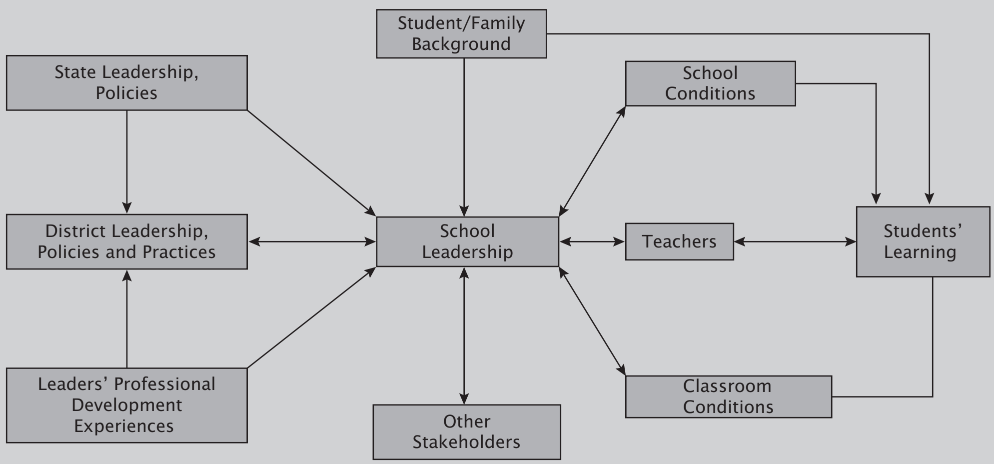
Multiple methodological approaches

The study used qualitative and quantitative methods to gain certain advantages associated with multiple methods research. The advantages typically include rich opportunity for cross-validating and cross-fertilising procedures, findings and theories.

Framework guiding the study

According to information summarised in the following figure, features of state and district policies, practices and other characteristics interact with one another and exert an influence on what school leaders do. These features also influence conditions in schools, classrooms and the professional community of teachers. Other stakeholder groups, including the media, unions, professional associations, and community and business groups also influence school leadership practices. And, of course, leaders are influenced by their own professional learning experiences and by student and family backgrounds. The study examines all kinds of complex relationships among these variables.

Finally, the study demonstrates that leaders, to be successful, need to be highly sensitive to the contexts in which they work. From one perspective, such contexts moderate (enhance or mute) the influence of any given set of leadership practices. From a more practical perspective, different contexts call for quite different enactments of the same basic set of successful leadership practices. Based on the findings, it synthesises implications for policy and practice.



Leadership Influences on Student Learning

Content

Part 1: What Do School Leaders Do to Improve Student's Achievement?
- 1.1. Collective Leadership Effects on Teachers and Students
- 1.2. Shared Leadership: Effects on Teachers and Students of Principals and Teachers Leading Together

(Continued)

(Continued)

1.3. Patterns of Distributed Leadership by Principals: Sources, Beliefs, Interactions and Influences
1.4. Leadership Practices Considered Instructionally Helpful by High-performing Principals and Teachers
1.5. Instructional Leadership: Elementary Versus Secondary Principal and Teacher Interactions and Student Outcomes
1.6. Poverty, Size, Level and Location: The Influence of Context Variables on What Leaders Do and What They Accomplish
1.7. A Synthesis of Implications for Policy and Practice about School Leadership

Part 2: Districts and Their leaders: How They Foster School Improvement and Student Learning
2.1. How Districts Harness Family and Community Energies for School Improvement? Principals' Efficacy: A Key to District Effects on Schools and Students
2.2. How Districts Build Principals' Sense of Efficacy for School Improvement
2.3. Ensuring Productive Leadership Succession
2.4. Data Use in Districts and Schools: Findings and Limitations
2.5. District Approaches to Improving Teaching and Learning
2.6. A Synthesis of Implications for Policy and Practice about District Leadership

Part 3: State Leadership and Relationships with Districts
3.1. State Political Cultures and Policy Leadership
3.2. The Changing Leadership Role of State Education Agencies
3.3. District and School Responses to State Leadership
3.4. State Leadership for School Improvement: A Synthesis of Implications for Policy and Practice

Conclusion

Source: Adapted from Louis et al.

experience, observation or experiment. The word 'empirical' refers to the use of working hypothesis that can be tested using observation and experiment. Experiments and observation produce empirical data.

Micro-genetic Research

Micro-genetic research is a product of Jean Piaget's research in genetic epistemology (the theory of knowledge, especially with regard to its methods, validity and scope). The renowned Swiss psychologist used this method to find out how knowledge grows.

According to Kuhn (1995), one of the major methodological flaws of traditional research designs for analysing learning change processes is that they fail to directly observe change while it actually occurs. For a long time, developmental psychologists devoted themselves to understand how children's knowledge and understanding changes, but they did not directly examine the process by means of which these changes occur. Instead, the typical method was to study what children of one age understand and compare to what children a few years older understand. Researchers then made inferences about what developed between the two ages. But clearly these inferences were indirect. No one had observed the change occurring within the individual child.

Kuhn and Phelps (1982) in in their seminal research work argued that cross-sectional and longitudinal designs allow researchers to view only the products, never the processes, associated with change. In both designs, children's target behaviour is essentially compared at various ages—cross-sectional studies basically compare different children's behaviour, while longitudinal studies, in contrast, compare the same children's behaviour at various ages. More precisely, cross-sectional studies produce a 'snapshot' of a population at a particular point in time. These studies look at people who differ on one key characteristic at one specific time. The design generates feedback on target behaviour's characteristics in large group settings at various ages.

In contrast, they further argue that longitudinal designs analyse the duration of social phenomena and provide critical feedback on changes within cases, rendering a comparison of these changes across various cases feasible. Hence,

> [L]ongitudinal studies enhance the probability that researchers will derive insight on multiple individuals' stable/unstable patterns over time. Nonetheless, longitudinal designs are frequently based on a few observations that are collected over greatly spaced intervals because of their time-consuming nature. In addition, this design potentially conceals intra-individual variability, so essential for grasping specifically how change occurs. Another limitation of this design is its lengthy time intervals between observations are too long to apprehend the progressive process of change. (Kuhn 1995)

Thus, the information derived from this design is analogous to that of snapshots occurring across wide intervals of time rather than to the continual flow of information associated with movies (Calais 2008; Siegler and Crowley 1991). Consequently, transitional behavioural patterns that evolve over short periods of speedy change, critical for ascertaining how, when and why domain specific transitions occur, remain obscured due to a snapshot approach (Lavelli et al. 2005).

> 'Micro-genetic research examines change as it occurs, thus attempting to identify and explain its underlying mechanisms. It is an intensive observation of the same individual over a long period of time; this could be for many weeks or even months. It examines both large-scale and small-scale changes in learners' use of strategies over time. Data can be analysed via either quantitative or qualitative methods, depending on the types of data that are collected' (Kuhn 1995).

As noted by Chinn (2006), most educational research using a micro-genetic approach has examined learners' usage of cognitive strategies (e.g., problem-solving). Micro-genetic studies are time-consuming and can be expensive. But they also can provide researchers with rich and detailed information concerning cognitive processes in learners. An example of a micro-genetic study would be an examination of a kindergartener's strategy usage in solving simple addition problems over a three-month period.

Micro-genetic analysis studies the process of learning and achievement of competency by observing the performance of people through time (Box 4.3). This method seeks to notice subtle changes in students that would go unnoticed by analysis methods with less frequent assessment intervals. Teachers, managers and researchers who use micro-genetic analysis seek to discover how people learn and the efficiency of the teaching methods used. In other words, this method permits the researcher to closely observe 'process of change', instead of 'products'.

BOX 4.3: 'The Development of Organisational Skills'

Aim of the study

The aim of this study was to examine the performance of Grade 5 children in the reproduction of the content of a new text—directly after they had read it (immediate recall), and one week later (delayed recall)—and to investigate the relationship between performance, self-reported memory strategies and working memory capacity (WMC).

Methods

Participants

Thirty-three boys and 33 girls participated in the study. All were in Grade 5 (age range from 11 years 4 months to 12 years 2 months) within the Swedish educational system and were drawn from three public compulsory schools in northern Swedish municipalities. All children gave their consent to participate, and informed consent was obtained from the parents. No one declined to participate. Of the 66 children, 65 took part in all measurements, and all showed age-adequate writing and reading skills. One child failed to complete the listening span and delayed text recall tasks and those partial results were, therefore, excluded from the analysis.

Instruments and procedures

The children were given instructions before each task, offered an opportunity to practice and encouraged to ask questions to prevent misunderstanding.

Text memory

For measurement of text memory, the children were asked to read a standardised text that was new to them with the illustrations removed. The text conformed to a nationally accepted assessment tool of appropriate textbooks for children in Grade 5; it was a 193-word article about a man who worked as a painter in Sweden in the fourteenth century.

The children were asked to read the article with the intention of later reproducing the content. When finished, they were instructed to turn over the paper; the reading speed was then noted by one of the authors. Text reading and the subsequent text recall tasks were conducted in three separate classes with all children present. The memory assessment was based on how well the participants were able to reproduce the content on two occasions. The first occasion was immediately after they had read the text (immediate recall); the second was one week later (delayed recall). On both occasions, they were given instructions to reproduce all details by writing down all that they could remember from the text without worrying about grammar or spelling.

The analysis was based on a system in which the article is divided into chapters consisting of 3 to 10 words depending on the content. Exact reproduction of a chapter scored 1 point, partial reproduction scored 0.5 points and one word or nothing reproduced from the chapter scored 0.

Memory strategies

Questions about the memory strategies that the participants employed during encoding for later reproduction of the content were asked individually in a separate room after the immediate text recall task.

Statistical analysis

All analyses were done with SPSS 18. To evaluate the frequency distribution of the self-reported strategies (no strategy, repetition, visualisation, combination), denoted as strategies, a chi-square analysis of goodness of fit was conducted. Test of linearity assumption between the independent and dependent variables was also calculated. The analysis of skewness and kurtosis was taken to measure reading speed.

Results

The results revealed that more complex strategies were associated with better performances, and that children with high WMC outperformed children with lower WMC in immediate and delayed text recall tasks. Hierarchical regression analyses showed that memory strategy and WMC were the strongest predictors for both immediate and delayed recall tasks. It is argued that self-reported memory strategies are possible to be used as estimates of strategy proficiency. The awareness of the importance of memory strategies and children's WMC in education were further discussed.

Source: Jonsson.

The whole concept of micro-genetic approach rests on the following three critical principles:

- Observations must span a known period of change.
- Density of observations must be high in comparison with the rate of the change.
- Observations are analysed intensively to establish the process that gave rise to them.

Micro-genetic research design helps the researcher to pinpoint different groups of people which may require different treatment or intervention styles. The approach can be used to answer questions that cannot be responded by other approaches. Micro-genetic studies tell us what children know and how they get there.

Most educational research using a micro-genetic approach has examined learners' usage of cognitive strategies (e.g., problem-solving). Micro-genetic studies are time-consuming and expensive, but they can also provide researchers with rich and detailed information concerning cognitive processes in learners. An example of a micro-genetic study would be an examination of a kindergartener's strategy usage in solving simple addition problems over a three-month period.

In recent years, micro-genetic designs are increasingly and widely used to investigate a range of different domains. Some of these domains include early emotional development, mother–infant communication, motor development, early language development, social writing, attention, memory, young children's problem-solving strategies and the effects of instructional procedures. In the field of education, micro-genetic analysis brings more rigour to our teaching techniques and helps us assess our successes and failures as an education provider.

In summary, 'micro-genetic designs, in contrast to traditional designs, not only allow researchers to directly observe change processes but also to observe short-lived transitional behaviours that would otherwise typically remain undetected within analyses of a more aggregated nature' (Postlethwaite 2005). In addressing the topic of how change evolves, micro-genetic designs reflect facets of change along both quantitative and qualitative dimensions, while simultaneously elucidating the nature of these transitional states. Consequently, micro-genetic designs facilitate the investigation of intra-individual variability, that is, individual behaviours' stability and instability across time and varying conditions. The ability of micro-genetic designs to facilitate researchers' ability to identify conditions most conducive for change to occur also enables them to formulate hypotheses regarding potential parameters associated with change, and to conduct micro-genetic experimental studies to verify their hypotheses (Lavelli et al. 2005). According to Siegler (2006), detailed analysis of which mechanisms are associated with shorter versus longer time periods may signify that differences currently evident are actually more substantial. In articulating our understanding of how change occurs, micro-genetic studies also enable us to both explain and describe such processes.

Single-subject Research

In a single-subject research, the investigation is directed exclusively for one participant or many participants viewed as one group, that is, the method examines a variable at a baseline stage (prior to the start of an intervention) and then later examines how this variable changes at different time intervals, as an intervention is introduced. In this type of research, we do not use control or comparison groups. Here our particular interest is to see whether or not patterns replicate over time within the same subject; in addition, we also examine whether or not similar patterns can be generated in new subjects.

Single-subject studies are particularly common in the special education literature. The methodology, however, can be used in other areas of educational research as well. As an example, a single-subject study is best suited to study an examination of the effect of classical music on the ability of a learning-disabled child to solve single-digit addition problems. There are two important prerequisites for this example, namely, an assessment of the child's baseline addition skills followed by the measurement of the student's skills in the presence of music. You will alternately start and stop the music while assessing continuously the student's problem-solving skills (Box 4.4).

BOX 4.4: 'A Study on Foreign Families with Children with Disabilities in Manila, Philippines: Challenges in Local Schools'

This study focuses on the situations and challenges of foreign parents or immigrants in the Philippines. Qualitative interviews and surveys were carried out. Parents' levels of satisfaction and knowledge about the special education programmes were measured. The main respondents are foreign families with children with disabilities enrolled in Philippine schools. Qualitative interviews were carried out with parents, medical staff and special education teachers to explore the current status and the dilemma that immigrants have been facing. Survey was conducted to foreign parents. Mothers were the respondents for this study. Results indicated that the inclusion of a foreign child with disability into a special education school or mainstream school in the Philippines is a challenging and dynamic process that starts from the very moment that the Filipino teacher accepts the foreign student in the classroom. This concept is termed as 'troubling silence' in this paper. Factors discussed include the underlying values in accepting foreign students, the cultural barriers and the conflicting views of parents from developed and developing nations. Cultural politics enclosing the special education field that privileges certain groups was also discussed. In spite of legislation and the desires of parents, the practice of special education does not always reflect the values of equity and rights. Failure to establish collaborative and trusting relationships between teachers, parents and professionals poses a major challenge and can have a serious impact on the outcomes of integrating foreign children in Philippine schools.

Source: Adapted from Belarga and Nakamura (2010).

> 'Single-subject research is a scientific methodology used to define basic principles of behaviour and establish evidence-based practices' (Heffner 2004).

Heffner (2004) suggests that the two most commonly used single research designs are A-B-A-B (a focus on the individual) and multiple baseline design (a design in which individual is used as his or her own control observation). The following explanation would enable you understand clearly the A-B-A-B research design.

Heffner explains that 'the A-B-A-B design represents an attempt to measure a baseline (the first A), a treatment measurement (the first B), the withdrawal of treatment (the second A), and the re-introduction of treatment (the second B)'. In other words, the A-B-A-B design involves two parts: (a) gathering of baseline information, the application of a treatment and measurement of the effects of this treatment and (b) measurement of a return to baseline or what happens when the treatment is removed and then again applying the treatment and measuring the change.

This clearly suggests that single-subject research is 'experimental' (Horner and Morton 2005) and typically used as an alternative to group designs, particularly with individuals with low incidence of disabilities or in research where high numbers of participants are unavailable. Researchers using single-subject research ideally manipulate only one variable in their attempt to evaluate a functional relation between independent and dependent variables.

In single-subject research, researchers often use more than one individual, but results are examined by using each individual at his or her own control, rather than averaging results of different groups. Comparisons are made on the behaviour of one individual to that same individual at a different point in time.

Single-subject research has an important role to play in identifying and documenting solutions for students (individuals) with disabilities. However, this method needs much more evidence on what works for whom, under what conditions, for which tasks and so on. Although students with disabilities—even those with the same diagnosis—often experience unique needs, solutions may be adaptable in different environments, and knowledge sharing can inform others working on assistive solutions (Horner and Morton 2005).

Some examples of questions that might be answered through the utilisation of a single-subject experiment are:

- What is the impact of a self-assessment strategy on an elementary child's ability to examine strengths and weaknesses in his or her writing?
- What are the effects of oral versus silent reading on the reading fluency of struggling readers?
- How do direct teaching narrative and informational text structures aimed at high school students with learning disabilities affect their narrative and informational text reading comprehension?

Inductive Research

Researchers often use two broad methods of reasoning: the 'deductive' and 'inductive' approaches.

Inductive reasoning starts from specific observations and penetrates to broader generalisations and theories. Informally, researchers sometimes call this as a 'bottom-up' approach. In inductive reasoning,

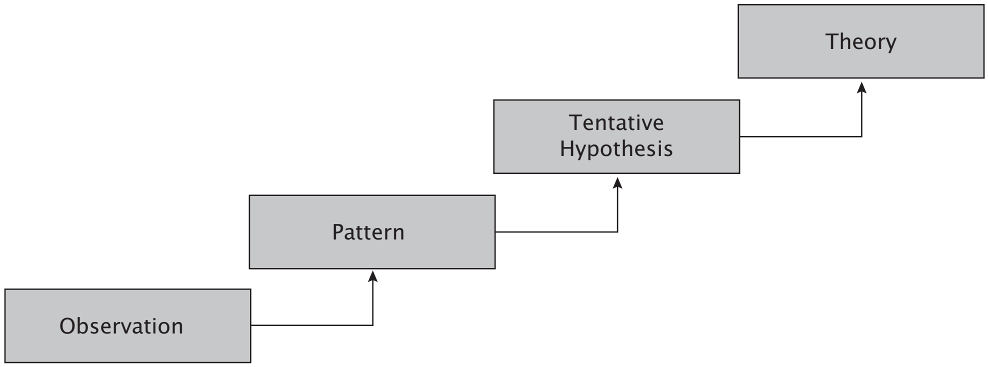


FIGURE 4.3 Inductive Research Process

Source: Authors.

we begin with specific observations and measures to detect patterns and regularities, formulate some tentative hypotheses that we can explore and finally end up developing some general conclusions or theories (Figure 4.3).

> 'Inductive research works from the bottom-up, using the participants' views to build broader themes and generate a theory interconnecting the themes' (Creswell and Plano 2011).

In this approach, the researcher begins with specific observations and measures, then detect patterns and regularities, formulate some tentative hypotheses to explore and finally ends up developing some general conclusions or theories.

Inductive reasoning is basically more open-ended and exploratory, especially at the beginning. It is concerned with the generation of a new theory emerging from the data. Inductive approach to research is generally qualitative research approach. However, there are no set rules as some qualitative studies may have a deductive orientation as well.

Let us consider some examples of inductive reasoning.

Example

You have a very good friend circle. (Premise)
Therefore, you are very good. (Conclusion)

In the above example, the person is being judged. The judgement may not necessarily be true. Even if it is, you can never say if it is temporarily or permanently true.

Example

All the swans that I have seen till date are white in colour.
Therefore, all swans are white.

Example

This year began very well for me.
So this is a lucky year.

In these two examples, there is a sense of a generalised judgement, which may or may not turn out to be true, whereas in deductive reasoning, there is no judgement. The conclusions are mostly true, based on the given situation.

Consider another example. Imagine that a researcher is taking a test in statistical methods. He notices a mosquito kept buzzing around his ear and is distracting him. So as a researcher, he or she wonders if noise distraction has any effect on his or her test taking. In order to study this research question scientifically, he or she sets up an experiment involving 100 students taking a test with some noisemaker in the background. He divides these 100 sample students into five groups and records different level of noise for each group—from quiet to obnoxiously loud.

After all the five groups have completed tests, he or she compares their different scores to see if there is a difference in scores. If the scores continue to decline steadily as the noise increases, then he or she would draw a conclusion that as distraction increases, test scores will generally decrease. If, on the other hand, the majority of their scores increase with the noise, then he or she would conclude that as distraction increases, test scores generally increase—a simple correlation. To reiterate, inductive reasoning draws conclusions from evidence.

Thus, inductive research is the more common way that educational researchers use to conduct experiments (Box 4.5). They have an idea of something to study more in depth. For this to explore, they go and collect data through experiments, observations or surveys. With all of the data in hand, they analyse it to draw conclusions.

BOX 4.5: 'More Than Just a Punctuation Mark: How Boys and Young Men Learn about Menstruation'

Abstract

Parents, peers, schools and the media are the primary contexts for educating young people about sexuality. Yet girls receive more sex education than boys, particularly in terms of menstruation. Lack of attention to how and what boys learn about menstruation has consequences for their private understanding about the biology of reproduction and also for social and cultural ideologies of gendered relationships. In this qualitative study, 23 written narratives from male undergraduates (aged 18–24 years) were analysed using grounded theory methodology to explore how young men perceive their past and present learning about this uniquely female experience. By looking for patterns across all 23 men's narratives, the researchers were able to develop a general theory of how boys and young men learn about this aspect of girls and women's biology. They conclude that sisters play an important role in boys' early understanding of menstruation, that menstruation makes boys feel somewhat separated from girls and that as they enter young adulthood and form romantic relationships, young men develop more mature attitudes about menstruation.

The approach followed for this study is an inductive approach. The researcher began his or her investigation by collecting data that are relevant to the research topic. Once a substantial amount of data were collected, the researcher then took a breather from data collection, stepping back to get a bird's eye view of his data. At this stage, the researcher looked for patterns in the data, working to develop a theory that could explain those patterns. Thus, when the researcher took an inductive approach, he or she started with a set of observations and then he or she moved from those particular experiences to a more general set of propositions about those experiences. In other words, he or she moved from data to theory or from the specific to the general as shown below in Figure A outlining the steps involved with this research study.

(Continued)

(Continued)

Inductive Research

Gather Data	→	Look for Patterns	→	Develop Theory
Specific level of focus		*Analysis*		*General level of focus*

There are many good examples of inductive research, but we will look at just a few here. One fascinating recent study in which the researchers took an inductive approach was Katherine Allen, Christine Kaestle and Abbie Goldberg's study (2011). To understand this process, Allen and her colleagues analysed the written narratives of 23 young men in which the men described how they learned about menstruation, what they thought of it when they first learned about it and what they think of it now. By looking for patterns across all 23 men's narratives, the researchers were able to develop a general theory of how boys and young men learn about this aspect of girls' and women's biology. They conclude that sisters play an important role in boys' early understanding of menstruation, that menstruation makes boys feel somewhat separated from girls, and that as they enter young adulthood and form romantic relationships, young men develop more mature attitudes about menstruation.

In another inductive study (Ferguson et al. 2011), 'Enhancing Empowerment and Leadership among Homeless Youth in Agency and Community Settings: A Grounded Theory Approach', researchers analysed empirical data to better understand how to best meet the needs of young people who are homeless. The authors analysed data from focus groups with 20 young people at a homeless shelter. From these data, they developed a set of recommendations for those interested in applied interventions that serve homeless youth. The researchers also developed hypotheses for people who might wish to conduct further investigation of the topic.

Source: Adapted from Allen, Kaestle and Goldberg.

In the most basic sense, inductive reasoning is a process which typically consists of taking past experiences and using them to explain a present or future circumstance. In inductive reasoning, we extrapolate from experience about what will happen. In inductive reasoning, the researcher relies largely on taking individual instances and compiling them to construct a conclusion. The underlying assumption in inductive reasoning is that known cases can provide information about unknown cases.

Deductive Research or Reasoning

In deductive approach, researchers follow all the steps of the inductive approach described above and reverse their order. They work from the more general information to the more specific. They start with a social theory that looks to them compelling and then test its implications and consequences with data, that is, they move from a more general level to a more specific one. A deductive method to research typically associates with scientific investigation. The researcher analyses the work done by others, reads existing theories on the phenomenon he or she is interested in and then tests hypotheses that emerge from those theories—a confirmation (or not) of our original theories. Figure 4.4 outlines the steps involved with a deductive approach to research.

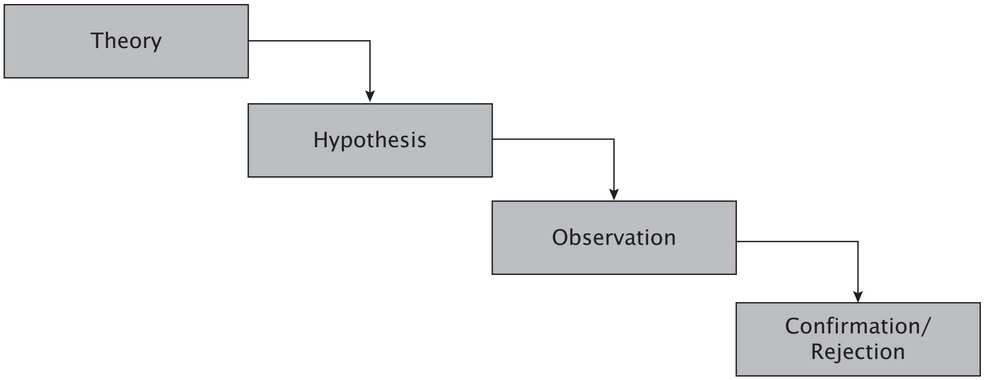


FIGURE 4.4 Deductive Research Process

Source: Authors.

> 'A deductive approach is concerned with developing a hypothesis (or hypotheses) based on existing theory, and then designing a research strategy to test the hypothesis' (Wilson 2013).

The basic principle on which deductive reasoning is based is a well-known mathematical formula:

If A = B (premise)
And B = C (premise)
Then A = C (conclusion)

The conclusion drawn in the above example is an obvious fact in the premise.

Examples of deductive reasoning: given below are a few examples that will help you understand this concept better.

Example

All mangoes are fruits.
All fruits grow on trees.
Therefore, all mangoes grow on trees.

Vicky is a bachelor.
All bachelors are single.
Hence, Vicky is single.

The above examples are valid and sound. Here are a few valid, but unsound examples:

All flight attendants know how to swim.
Ralph knows how to swim.
Hence, Ralph is a flight attendant.

The above conclusion is untrue, because it is not necessary that only flight attendants know how to swim. Absolutely any swimmer can swim.

Sometimes, deductive reasoning can be established even without the help of syllogisms—subtle or crafty arguments consisting of a major and minor premise and conclusion. A common term is present in the two premises but not in the conclusion, which may be invalid (e.g., all dogs are animals; all animals have four legs; therefore, all dogs have four legs). Given below are some examples of deductive reasoning without syllogism.

Statement 1: Every day I go to work.
Statement 2: This journey from my home to my office takes me one hour (premise).
Statement 3: I have to report at eight o' clock in the morning (premise).
Statement 4: So, if I leave my house at seven o' clock in the morning, I will reach office in time (conclusion).

From the second and third statements in the above example, the fourth statement is concluded.

Do not touch the dog; it will bite you.

Through this statement, it is understood that dogs bite (or that a specific dog bites).

There are also chances of deductive reasoning examples that go from specific to general. These are rare and generally have a lot of premises, each of which follows upon the previous one. Given below is one such example:

The members of Khan family are Qamar, Shams and Najam.

Qamar is tall.
Shams is tall.
Najam is tall.

Therefore, all members of the Khan family are tall.

The above examples enable you understand what deductive reasoning is. It may seem simple, but it can go wrong. If a defective reasoning gives a faulty conclusion, it simply means that the given premise was faulty or incorrect. Therefore, the premises used in deductive reasoning are in many ways the most important part of the entire process of deductive reasoning. This is what we have proved in above examples. If our premises are wrong, the entire foundation of the whole line of reasoning will also be faulty and, by inference, the conclusions derived will also be faulty. However, at times, even if the logic is executed properly, there is likelihood for the conclusion to be wrong. Thus, for the purpose of minimising the chances of this mistake happening, it is best not to assume anything and to accept only what has been mentioned.

The difference between deductive and inductive research can best be understood from Table 4.1.

TABLE 4.1 Difference Between Deductive and Inductive Research

Inductive Research	*Deductive Research*
• Is concerned with the generation of new theory emerging from the data	• Is aimed at testing a theory
• Usually uses research questions to narrow the scope of the study	• Usually begins with a hypothesis
• Is usually focused on exploring new phenomena or looking at previously researched phenomena from a different perspective	• Emphasis is generally on causality
• Generally associated with qualitative research **(however, there are no set rules and some qualitative studies may have a deductive orientation)**	• Commonly associated with quantitative research

Source: Authors.

BOX 4.6: Example: Deductive Research

'Negative Classroom Environment Adversely Affects Children's Mental Health'

Children in classrooms with inadequate material resources and children whose teachers feel they are not respected by colleagues exhibit more mental health problems than students in classrooms without these issues, finds a new study in the March issue of the *Journal of Health and Social Behavior*.

The study relies on a nationally representative sample of approximately 10,700 first graders, whose parents and teachers were interviewed.

As part of their study, the authors considered how the classroom environment impacted four components of mental health: learning (e.g., attentiveness), externalising problems (e.g., fights), interpersonal behaviour (e.g., forming friendships) and internalising problems (e.g., anxiety and sadness).

Children in classrooms with inadequate material resources and children whose teachers felt their colleagues did not respect them experienced worse mental health across all four measures.

The material resources ranged from basics such as paper, pencils and heat to child-friendly furnishings, computers, musical instruments and art supplies.

While the study focuses on first graders, the authors expect similar results for older children. But they are not sure.

Source: Adapted from Milkie and Warner (2011).

The first most important point to bear in mind when considering whether to use inductive or deductive approach is the purpose of the research and second is the methods that are best suited to either test a hypothesis, explore a new or emerging area within the discipline or to answer specific questions (Box 4.6).

To conclude, deductive reasoning is a top-down approach, that is, it works from the general to the specific. Empirical research starts with the review of only those theories that have been developed in conjunction with a topic of interest of the researcher. The approach allows the researcher think about research studies that have already been completed. It is on the basis of this typical reflection, he or she develops an idea about extending or adding to that theoretical foundation. From the topical idea, the researcher works to develop a hypothesis. He or she, in the process of conducting a new study, will test this new hypothesis. Specific data that the researcher has collected and analysed for the new study will be the basis of the test of the hypothesis. The specific data will allow either to confirm the hypothesis or not.

Diagnostic Research

Diagnostic research is also called clinical research or laboratory research. In this method, the researcher seeks to control conditions and variables to determine whether a clinical intervention produced the desired (expected) effects or if other factors were responsible for the desired effects.

Like descriptive research design, a diagnostic research design basically takes an interest in detailed description of the phenomenon, group or community. It is concerned with discovering and testing certain variables with respect to their associations or dissociation (Box 4.7). The design is also concerned

BOX 4.7: 'Small Group Math Games Help Children Learn More Math'

Ashley is a graduate student who wanted to conduct a research study to test whether small group math games helped children learn more maths. She designed an experiment to research this question, which included three teachers facilitating small group math games with children. It was important to pick the right games and establish exactly how the games would be facilitated so that each teacher was doing exactly the same thing. These were good reasons to conduct a pilot study before the main study.

For Ashley's graduate research study, she conducted a pilot study to see what small group math game would work best, including how long it would take to play the game. She also established the protocol for facilitation and then practised that protocol with a large number of children. Although the pilot study helped perfect the implementation of the main study, there were still limitations to consider.

Source: Tracy Payne, study.com/academy

Example: Feasibility Study

'Nigerian Virtual Library for Universities and Other Institutions of Higher Education'

The immediate objective of the project is to conduct a detailed feasibility study for the creation of a virtual library for universities and other institutions of higher education in Nigeria with an action plan for its implementation.

The feasibility study will assess what elements are needed and what steps will have to be deployed in order to develop a virtual library for universities and other institutions of higher education in Nigeria. It shall serve as an action-oriented policy and project development tool for the implementation of such a project in Nigeria.

Aspects to be addressed include issues relative to user needs analysis; options for affordable connectivity and access; content provision, creation and management; workflow models for acquisition, cataloguing and dissemination; standardisation for digitisation and cataloguing; training for staff and users; maintenance; project funding; project management and project evaluation.

In particular, the feasibility study will address the following issues:

- **Content**
  What are the needs of the users? What type of information should be made available? How should this information be made available—free or via subscription? Which are the legal issues (copyright, fair use, legal deposit for digital born contents)? Should only Nigerian publications be made available? What languages will be used (English [official], Hausa, Yoruba, Igbo [Ibo], Fulani)? What about foreign monographs, serial publications, online materials (including courseware)? What about digitising already available material, both nationally and internationally? What standards to be used for digitisation and cataloguing of electronic contents?
- **Capacity-building and training**
  Once the system is devised, what sorts of skills will be necessary? Who needs to be trained? For what functions do these people need to be trained (to input data, to upkeep the library, to consult information online, to create alternative methods of delivery [i.e., CD-ROM], to ensure technical maintenance of the equipment…)? What are the current capacity and skills? What are the possibilities of training trainers? What are the possibilities of training users?

Source: UNESCO Portal.

with an existing social problem and its basic nature and cause(s) as well as its treatment. The main objectives of this research design are:

- To diagnose the problem, to accurately specify the characteristics, to determine the frequencies of significant variables and to find out whether certain variances are associated
- To objectively define questions which are to be answered

Diagnostic studies enable the researcher identify a condition, disease, disorder or problem by systematic analysis of the background or history. The method helps them analyse and examine the signs or symptoms, evaluate the research or test results and investigate the assumed or probable causes. Effective prognosis (the likely course of a disease or ailment) is not possible without effective diagnosis.

Pilot Research/Feasibility Study

Pilot research is a research project that is conducted on a limited scale that allows researchers to get a clearer idea of what they want to know and how best they can find it without the expense and effort of a full-fledged study. They are used primarily to try out (validate) survey questions and to refine research hypotheses.

A pilot study is a standard scientific tool for 'soft' research, allowing scientists to conduct a preliminary analysis before committing to a full-blown study or experiment.

A feasibility study is an analysis of the viability of an idea. The feasibility study focuses on helping answer the essential question of 'should we proceed with the proposed project idea?' All activities of the study are directed towards helping answer this question.

Source: Hofstrand and Holz-Clause, What is a Feasibility Study?

Pilot studies last for shorter amounts of time and usually involve a smaller number of participants, sites or organisations (Box 4.7). A pilot study is a standard scientific tool for 'soft' research, allowing scientists to conduct a preliminary analysis before committing to a full-blown study or experiment. For instance, a small chemistry experiment in a college laboratory costs very little and mistakes or validity problems easily rectified as compared to a medical experiment taking samples from thousands of people from across the world.

A pilot study may also be viewed as a feasibility study. A feasibility study is completed to determine whether or not a full study on the topic of research can be accomplished. Feasibility studies are extremely useful and practical when there is concern that a full-scale study may not be possible due to concerns about cost, procedures, personnel and other issues. Pilot studies are not simply exploratory in nature. In fact, researchers design them with a clear purpose of developing some conclusions and pushing an area of research.

A feasibility study, on the other hand, can be defined as an 'analysis or research into the practicality of a proposed plan or method, based on factors like marketplace, competition, available technology, manpower, and financial resources' (Postlethwaite 2005). A feasibility study needs to provide the business or business promoter with an answer to the fundamental question, 'Does this idea have the potential to succeed and will it work?' There is no easy answer to this question, and in order to find one, the educational planner must undertake pertinent research under the appropriate headings.

Educational planners and administrators use feasibility studies in formulating a structured approach for assessing and weighing up the opportunity at hand while simultaneously acting as a safeguard against wasted investment or resources.

Preparation, undertaking and documenting of feasibility studies demand a great deal of work. It is critical that the researcher(s) undertaking these studies dedicate the appropriate time to each stage of the process to ensure that a comprehensive and objective study is conducted.

By conducting a thorough feasibility study, the researcher will be provided with:

- Clear supporting evidence for recommendations to assist with decision-making
- A signpost towards challenges that will need to be addressed
- Valuable information about the target market
- A definitive answer on whether or not to pursue the business idea in question

Pilot studies are mini versions of a full-scale study (also called 'feasibility' studies), as well as the specific pretesting of a particular research instrument such as a questionnaire or interview schedule. Pilot studies are a crucial element of a good study design. We should, however, remember that conducting a pilot study does not guarantee success in the main study, but it certainly increases the likelihood. Pilot studies bring to realisation a range of important functions. They can provide valuable insights for other researchers.

Mixed Methods Research

Researchers use mixed methods to conduct research that involves collecting, analysing and integrating 'quantitative' (e.g., experiments, surveys) and 'qualitative' (e.g., focus groups, interviews) research. The approach is useful when the researcher aims to integrate 'qualitative' and 'quantitative' methods. The method provides a comprehensive understanding of the research problem than provided either of each alone.

We have noticed in Chapter 3 that quantitative data comprises close-ended information such as that found to measure attitudes (e.g., rating scales), behaviours (e.g., observation checklists) and performance instruments. Here the researcher analyses statistically scores collected through questionnaires or contained in checklists to answer research questions or to test hypotheses.

On the other hand, we have also observed in Chapter 2 that qualitative data consists of open-ended information that the researcher usually collects through interviews, focus groups and observations. While analysing the qualitative data (words, text or behaviour), the researcher typically aggregates data into categories of information and highlights the diversity of ideas obtained during the process of data collection.

The researcher, by mixing both quantitative and qualitative research and data, attains an in-depth understanding and corroboration, while offsetting the weaknesses inherent to using each approach independently. One of the typical characteristics of the mixed method is the possibility of triangulation, that is, the use of several means (methods, data sources and researches) to examine and investigate the same phenomenon. The mixed method approach overcomes the limitations of single design.

When to Use the Mixed Method

According to Creswell (2013), the mixed methods are particularly suitable when one wants to:

- Validate or corroborate the results obtained from other methods
- Use one method to inform another method
- Continuously look at a research question from different angles and clarify unexpected findings and/or potential contradictions
- Elaborate, clarify or build on findings from other methods
- Develop a theory about a phenomenon of interest and then test it
- Generalise findings from qualitative research

Advantages of Mixed Methods Research

Using a mixed methods research design has several advantages. It:

- Compares qualitative and quantitative data
- Reflects participants' point of view and offers them a voice
- Fosters scholarly interaction and multidisciplinary team research approach
- Collects rich comprehensive data.

Disadvantages of Mixed Methods Research

Mixed methods impose serious challenges when it comes to implementation, especially when they are used to investigate complex interventions. Several limitations of the methods are as follows:

Limitations

- Complexity of evaluations: Mixed methods studies are complex and difficult to plan.
- Dependence on multidisciplinary team of researchers: The multidisciplinary team of researchers may not have the knowledge of the subject under examination.
- Larger resources: The method is relatively labour intensive and requires larger resources and time compared to those needed to conduct a single method study.

The integration of quantitative and qualitative data in the form of a mixed methods study has great potential to strengthen the rigour and enrich the analysis and findings of our research studies. However, this decision should be determined by the overall purpose of the research.

Summary

From Chapters 2, 3 and 4 on research methods, we have noticed that research methods are generalised and established ways of approaching research questions. Research methods are divided into qualitative

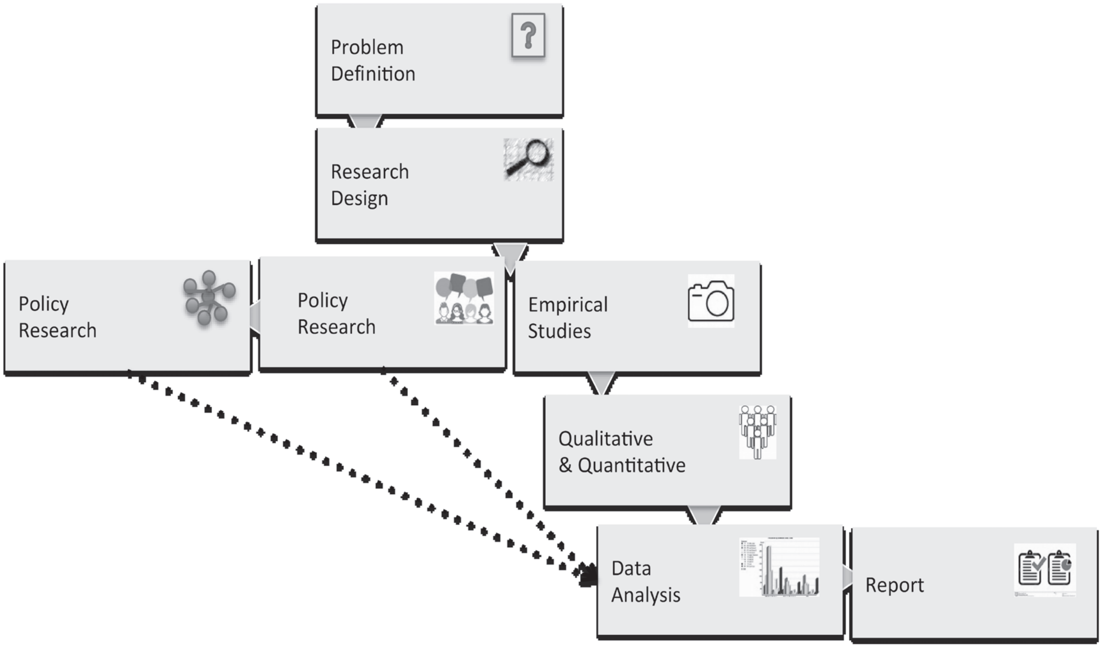


FIGURE 4.5 Specific Research Activities

Source: Authors.

and quantitative approaches and involve the specific study activities of collecting and analysing research data in order to answer the particular research question (Figure 4.5).

All research methods and designs cannot be applied to all research questions. For instance, qualitative methods should be used when the aim is an in-depth, contextual analysis of a phenomenon. These methods are highly useful when our research objective is to answer types of questions involving 'what' and 'who'.

Similarly, quantitative methods should be used when the research aim is to produce generalisable results that show prevalence, incidence, statistical relationships between variables and causation. These methods are highly useful when a researcher has to answer 'why' and 'where' types of questions, but may lack a deep understanding of a phenomenon, particularly if the research area is of a sensitive nature.

A researcher can deploy data collection methods in various study types, including:

Empirical Studies: basically field-based studies for which the researcher collects primary data
Desk Reviews: primarily non-field based studies. The researcher relies heavily on secondary data for his analysis and synthesis to articulate new findings
Research Analysis and Evaluation: either field studies or desk reviews. Here the researcher recommends the best policy or programme option(s) before implementation or to articulate the impact of a policy or programme after implementation

All these data collection, analysis and presentation of results have been discussed in subsequent chapters.

Self-test Exercise

Exercise 4.1: Student failure is a multi-factorial phenomenon, for which there is an imperative to be investigated as it is directly linked to economic development and social stability.

4.1.1. In this context, how would you design a research study to identify parameters which determine student failure in order to reach interpretative conclusions and provide suggestions on a political level?

Exercise 4.2: China just began a major education reform effort that is aimed at reducing the importance of standardised testing in determining school quality and including factors such as student engagement, boredom, anxiety and happiness. It also seeks to cut back on the amount of schoolwork students are given. The approach followed by China in its policy reform process is the opposite of the education reform path, which in recent years has increased the importance of test scores for accountability purposes. Chinese documents explaining the reason for the reform are remarkable, noting that the obsession with test scores 'severely hamper student development as a whole person, stunt their healthy growth and limit opportunity to cultivate social responsibilities, creative spirit and practical abilities in students'.

4.2.1. How would you explain and describe the following phenomenon by undertaking a conceptual research:

A. Severely hamper Chinese students' development as a whole person?
B. Stunt their healthy growth?
C. Factors limiting opportunity to cultivate social responsibilities, creative spirit and practical abilities in students?

Exercise 4.3: Which of the following research questions do you consider appropriate for undertaking a single-subject research? Why?

4.3.1. Whether left-brained students acquire more calculus skills than right-brained students.
4.3.2. Whether a person's dandruff is greater when using shampoo X than when using shampoo Y.
4.3.3. Whether students' solution of novel algebra problems involving 2 unknowns is greater when they have been provided with manipulatives than when they receive procedural instruction.

Exercise 4.4: Mr X is about to park his car on C Street. But he does not want to cause traffic jam. So he looks back to his past experience and tries to predict the probability of a traffic jam in his parking.

4.4.1. How does he make this prediction?

Exercise 4.5: For your own research project entitled 'Are the university and college level courses offered today in the field of business more suited to the industrial age than the information age?'

4.5.1. Which approach you would adopt—an inductive approach, a deductive approach or a combination of the two—to undertake research on the above topic. List three reasons for your choice.

Exercise 4.6: The more you know about kids and their abilities and behaviours, the better you will be able to differentiate or individualise their instruction and make adjustments or adaptations to ensure that they are moving on the right course to achieve standards, or developing confidence, critical thinking and other skills and abilities they need. In order to meet these objectives, how would you answer the following questions?

4.6.1. What diagnostic procedures would you carry out in your classroom work?
4.6.2. How would your teaching alter in the light of diagnostic assessment?
4.6.3. To what extent would you be aware of each pupil's assessment history?
4.6.4. Are there any ways in which you could make greater use of diagnostic procedures to improve learning and teaching?
4.6.5. What are your views on centralised, standardised testing and how can results be most effectively used?

Exercise 4.7: In a research study, five-year-old children have been identified who have not yet mastered number conservation. Four training sessions are being organised to address this shortcoming. The study uses three conditions to present children with conservation problems. Some children are provided with feedback only, regarding their answers. Others, however, are initially requested to explain their reasoning; then they are provided with feedback about their answers. The third group is initially provided with feedback regarding their answers. Then the researcher asks the question 'How do you think I knew that?'

4.7.1. Can you develop a micro-genetic research methodology synopsis to identify the source, path, rate, breadth and variability of cognitive change?

References

Allen, K.R., C.E. Kaestle and A.E. Goldberg. 2011. 'More Than Just a Punctuation Mark: How Boys and Young Men Learn about Menstruation'. *Journal of Family Issues* 32 (2): 129–156.

Belarga, O. and Y. Nakamura. 2010. 'A Study on Foreign Families with Children with Disabilities in Manila, Philippines: Challenges in Local Schools'. *International Journal of Learning* 17 (1): 293.

Calais, G.J. 2008. 'Microgenetic Analysis of Learning: Measuring Change as It Occurs'. *National Forum of Applied Educational Research Journal* 21 (3): 1–7.

Chinn, C.A. 2006. 'The Microgenetic Method: Current Work and Extensions to Classroom Research'. In *Handbook of complementary methods in education research*, edited by J.L. Green, G.Camilli and P. Elmore, 439–456. Washington, DC: American Educational Research Association.

Creswell, J.W. and C.V.L. Plano. 2011. *Designing and Conducting Mixed Methods Research*. Los Angeles, CA: SAGE Publications.

Creswell, J.W. 2013. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, CA: SAGE Publications.

Ferguson, Kristin M. (2011). 'Enhancing Empowerment and Leadership Among Homeless Youth in Agency and Community Settings: A Grounded Theory Approach'. *Child and Adolescent Social Work Journal*, 28, (1): 1–22.

Glaeser, B. and V. Dickson. 2013. *Quantitative and Qualitative Research—Working Together to Inform Education for Students with Learning Disabilities*. Overland Park, KS: Council for Learning Disabilities.

Hadjerrouit, S. 2009. *Didactics of ICT in Secondary Education: Conceptual Issues and Practical Perspectives*. Kristiansand: Norway University of Agder.

Heffner, C.L. 2004. *Research Methods*. Florida: AllPsych and Heffner Media Group, Inc. Available at: https://allpsych.com/researchmethods/ (accessed on 31 January 2018).

Hofstrand, Don and Holz-Clause, Mary. *What is a Feasibility Study?* Available at: https://www.extension.iastate.edu/agdm/wholefarm/html/c5-65.html (accessed on 29 January 2018).

Horner, M.W. and Morton E. O'Kelly. 2005. 'A Combined Cluster and Interaction Model: The Hierarchical Assignment Problem'. *Geographical Analysis* 37 (3): 315–335.

Johnson, A.P. 2005. *A Short Guide to Action Research*. Boston, MA: Pearson Education.

Jonsson, B. 2014. *Self-reported Memory Strategies and their Relationship to Immediate and Delayed Text Recall and Working Memory Capacity*. Umeå: Department of Psychology, Umeå University Sweden.

Klein, S.H. 2012. *European Retail Research*, Vol. 26, Issue 2. Wiesbaden: Springer Gabler.

Kuhn, D. and E. Phelps. 1982. 'The Development of Problem Solving Strategies'. In *Advances in Child Development and Behaviour*, 17th ed., edited by H. Reese: 1–44. New York: Academic Press.

Kuhn, T.S. 1995. *The Structure of Scientific Revolution*. Chicago, IL: University of Chicago Press.

Lavelli, M., A.P.E. Pantoja, H. Hsu, D. Messinger and A. Fogel. 2005. 'Using Microgenetic Designs to Study Change Processes'. In *Handbook of Research Methods in Developmental Science*, edited by D.M. Teti, 40–65. Malden, MA: Blackwell Publishing.

Milkie, M.A. and C.H. Warner. 2011. 'Classroom Learning Environments and the Mental Health of First Grade Children'. *Journal of Health and Social Behavior* 52 (1): 4–22.

Niglas, K. 2001. *Paradigms and Methodology in Educational Research*. Paper presented at the European Conference on Educational Research, Education Line-Lille, 5–8 September 2001, Tallinn Pedagogical University Narva.

Postlethwaite, N. 2005. *Educational Research: Some Basic Concepts and Terminology Module 1*. Paris: UNESCO International Institute for Educational Planning.

Saunders, N.K., P. Lewis and A. Thornhill. 2009. *Research Methods for Business Students*. Harlow: Pearson Education Limited.

Shavelson, R.J. and L. Towne, eds. 2002. *Scientific Research in Education*. Committee on Scientific Principles for Education Research, Centre for Education, Division of Behavioural and Social Sciences and Education, National Research Council. Washington, DC: National Academy Press.

Siegler, R.S. and K. Crowley. 1991. 'The Microgenetic Method: A Direct Means for Studying Cognitive Development'. *American Psychologist* 46 (6): 606–620.

Siegler, R.S. 2006. 'Microgenetic Analyses of Learning'. In *Handbook of Child Psychology*, Vol. 2, edited by W. Damon and R.M. Lerner (series eds.) and D. Kuhn and R.S. Siegler (vol. eds.). Hoboken, NJ: Wiley.

Stevenson, W. 2014. *Conceptual vs. Empirical Research: Which is Better?* Enago. Available at https://www.enago.com/blog/conceptual-vs-empirical-research/ (accessed on 17 November 2017).

Wahlstrom, K.L., K.S. Louis, K. Leithwood and S.E. Anderson. 2010. *Investigating the Links to Improved Student Learning*. Toronto: Final report of Research Findings, Learning from Leadership Project, Ontario Institute for Studies in Education, University of Toronto.

Wilson, J. 2013. *Essentials of Business Research*, 2nd ed. London: SAGE Publications.

5 CHAPTER

Formulating Research Questions and Hypotheses

Introduction

We have noticed in earlier chapters that a meaningful research study follows heavily the process of iteration, that is, the process of returning again and again to the research questions, research methods and data. This process helps the researcher formulate new ideas and also to perform revisions and improvements. Basically, research is a step-by-step process. Thus, the research should be fluid, flexible and open to change. For research to be a successful undertaking, it is imperative to examine and re-examine the possibility of adding new research questions, variables to be ignored and other changes made. At the same time, it is important to examine several study methods at the disposition of a researcher to ensure a comprehensive approach to the research question. As a matter of fact, there is no single formula for developing a successful research study, but it is important to realise that the research process is cyclical and iterative.

The key objective of an investigation is to examine a given issue and suggest solution(s) for its resolve. Research must always be of high quality in order to produce knowledge that is applicable beyond the premise of the research setting, and its implications must go beyond the group that has participated in the research. In other words, the results of our research should have implications for policy formulation and project implementation.

We have explained earlier in Chapters 2, 3 and 4 the several types of research designs and methods that educational researchers and social scientists use to study diverse issues, no matter whether the research is very informal, simple process or it is a formal, somewhat a sophisticated process. No matter what type of process the researcher follows, all research begins with a generalised idea in the form of a research question or a hypothesis.

This chapter describes the several ways for:

- Finding a research topic
- Identifying a research question
- Formulating the hypothesis

Finding a Research Topic

For your research study, the first thing you have to decide is the topic you plan to investigate. Selecting a topic for your research study is not always an easy task. For some of you, the topic may pop into your mind immediately just like that. For many others, however, this may not happen like that. They may need to be more systematic to carefully choose a topic that they wish to explore further. You may find that you have one or two ideas, or you may have many ideas. Before making a final decision about the topic in which you want to invest a lot of time and effort, you should make sure that you have done a great deal of thinking and have read and consulted subject-specific journals and referred to the type of research that was being done on your topic. Your topic can be related to a career aspiration.

The competence to develop a good research topic is an important skill. Your supervisor may assign you a specific topic, but most often supervisors require you to select your own topic of interest. When deciding on a topic, there are a few things that you will need to consider and do seriously:

- Brainstorm for ideas. This will help focus your topic by discussing issues that did not occur to you at first.
- Check the procedures in your institution before selecting the topic.
- Identify and choose a topic that will enable you to read and understand the literature.
- Be sure that your topic is manageable and that material is available.
- Make a list of keywords.
- Be flexible.
- Define and make sure that your topic is a focused research question.
- Look at available research close to your topic and read more about your topic.
- Formulate a thesis statement.

When looking for the topic of your research, many places and people may inspire you. However, the topic you select should be one that can be dealt with in an appropriately academic manner within the means (resources) and time constraints of your research. You have to be sure that:

- The topic will sustain your interest during the months to come
- The selected topic can be approached with analytic distance
- There is enough literature and reference material you can lay hands on easily
- The topic is manageable with the time and resources available

Thus, you should always avoid too broad a topic or one that is overly ambitious. In order to keep your study into manageable limits, the focus of your research should bring the following three together:

- An area of social life.
- A type of research methodology that you would like to use.
- A body of theory that you are interested in exploring.

Bringing all three together is a way of narrowing the focus of the research study into manageable limits.

How Can You Identify the Research Questions?

Upon finalising a focused topic for research, you might be tempted to immediately jump in researching that topic. However, if you do that, you will most likely find yourself surrounded with huge pile of information. You should avoid that and do your research with a goal in mind. Thus, you should always frame your topic as a research problem, that is, the answer of your research question helps solve a problem that your audience cares. In this way, you can ensure both time saving and improved quality and relevance of your research. Thus, formulating a research question is vital in any investigation.

> 'A research question is the methodological point of departure of scholarly research in both the natural and social sciences. The research will answer any question posed. At an undergraduate level, the answer to the research question is the thesis statement. A research question is an answerable inquiry into a specific concern or issue' (*Oxford Dictionary*).

According to Kowalczyk (2015):

> A metaphor for a research project is a house. Your data collection forms the walls, and your hypothesis that guides your data collection is the foundation. So what is the research question? It is the ground beneath the foundation. It is what everything in a research project is built on. Without a question, you cannot have a hypothesis. And without the hypothesis, you would not know how to study what you are interested in.

This metaphor clearly suggests that your research question is the basis of your investigation as it tells you where you are going. It is vital for you to write a good research question. If the foundation of your house is difficult to hold or stand on because it is built on something slippery, wet or slimy, then everything following that will be about correcting that initial issue (foundation) instead of making an awesome home/research project.

The first key step for your research process is the development of solid research questions and hypotheses. The first thing you have to do in this process is to take a preliminary review of the existing literature for your topic. A research question highlights a relationship between two or more variables but phrases the relationship in terms of some question. If your research question is written meaningfully and appropriately, it will guide the implementation of your research project systematically and will provide clear guidelines for the construction of a logical argument. The research question should be a clear, precise and focused question. It should summarise the issue that you will investigate. While formulating the research question, you should:

- Specify your specific concern or issue
- Determine what you want to know about the specific concern or issue
- Ensure that the question is answerable
- Ensure that your question is not too broad or too narrow

TABLE 5.1 Examples of Research Questions

Very narrow: What is the obesity rate in primary schoolchildren?	**Less narrow**: How does the parents' level of education impact child obesity rates among primary schoolchildren?
This is very narrow because it can be explained with a simple statistic. Questions involving answers in 'YES' or 'NO' should typically be avoided.	This question expresses the reasonable (correct) amount of specificity and research would entail the opportunity for an argument to be formed.
Too broad and unfocused: What are the obesity effects of primary schoolchildren in country X?	**More focused**: How does childhood obesity correlate with learning achievement (academic performance) in primary schoolchildren?
This is a very broad question for a given research methodology to investigate. It is also too broad to be discussed in a typical research study.	It is very clear and focused question for which data can be collected, analysed, interpreted and discussed.
Highly objective: How much time do primary schoolchildren spend on physical exercise (activity) per day?	**More subjective**: What is the relationship between physical exercise levels and child obesity?
For this question, the researcher can collect data but it does lend itself to collecting data that can be used to create a valid argument because data will just provide factual information.	This question is more subjective. It is pertinent to the formation of an argument based on the results and analysis of data.
Very simple: How does primary school systems address childhood obesity?	**More complex**: What are the effects of intervention programmes in the primary school system on the rate of childhood obesity among Grade 4 and Grade 6 pupils in Country X?
For this question, the researcher can obtain information without collecting specific and unique data. The answer to this question can be obtained with a simple online research. It would not provide researcher an opportunity for a critical analysis of data.	The question is relatively more complex. However, it is good for investigation and evaluation of a particular intervention programme as well as to form an argument that may be discussed.

Source: Adapted from Kowalczyk (2015).

Considering the information above, Table 5.1 provides examples of flawed research questions as well as questions that are well designed so as to help you understand the importance of a research question in your research study.

A well-researched and thought-out question will, therefore, not only help focus your ideas but will also ensure an appropriate data collection mechanism. Your research question should be such that it enables you determine 'what', 'where', 'when' and 'how' the data will be collected. It should be seen as an important link between the conceptual and logistic aspects of your research plan. Your research question

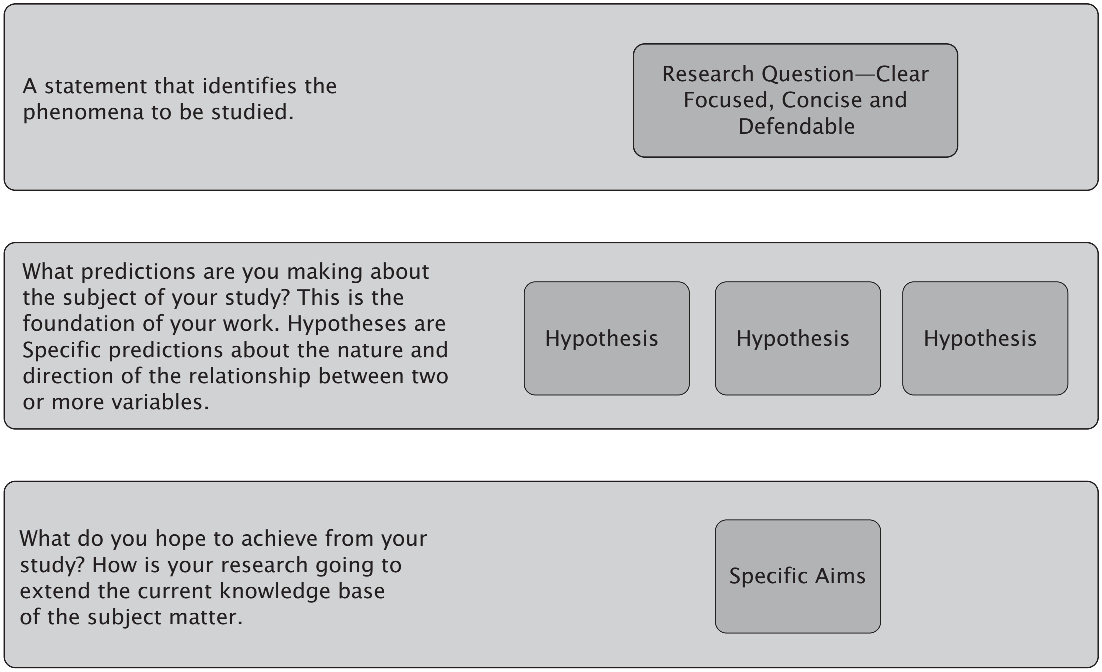


FIGURE 5.1 Research Process Hierarchy: Tips for Formulating a Research Question and Study Design

Source: Adapted from US Army Institute of Surgical Research.

could only be a good research question if it identifies the theoretical construct, transcends the data and has recognisability, significance, robustness and the capacity to surprise (Figure 5.1).

Research questions are generally of two types:

- 'Descriptive questions' involve observations to measure quantity. In these questions, there are no comparison groups/interventions. Purely descriptive questions 'do not' require hypothesis.
- 'Analytical questions', on the other hand, involve comparisons/interventions to test a hypothesis.

Remember that a research question is not the same as a thesis title, research problem, hypothesis or research focus, although they are interrelated and support one another. These differences can be explained as follows:

- 'A research question' summarises the significant issue your research will investigate.
- 'The thesis title' consists of the topic and outcome of a research project.
- 'Research problem' explains the knowledge gap your research will address.
- 'Research hypothesis' is the predicted answer to the research question that can be tested and is based on prior research.
- 'Research focus' specifies the scope or domain of inquiry.

How Can You Define a Hypothesis?

We have seen above the importance and significance of writing an appropriate and relevant research question. In fact, the 'heart and soul' of a research is the research question/problem statement. We also looked at some key components of these research questions/problem statements.

In this section, we will look at the third dimension of a research, namely, our 'guesstimate', if possible, as to what the possible 'answer to the research question' might be. Such statements are called hypotheses.

A hypothesis is one of the key tools in research of advancement of knowledge. It is consistent with existing knowledge and conducive to further enquiry. It is a prediction about the outcome of a study. This prediction is what we believe will hold true for the entire population.

> 'A hypothesis is "a declarative statement that attempts to predict the relationship between two or more variables based on statistical consideration. It is a specific, testable prediction about what a researcher expects to happen in his/her study"' (USAISR 2010).
>
> 'Hypothesis is a formal statement that presents the expected relationship between an independent and dependent variable' (Creswell 1994).
>
> 'A hypothesis is "a tentative assumption made in order to draw out and test its logical or empirical consequences"' (Dictionary definition).

Basically, a hypothesis is a premise. It is an educated guess about a relationship. It is a claim that we want to test or investigate. It provides direction and the way in which we investigate things and statistic. For hypothesis testing, we cannot go into a laboratory and do an experiment but what we can do by undertaking a survey. A hypothesis focuses on the key variables contained in the research question.

Being a prediction, a hypothesis is stated in sentence form. The following example will clarify the difference between a problem statement and its related hypothesis.

Example

Problem
This study is designed to determine the effects of a peer-assisted method of teaching reading, as compared to the conventional method in terms of reading comprehension.

Statement
This problem implies a controlled setting. The 'new method' is the 'treatment' and the conventional is the 'control' (the research method is 'experimental').

Hypothesis
Students taught by the peer-assisted method of teaching reading will score significantly higher on a reading comprehension test than students taught by the conventional method.

It is important to note the key difference between the above two forms (research question/problem statement and its related hypothesis). The research question or problem statement is in 'open-ended' form, while the hypothesis expresses a definite outcome or set of outcomes that the researcher might predict. It, in other words, suggests one key property: 'all hypotheses should possess: namely, they must be in testable form'.

It should be borne in mind that if your hypothesis is not well articulated (written) then it becomes 'non-testable'. Such hypotheses tend to come in two forms. First, you have accidently left out a basis for comparison. For example, consider the hypothesis: 'Students taught by method A will be better readers'.

This hypothesis is non-testable, as we do not know what 'better' means here. 'Better' implies a comparison, but we do not know what is being compared. Note that in above example, peer teaching (Method B) is being compared to 'conventional' method (Method A).

Second, if a hypothesis is based on a researcher's value judgement or his or her opinion, it will again be a non-testable hypothesis. Let us consider the following example: 'All junior high school age boys should be required to take a course in economics'. This is more of a value judgment or opinion. This hypothesis to be testable should be written as: 'Junior high school age boys who have taken a course in economics will exhibit significantly less sex-role-stereotyping behaviour than junior high school age boys who have not taken a course in economics'. This particular hypothesis can be tested, and then, of course, retained or rejected once the study evidence is in.

The discussion presented in this section is limited to the hypothesis-testing procedures applied to a selected batch of the following parameters:

- The population mean μ
- The population proportion, π, for one-sample tests of significance
- The difference of two population means, $\mu_1 - \mu_2$, for independent samples
- The difference of two population proportions, $\pi_1 - \pi_2$, for two-sample tests

How Can You Formulate a Hypothesis?

Hypothesis formulation is a key component of descriptive statistics. It examines the relationships between closely observed variables. In fact, formulating a hypothesis is an important phase that arrives only after a research question, comprising all the variables, has been identified without a single flaw by the researcher or a group of researchers. Figure 5.2 highlights the process of hypothesis formulation.

When you try to come up with a good hypothesis for your education research or experiments, you are required to ask yourself the following questions:

- Is my hypothesis based on my research topic?
- Can I test my hypothesis—verifiable or falsifiable?
- Does my hypothesis include independent and dependent variables?
- Is my hypothesis a moral or ethical question?
- Is my hypothesis too specific or too general?
- Is it a prediction of consequences?
- Is it considered valuable even if proven false?

Before you come up with a specific hypothesis, you should always spend sometime doing background research on your topic. Once you have completed the literature review, start thinking and identifying all potential questions you still have.

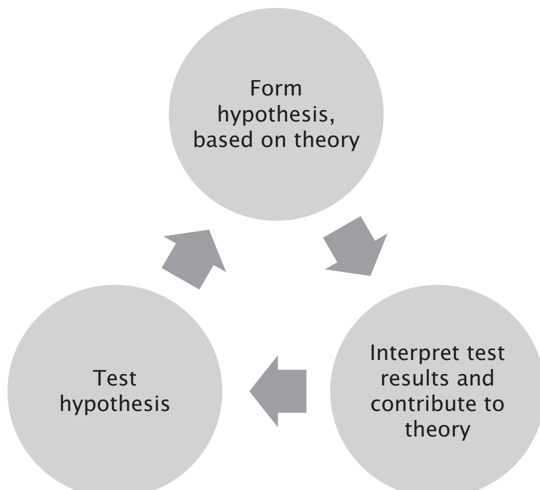


FIGURE 5.2 Hypothesis Formulation Process

Source: Authors.

In short, the hypothesis is directly related to a theory but contains operationally defined variables and is in testable form. Hypotheses allow us to determine, through research, if our theory is correct, that is: 'Does prior work experience result in better grades?' When doing research, we are typically looking for some type of difference or change between two or more groups (Christopher 2017).

How Can You Differentiate Between Null Hypothesis and Alternative Hypothesis Formulation?

When you analyse a data set to determine if two factors might be related, you have to form two hypotheses. The first is the 'null hypothesis' (denoted by H_o), which says: 'Factor A does not affect Factor B'. It is usually the hypothesis that the researcher wants to gather evidence against or is the hypothesis to be tested. The H_o carries the symbol '=' or at certain times bear two symbols such as '≥' (equal or greater than) or '≤' (equal or less than), thus making H_o composite in nature.

On the other hand, the 'alternative hypothesis' (denoted by H_a) tells us: 'Factor A does affect Factor B'. The alternative hypothesis is usually the hypothesis for which you want to gather supporting evidence by way of observation that could be obtained from your sampling experiment. The alternative hypothesis is the opposite of the null hypothesis. For a two-tailed test, the symbol used in the H_a statement is '≠' (not equal) and for a one-tailed test, it is either '<' (less than) or '>' (greater than). The alternative hypothesis carries only 'one symbol' at all times.

Thus, every study has two hypotheses; one stated as a difference between groups and one stated as no difference between groups.

Let us first understand what a 'null hypothesis' is. In simple words, a null hypothesis is the default hypothesis—the currently established value of a parameter. It is empty or zero and statistically denoted with a small '0' (zero). What does this mean? It is the currently accepted value for a parameter. For instance, when we are looking at a parameter of the population, say, the mean of the IQ, we are going to have some currently established value set up by previous studies. We typically call that 'H_0' (H naught), that is, the currently accepted value.

> 'The null hypothesis is a type of hypothesis used in statistics that proposes that no statistical significance exists between the two variables in the hypothesis or in a set of given observations. The null hypothesis is usually the hypothesis that samples observations result purely from chance' (Investopedia).[1]

An inherent assumption of a null hypothesis is that the groups do not differ. For instance, a null hypothesis would state that, 'students with prior work experience get higher grades than students without prior work experience'. It means that the null hypothesis attempts to show that no variation exists between variables, or that a single variable is no different than zero. It is presumed to be true until statistical evidence nullifies it for an alternative hypothesis.

Null hypotheses are used in quantitative analysis to test theories about markets, investing strategies or economies, education and social sciences to decide if the idea is false or true.

There is, on the other hand, another hypothesis typically called the 'alternative hypothesis' or 'research hypothesis'. An alternative hypothesis is a challenge, because it involves the claim to be tested. We call an alternative hypothesis 'H_a'. We call it H_a because it is an alternative.

---

[1] See http://www.investopedia.com/terms/p/p-value.asp (accessed on 5 June 2017).

'An alternative hypothesis states that there is statistical significance between two variables. It is hypothesis the researcher is trying to prove'.

To make the concepts further clear, let us consider the following example.

Example

It is believed that a community secondary school produces graduates who on an average score 80 marks in mathematics. A teacher claims that the school after a given reform in the examination system in the school no longer produces graduates with 80 marks in mathematics.

This teacher, in fact, presents an alternative hypothesis that we are going to test, that is, why it is called the alternative or research hypothesis. How can we write H_0 and H_a? For this particular case, both the null hypothesis and the alternative hypothesis can be formulated as follows:

Null Hypothesis	H_0: $\mu = 80$
Alternative Hypothesis	H_a: $\mu \neq 80$

In our example of peer-assisted method and traditional method of teaching discussed above, the two hypotheses will be as follows:

Null Hypothesis (H_0)
'Students taught by the peer-assisted method of teaching reading will not score significantly higher on a reading comprehension test than students taught by the traditional method'.

Alternative Hypothesis (H_a)
'Students taught by the peer-assisted method of teaching reading will score significantly higher on a reading comprehension test than students taught by the traditional method'.

The null hypothesis is what we test through the use of statistics. Since we are testing the null hypothesis, we can assume that if the null hypothesis is not true then some alternative hypothesis to the null hypothesis must be true. It is the alternative hypothesis.

In order to make your research as specific and authentic as possible, you have to look for one of two outcomes, either the null or the alternative hypothesis. To conclude that there is no difference between the two groups means we are accepting our null hypothesis. If we, however, show that the null is not true then we must reject it and therefore conclude that the alternative hypothesis must be true.

The important thing to note here is that the null hypothesis and alternative hypothesis are mathematically opposite.

How Can You Test a Hypothesis?

Hypothesis testing is a very significant topic of statistics. This technique is associated with inferential statistics. The technique is intensively used by researchers from all sorts of disciplines including psychology, education, marketing and medicine for hypothesis testing or claim about a population being studied.

Hypothesis tests belong to the theory of probability, a specialised field of mathematics. Researchers use probability to quantify how likely it is for an event to occur. Since all inferential statistical methods relate with rare events, which is why probability is used so extensively. The underlying idea here is that while testing a claim, we must differentiate between two different features of an event:

- An event that easily occurs by chance
- An event that is highly unlikely to occur by chance

In case of an event which is highly unlikely to happen or occur, statisticians express this by stating that a rare event has happened or that they started with a false assumption. In other words, the assumption was not true.

What Are the Essentials of Hypothesis Testing?

For introducing the topic of hypothesis testing, let us begin with some important terms frequently used in hypothesis testing. These are given in Table 5.2 together with their main features. These concepts are explained in more details in this tag.

TABLE 5.2 Essentials of Hypothesis Testing

A: BASIC CONCEPTS USED IN HYPOTHESIS TESTING	
Terms	***Main Features***
• Hypothesis	• A claim (assertion) about a population parameter
• Hypothesis Testing	• A procedure to evaluate a hypothesis about the population with actual results obtained from a sample of data • The purpose is to choose between two contradictory alternatives of the value of a population parameter (null hypothesis and alternative hypothesis)
• Null Hypothesis	• The default hypothesis (H_0) • Refers to the value reached with the fewest assumptions. This is often but not always the 'status quo' or currently assumed value • Always contains '=', '≤' or '≥' sign • May or may not be rejected by sample data, cannot be proved
• Alternative Hypothesis	• Complement of the null hypothesis (H_a) • Challenges the current belief by suggesting the null hypothesis parameter is wrong • Never contains '=', '≤' or '≥' sign • May or may not be accepted depending on analysis of sample data • Is generally the hypothesis that the researcher is trying to test, so it is sometimes called the 'test hypothesis'
• Critical Value	• This value of a test statistic creates a line for decision-making—it answers the question of how far is far enough
• Type I Error	• Rejecting a null hypothesis when it is true. Imagine it as a 'false alarm'. Serious type of error • The probability of this is α (the level of significance of the test, chosen by the researcher in advance of data analysis)
• Type II Error	• Maintaining null hypothesis when it is actually false. Imagine it as a 'missed opportunity' • The probability of this is β
• Confidence Coefficient	• The probability of not rejecting H_0 when H_0 is true $(1 - \alpha)$
• Confidence Level	• Confidence coefficient expressed as a percentage: $(1 - \alpha) \times 100\%$

• Power of a Statistical Test	• The probability of rejecting H_0 when H_0 is false $(1 - \beta)$
• Level of Significance	• The probability of rejecting the null hypothesis when it is true (making a Type I error) • Defines the rejection region (RR) of the sampling distribution • Designated by α • Selected by the researcher at the beginning • Used to calculate the critical value(s) of the test
• Z-test of Hypothesis for the Mean (σ known)	Convert sample statistic $\overline{X}$ to a z test statistic
• Steps of Critical Value Hypothesis Testing	• Formulate and state the null (H_0) and alternative (H_a) hypotheses • Specify the desired significance level (α) and sample size (n) • Select the appropriate test statistic distribution (z, t, other) • Calculate the resulting RR using critical values derived from your choice of α in Step 2 (z-crit, t-crit and so on) • Take a random sample and use its data to calculate the test statistic z-stat, t-stat and so on) • Compare values to make the statistical decision and state conclusion in terms of the question
• p–value	• Probability of randomly choosing a sample out of all possible samples of the same size with a resulting test statistic. Also called 'observed level of significance' • Mathematically: $P(\lvert z \rvert \geq \lvert z\text{-test} \rvert)$
• p-value Approach to Testing	• If p-value $\leq \alpha$, reject H_0 • If p-value $\geq \alpha$, do not reject H_0
• Lower Tail Test	H_0: $\mu \geq$ value H_a: $\mu <$ value
• Upper Tail Test	H_0: $\mu \leq$ value H_a: $\mu >$ value

B: STEPS INVOLVED IN THE HYPOTHESIS TESTING PROCEDURE

Step 1	State the null hypothesis, H_o
Step 2	State the alternative hypothesis, H_a
Step 3	Specify the level of significance, α
Step 4	Determine the critical region and the appropriate test statistic
Step 5	Compute the equivalent test statistic of the observed value of the parameter
Step 6	Make your decision either REJECT H_o or ACCEPT H_o

C: SYMBOLS AND TERMINOLOGY USED IN HYPOTHESIS TESTING*

Measure of Characteristic	*Parameter (Population Measure)*	*Statistic (Point Estimator of Parameter)*	*Standard Error of the Sample Statistic*
Mean	μ	$\overline{x}$	$\sigma_{\overline{x}}$ or $s_{\overline{x}}$
Proportion	π	p	σ_p
Difference of 2 mean	$\mu_1 - \mu_2$	$\overline{x}_1 - \overline{x}_2$	$\sigma_{\overline{x}_1} - \sigma_{\overline{x}_2}$
Difference of 2 proportions	$\pi_1 - \pi_2$	$p_1 - p_2$	$\sigma_{p_1 - p_2}$

***Source:** Vasquez (2003).

Hypothesis Testing Methodology

We have noticed above that hypothesis testing typically begins with some theory, claim or assertion about a particular parameter of a population. A key factor that you should always remember that hypothesis testing requires that you state a claim unambiguously. Let us consider the following example.

Example

It is believed that a community secondary school produces graduates who on an average score 80 marks in mathematics. A teacher claims that the school after a given reform in the examination system in the school no longer produces graduates with 80 marks in mathematics.

Our initial hypothesis in this case is that the examination system is working properly, so the mean score is 80 marks, and no corrective action (reform) is needed. Thus, the null hypothesis should be stated as:

$$H_0: \mu = 80$$

Whenever a null hypothesis is specified, an alternative hypothesis is also specified, and it must be true if the null hypothesis is false. The alternative hypothesis, H_a, is the opposite of the null hypothesis, H_0. This could be stated for the above case as:

$$H_a: \mu \neq 80$$

The alternative hypothesis draws the conclusion reached by rejecting the null hypothesis. A null hypothesis is rejected when there is sufficient evidence from the sample information that the null hypothesis is false. In the above case, if the score in math in the sample is sufficiently above or below the expected mean value of 80 marks specified by the school, we reject the null hypothesis in favour of the alternative hypothesis that the mean score is different from 80 marks.

If this is the case, then we have to review the examination system and take any action necessary to correct the problem. If, on the other hand, the null hypothesis is not rejected, we should continue to believe in the status quo, that the process (examination system) is working correctly and therefore no corrective measure is required. In this second circumstance, we have not proven that the process is working correctly. Rather, we have failed to prove that it is working incorrectly, and therefore we continue our belief (although unproven) in the null hypothesis.

The following key points summarise the null and alternative hypotheses:

- The null hypothesis, H_0, highlights the existing status of affairs (status quo) or the feeling being sure that something exists or that something is true.
- The alternative hypothesis, H_a, is the opposite of the null hypothesis. It represents a research claim or specific inference the researcher would like to prove.
- If the researcher rejects the null hypothesis, he should have enough statistical reasoning to prove that the alternative hypothesis is correct.
- If he does not reject the null hypothesis, he has failed to prove the alternative hypothesis.
- The failure to prove the alternative hypothesis, however, does not mean that he has proven the null hypothesis.
- The null hypothesis, $H_{0,}$ always refers to a specified value of the population parameter (such as μ), not a sample statistic (such as $\overline{X}$ or arithmetic means).

- The statement of the null hypothesis always contains an equal sign regarding the specified value of the population parameter (e.g., H_0: = 80).
- The statement of the alternative hypothesis never contains an equal sign regarding the specified value of the population parameter (e.g., H_a: ≠80 marks grams).

It should be noted that the way the alternative hypothesis is formulated depends on the scientific objective at hand. Most examples of alternative hypotheses are 'one-sided'. For instance, you might have come up with a measurable hypothesis that children have a higher IQ if they eat oily fish for a period of time. Then we would set up our hypotheses as:

$$H_0: \mu = \mu_0 \text{ versus } H_a: \mu < \mu_0;$$

where μ_0 is the mean of children not eating oily fish, that is, your alternative hypothesis, H_a, would be 'children who eat oily fish for six months will show a higher IQ increase than children who have not'. And your null hypothesis, H_o, would be: 'children who eat oily fish for six months do not show a higher IQ increase than children who do not'.

On the other hand, if we want to test if a new drug increases the mean survival time for people suffering from a particular type of cancer, then we would set up our hypotheses as:

$$H_0: \mu = \mu_0 \text{ versus } H_a: \mu > \mu_0;$$

where μ_0 is the mean survival time without medication.

Table 5.3 compares the two types of hypotheses.

TABLE 5.3 Comparison Between Null Hypothesis and Alternative Hypothesis

A hypothesis is a claim (assumption about a population parameter):

- **Population mean**
  Example: The mean monthly Internet bill in India is μ = \$ 16.
- **Population proportion**
  Example: The proportion of people in India with Internet connections is π = 0.68

NULL HYPOTHESIS, H_0	**ALTERNATIVE HYPOTHESIS, H_a**
• States the claim or assertion to be tested. Example: The average monthly Internet bill number of children per household in Country X is five (H_0: μ = 5) • It is always about a population parameter, not about a sample statistic (H_0: $\mu = 5$) (~~H_0: $\bar{x} = 5$~~) • Begins with assumption that the null hypothesis is true • Refers to the status quo or historical value • Always contains '=', '≤' or '≥' sign • May or may not be rejected	• Is the opposite of the null hypothesis, for example, the average number of children per household in Country X is not equal to 5 • Challenges the status quo • Never contains '=', '≤' or '≥' sign • It is generally the hypothesis that the researcher is trying to prove

Source: Authors.

The following examples will help you understand further the way the null and alternative hypotheses and level of confidence can be written. Notice we are not doing the testing in these problems. We are just writing the hypothesis.

Example

A World Food Programme (WFP) project officer has stated that the primary school girls participating in the midday meal programme attend school on an average for four days per week. A teacher in the school believes the programme is no longer results attendance of this size and to test this, he or she samples 100 girl students participating in the programme to perform a hypothesis test with 99 per cent level of confidence.

What we want to do here is to write down the H_o and H_a at 99 per cent level of confidence. When you read this problem, make sure that you have read it carefully—you need to see exactly what it is telling you. First thing it says the girls participating in the midday meal programme attend school on an average for four days per week. This is what the WFP project officer believes. But a teacher in the school now believes that the programme no longer results attendance of the order of four days per week. Notice all it really says is that the project officer believes the attendance of four days and the teacher believes that it is no longer four days. It does not say anything about greater than or less than four days. This is important because in some other problems we may have a greater than or less than and in some other we would not. So in this case our null hypothesis would be:

$$H_o: \mu = 4 \text{ days}$$

But when we read the next sentence, this is no longer the case. Thus, our alternative hypothesis would be:

$$H_a: \mu \neq 4 \text{ days}$$

When you are asked to write down the H_o and H_a, this is how you would write.

Here we do not have any greater than or less than signs because the actual problem does not say anything about greater or less than four days. It just says four days.

As we go further, we need to know how many samples do we actually choose to test the hypothesis. We take a sample of 100 girl students participating in the midday meal programme, that is, $n = 100$. The problem also tells us that we want to test it at 99 per cent confidence level. This means that the variable C is 0.99. We know that sometimes we use C and sometimes we use significance level (alpha) where

$$\alpha = 1 - \text{C or } 1 - 0.99 \text{ or } 0.01$$

Note, when C is very high alpha is very low.

Example

Doctors believe that pre-school children sleep on average no longer than 10 hours per day. A researcher believes that they sleep on average longer.

Now the difference in this example and example given above is in the wording. It is very different. Doctors believe that the pre-school children might sleep one hour, two hours and so on but no longer than 10 hours. It means that this is clearly not the case when we have equals or not equals because the wording of the problem is telling 'no longer than 10 hours'. Thus, the null hypothesis would accordingly be written as follows:

$$H_0: \mu \leq = 10 \text{ hours}$$

This means that the population mean is less than or equal to 10 hours. So alternative or challenge hypothesis (H_a) would be:

$$H_a: \mu > 10 \text{ hours}$$

Notice there is no 'equal' sign in the alternative hypothesis here.

Example

A school board claims that at least 70 per cent of students bring phone to school. A teacher believes this number is too high and randomly samples 30 students to test at a level of significance of 0.02.

In this example, we are not really asked to perform the test. What we are dealing with here is using the percentage the school board claims: at least 70 per cent of students bring a phone to class. In above two examples, we have numbers for averages. That is why we use number in these two cases. But in this case, it is actually a proportion. A proportion is just a percentage. Thus, in cases such as these, we construct our hypothesis test without including the mean value (parameter) of the population. The population parameter is expressed in terms of proportion. It is exactly the same concept as in two examples above. Therefore, the null hypothesis the school board claims is at least 70 per cent. So it is a proportion of students in that school.

Thus, *p*:

$$H_o: p >= 0.70$$

The alternative hypothesis will be:

$$H_a: p < 0.70$$

The number of cases that we sampled is 30 students for our experiment. Now in this problem we have not given all the information to solve it because we did not tell you what the results of our sampling were. Using the sample data, it is up to us now to figure out when it becomes true statistically significant if all 30 students answer 'Yes' and what if 19 of them answer 'Yes'. The question here is: How do we choose the threshold beyond which we are going to reject or not to reject the null hypothesis?

Thus, the H_a can be any expression: 'less than', 'greater than', 'not equal' and also 'equal to'.

Thus, there are two types of hypotheses. The 'null hypothesis' (H_0) is what you currently think is true. It usually comes from a previous survey or previous set of data—something that somebody has done in the past. Along comes somebody else with new information and proposes an 'alternative hypothesis' (H_a) that he or she believes is true. This is the hypothesis that you have to investigate.

In conclusion, in any research or investigation, the researcher identifies the research question, formulates the hypothesis and tests it. We have seen above that a hypothesis is some claim that he wants to investigate.

The Critical Value of the Test Statistic

In our example of secondary school graduates above, the null hypothesis is that the mean score in math for all the graduates is 80 marks (the population parameter specified by the school principal). We select a sample of graduates from the list of all graduates, record the marks obtained in math and compute the sample mean. This statistic is an estimate of the corresponding parameter (the population mean μ).

Even if the null hypothesis is true, there is likelihood for statistic $\overline{X}$ (the sample mean) to differ from the value of the parameter (the population mean, μ) because of variation due to sampling. However, it is expected that the sample statistic to be close to the population parameter if the null hypothesis is true. If the sample statistic is close to the population parameter, we have insufficient evidence to reject the null hypothesis. For example, if the sample mean is 79.9, we conclude that the population mean has not changed (i.e., $\overline{X} = 80$) because a sample mean of 79.9 is very close to the hypothesised value of 80. Intuitively, we think that it is likely that we could get a sample mean of 79.9 from a population whose mean is 80.

> 'The "critical value" is a factor used to compute the margin of error.
> Margin of error = Critical value multiplied by standard deviation of the statistic.
> Margin of error = Critical value multiplied by standard error of the statistic.
> When the sampling distribution of the statistic is normal or nearly normal, the critical value can be expressed as a t-score or as a Z-score'.

However, if there is a large difference between the value of the statistic and the hypothesised value of the population parameter, we conclude that the null hypothesis is false. For example, if the sample mean is 70, we conclude that the population mean is not 80 (i.e., $\mu \neq 80$), because the sample mean is very far from the hypothesised value of 80. In such a case, we conclude that it is very unlikely to get a sample mean of 70 if the population mean is really 80. Therefore, it is more logical to conclude that the population mean is not equal to 80. Here we reject the null hypothesis.

We can use normal distribution or t-distribution to help determine whether the null hypothesis is true. In the following paragraphs, we will see how do these tests really work and what does statistical significance actually mean.

Rejection Region and Non-rejection Region of Hypothesis

Once you have collected and analysed data, then the next thing you have to decide is: should we reject H_0 and accept H_a? Or do we have insufficient evidence to reject H_0? When making decisions, there are four possible scenarios and two of them involve errors:

Scenario 1. Accept H_0 when in fact H_0 is true (good decision)
Scenario 2. Reject H_0 when H_0 is false (good decision)
Scenario 3. Reject H_0 when in fact H_0 is true (an error)
Scenario 4. Accept H_0 when in fact H_0 is false (an error)

Based on the above four scenarios, the sampling distribution of the test statistic is divided into two regions: a RR (sometimes called the 'critical region') and a 'NR'. This is shown in Figure 5.3.

The two possible errors (scenarios 3 and 4) are called 'Type I Error' and 'Type II Error', respectively. The types of errors and 'do not reject' and 'reject' scenarios are explained in Table 5.4.

A 'Type I Error' is rejecting the null hypothesis H_0 when H_0 is true. A Type I Error is considered a serious type of error. The probability of a Type I Error is α, called 'level of significance' of the test set by the researcher in advance.

$$\alpha = P_r \text{ (Type I Error)} = P_r \text{ (}H_0\text{ is rejected when it is true)}$$

You must always choose 'α' prior to the test. Conventionally, α = 0.10, 0.05 and 0.01 and can be increased or decreased depending on a particular problem. Choice of α defines RR or size of p-value that results in rejection of null hypothesis.

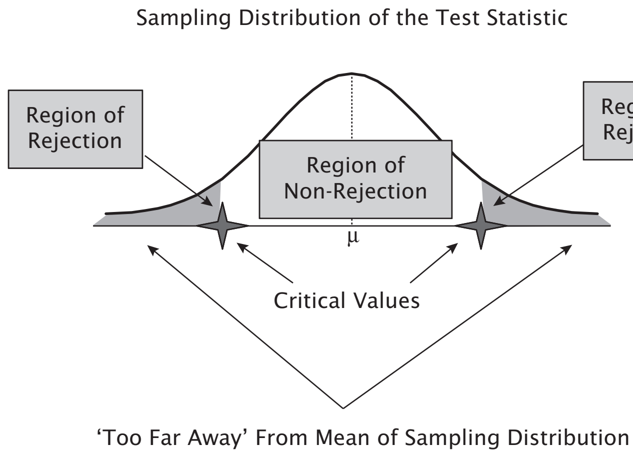


FIGURE 5.3 Rejection and Non-rejection Regions

Source: Authors.

TABLE 5.4 Possible Errors in Hypothesis Test Decision-making

	Possible Hypothesis Test Outcome	
	Actual Situation	
Decision	*H_0 True*	*H_0 False*
Do Not Reject H_0	No Error Probability $1 - \alpha$	Type II Error Probability β (Scenario 4)
Reject H_0	Type I Error Probability α (Scenario 3)	No Error Probability $1 - \beta$

Source: Authors.

A Type I Error is the incorrect rejection of a true null hypothesis. Usually, a Type I Error provides the researcher an evidence to conclude that a supposed effect or relationship exists when in fact it does not. Some good examples of Type I Errors are as follows: a medical test that shows a patient to have a disease when in fact the patient does not have the disease; a continuously ringing school fire alarm indicating a fire when in reality there is no fire or an experiment indicating that an education reform should resolve a given problem when in fact it does not.

Likewise, a 'Type II Error' is accepting H_0 when H_0 is false. It is a failure to reject a false null hypothesis. The probability of a Type II Error is denoted by β.

$$\beta = P_r \text{ (Type II error)} = P_r \text{ (not rejecting } H_0 \text{ when it is false)}$$

Examples of Type II Errors are as follows: a blood test failing to diagnose the disease it was meant to detect in a patient who really has the disease; a fire breaking out and the fire alarm does not ring or a clinical trial of a medical treatment failing to show that the treatment works when really it does.

What do these errors signify or explain? This can be easily understood from the above example of the secondary school. We make a Type I Error if we conclude that the population mean is 'not' 80 when it 'is' 80. This error causes us to adjust the examination system even though it is working properly. On the other hand, we make a Type II Error if we conclude that the population mean 'is' 80 when it is 'not' 80. Here, we would allow the process of the examination system to continue without adjustment even though adjustments are needed.

Confidence Coefficient and Level of Significance

'Confidence coefficient' is denoted by symbol 'C'. It is the complement of the probability of a Type I Error, that is, $C = (1 - \alpha)$. The confidence coefficient, $(1 - \alpha)$, is the probability that we will not reject the null hypothesis, H_0, when it is true and should not be rejected. The confidence level of a hypothesis test is $(1 - \alpha) \times 100$ per cent. In secondary school example above, the confidence coefficient measures the probability of concluding that the population mean is 80 marks when it is actually 80 marks.

Generally, it is either 95 per cent or 99 per cent. In hypothesis testing, it is shown in decimal terms, that is, 0.95 or 0.99. It simply shows how confident we are in our decision—how sure we are that we have made the right decision.

> 'The probability of making a Type I Error is called the "significance level" and is often denoted by α'.

If we do a hypothesis test at 99 per cent confidence, then we are 99 per cent sure with certainty that rejecting the null hypothesis was correct thing to do. But if we only have a level of confidence of 50 per cent, nobody is going to believe that the decision made is information. Thus, you should see that your level of confidence is around 90 per cent or 95 per cent.

The compliment of level of confidence is what we call 'level of significance', which is denoted by 'alpha (α)'. It is calculated as follows:

$$\text{Level of Significance } (\alpha) = 1 - \text{C (level of confidence)}$$

If the level of confidence is 95 per cent (0.95), then 'α' will be:

$$\alpha = 1 - 0.95 = 0.05$$

C and α always add up to 1, that is, $C + \alpha = 1$.

Power of Hypothesis Test

The complement of the probability of a Type II Error, $(1 - \beta)$, is known as the 'power of a statistical test'. It is 'P_r' (rejecting H_0 when it is false). Both α (level of significance) and β (power of test) are ideal when they are small (i.e., high power), but for a fixed sample size they trade off. In practice, the researcher controls α by 'choice' and β with 'sample size' (bigger sample, more power).

> 'The power of a statistical test, $(1 - \beta)$, is the probability of rejecting the null hypothesis (H_0) when it is false'.

In the secondary school graduate case, the power of the test is the probability that we will correctly conclude that the mean score is not 80 marks when it actually is not 80 marks.

It should be noted that the power of a test, $(1 - \beta)$, depends on a specific H_a parameter. Increasing the sample size (large samples generally permit the researcher to detect even very small differences between the hypothesised values and the population parameters), increasing the level of significance, denoted by α or increasing the difference between H_0 and H_a parameters will cause the power of the test to increase.

Normal Distribution

Before discussing the several sampling distributions commonly used in hypothesis testing, it is important to understand the basic features of a normal distribution.

Normal distributions are widely used statistical tools. They are often referred to as 'bell curves', given their bell-type shape, single-peaked and perfect symmetry. A normal curve represents a distribution of individuals and generally indicates that most individuals are typical or normal on a particular measurement.

The normal distribution serves many purposes. It can be used for calculating probabilities, examining students' performance on tests relative to the performance of other students, determining when certain students' scores are unusual or highly atypical and for deriving common metrics to compare students' scores across a variety of different assessments. The normal curve is the basis of many commonly used educational measures and many commonly used intelligence tests.

Formally, 'normal distribution' is defined in several different ways. But in simple terminology, it is that in a normal distribution, things tend to hover near the mean score—the closer a value is to the mean, the more we will see it; and the number of value on either side of the mean at any particular distance are equal.

> 'A normal distribution is one where the data are evenly distributed around the mean in a very regular way, which when plotted as a histogram will result in a bell curve'.
>
> 'A standard normal distribution is a distribution with a mean equal to zero and a standard deviation equal to one, that is $\mu = 0$ and $\sigma = 1$'.

The normal curve is generally measured in standard deviation units, referred to as 'Z scores'. The typical feature of a normal distribution is that it has a mean Z score of zero and a standard deviation of 1.0. Subtracting the mean score of the distribution from the individual's specific score and then dividing by the standard deviation of the distribution can determine a z score for any individual.

Examples of different shapes of normal distributions are shown in Figure 5.4. Although they differ in how spread out they are, the area under each curve is the same. The 'height' of a normal distribution is specified mathematically in terms of two parameters: the mean (μ) and the standard deviation (σ).

Following are the features of normal distributions:

- Normal distributions are symmetric around their mean.
- The mean, median and mode of a normal distribution are equal.
- The area under a normal curve is equal to 1.0.
- Normal distributions are denser at the centre and less dense at the tails.

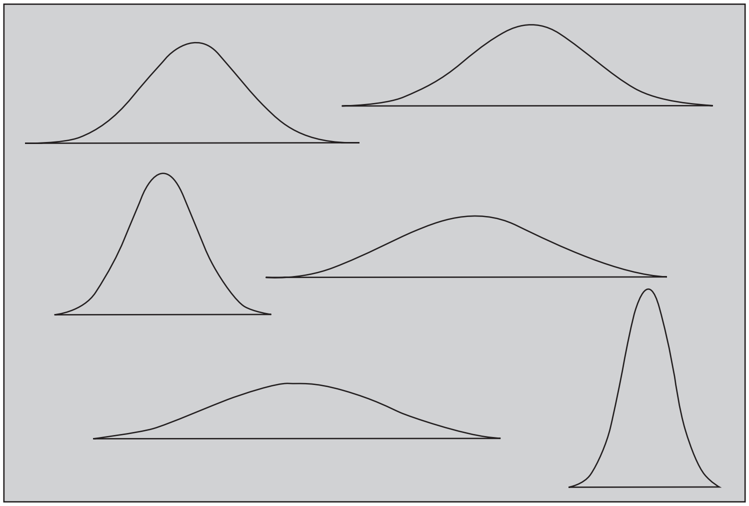


FIGURE 5.4 Different Shapes of Normal Distributions

Source: Authors.

- Normal distributions are defined by two parameters: the mean (μ) and the standard deviation (σ).
- Of the area of a normal distribution, 68.26 per cent, 95.44 per cent and 99.74 per cent are within ±1 standard deviation, ±2 standard deviations and ±3 standard deviations of the mean, respectively.

The Z scores and areas covered by a normal distribution are shown in Figure 5.5.

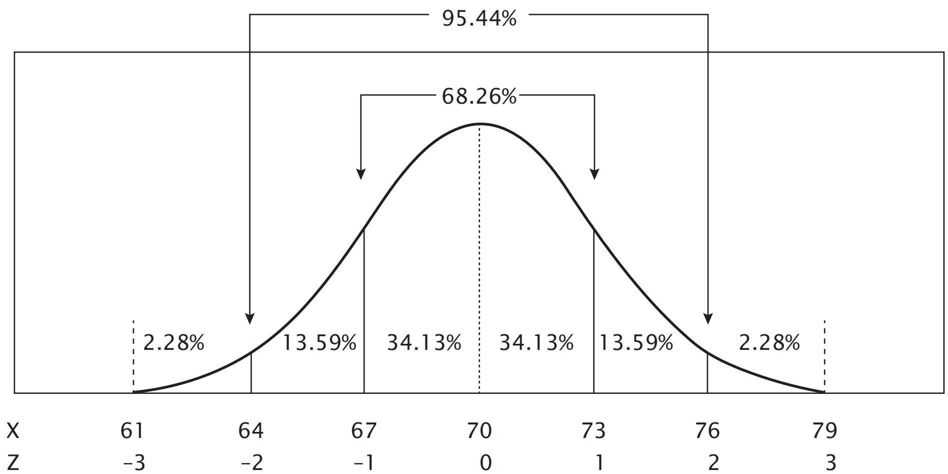


FIGURE 5.5 Standard Normal Distribution

Source: Authors.

Normal distributions can be transformed to standard normal distributions by using Equation 5.1. (You can also use 'Z score calculator' to find out these standard values.[2])

Computed Value of Z score $$Z = (X - \mu) / \sigma \quad (5.1)$$

Where
X = score from the original normal distribution
μ = mean of the original normal distribution
σ = standard deviation of the original normal distribution

A Z score always reflects the number of standard deviations above or below the mean a particular score is. Let us consider the following example.

Example

The English language test scores of Grade 5 of primary school pupils are normally distributed with a mean of 70 marks and a standard deviation of 3 marks.

Let us draw the distribution (Figure 5.5) and then convert it to standard normal distribution.

By using Equation 5.1, we get the computed value of Z score on the right-hand side of X (score from the original normal distribution). Thus, we can use Equation 5.1 for calculating all Z scores.

$$Z = (73 - 70) / 3 = 1$$

Areas under portions of the standard normal distribution are shown in Figure 5.5. About 0.68 (0.34 + 0.34) of the distribution is between standard deviation ±1, while about 0.95 of the distribution is between ±2 standard deviations.

Almost all statistical tests discussed below assume normal distributions. Fortunately, these tests work very well even if the distribution is only approximately normally distributed.

Area under a Portion of the Normal Curve

From our above example, if the test score is normally distributed with a mean of 70 marks and a standard deviation of 3 marks, we want to know:

1. What per cent of pupils scored less than 67 marks?
2. What per cent of pupils scored more than 67 marks?
3. What per cent of pupils scored between 61 and 65 marks?

i. Per cent of pupils scored less than 67 marks

This can be calculated by figuring out how many standard deviations below the mean the score of 67 are. Since 67 is 67 − 70 = − 3 points below the mean and since the standard deviation is 3, a score of 67 is −3 / 3 = −1 standard deviations below the mean. Or in terms of the formula,

$$Z = (67 - 70) / 3 = (-3) / 3 = -1.00$$

[2] See https://www.mathsisfun.com/data/standard-normal-distribution-table.html (accessed on 17 November 2017).

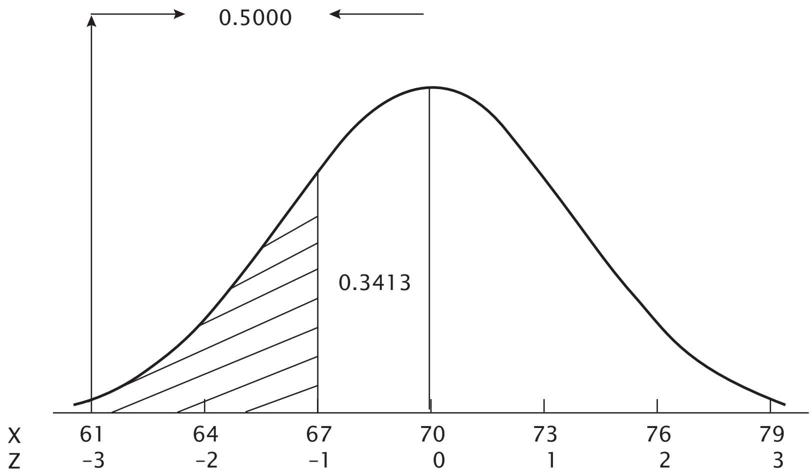


FIGURE 5.6 Per Cent of Pupils Who Scored Less Than 67 Marks

Source: Authors.

The '*z*-table' (Appendix 5.1) can be used to calculate that 0.3413 of the scores are less than or equal to a score –1.0 standard deviation below the mean. It follows that only 0.5000 − 0.3413 = 0.1587 of the scores are below a score of –1 standard deviation (67 marks). Thus, only 15.87 per cent of the scores are less than 67 marks (Figure 5.6).

Exercise

The IQ scores of a large group of students are normally distributed with a mean of 100 and a standard deviation of 15. What is the probability that a randomly selected student from this group will have an IQ score of below 128?

Solution

Compute for Z.

$$Z = 128 - 100/15 = 1.8667$$

Check the area with the computed Z. From the table, we get 0.4686. Since the table gives the area from 0 to z and 128 is above the mean 100, we add 0.5 to represent the other half of the curve below 100, which is 0 in standard form.

Answer

$$0.5 + 0.4686 = 0.9686 = 96.86\%$$

Exercise

If the price of a croissant is normally distributed with a mean of € 1.25 and a standard deviation of € 0.1, what is the probability that the croissant has a price less than € 1.00?

Solution

Compute for z.

$$Z = 1 - 1.25/0.1 = -2.5$$
$$1 - 1.25/0.1 = -2.5$$

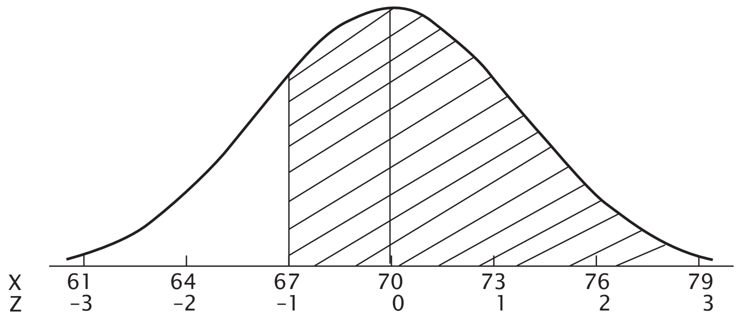


FIGURE 5.7 Per Cent of Pupils Who Scored More Than 67 Marks

Source: Authors.

Check the area for the computed z. From the table, we get 0.4938. Since the table gives the area from 0 to z and €1.00 is below the mean €1.25, we subtract 0.4938 from 0.5000 to find out the area below the mean. That is:

Answer

$$0.5000 - 0.4938 = 0.0.0062 = 0.62\%.$$

ii. Per cent of pupils scored more than 67 marks

We know from Figure 5.6 that the score between mean and the standard deviation −1 is 0.3413. We also know that the area between mean and +3 standard deviation is 0.5000. Using the *z*-table, we found that 0.3413 of the scores are less than or equal to a score −1.0 standard deviation below the mean. It follows that pupil scoring more than 67 marks will be 0.5000 + 0.3413 = 0.8413. Therefore, 84.13 per cent of the scores are above 67 marks (Figure 5.7).

Exercise

The average distance covered by a car in one hour is found to be 80 miles with a standard deviation of 6 miles. In the distribution, what is the per cent of the distances more than 88 miles?

Solution

Compute for z.

$$Z = 88 - 80 / 6 = 1.3333$$

From *z*-table, we get a value of 0.4082 corresponding the distance between 80 and 88 miles. Since 0.5000 is the value covered to the right-hand side of the mean (88 miles), we deduct 0.4082 from 0.5000 to represent the remaining area above 88 miles. That is,

Answer

$$0.5000 - 0.4082 = 0.0918 = 0.0918 = 9.18\%.$$

iii. Per cent of pupils scored between 62 and 65 marks

For calculating the percentage of pupils who scored marks between 62 and 65 (Figure 5.8), we have to calculate first two Z scores: Z score for 62 marks and z scores for 67 marks as follows:

Z score for 65 marks:

$$Z = (65 - 70) / 3 = -5 / 3 = -1.6666$$

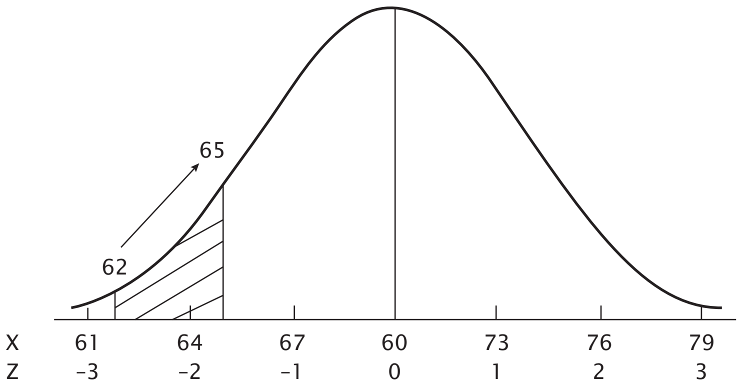


FIGURE 5.8 Per Cent of Pupils Who Scored Between 62 and 65 Marks

Source: Authors.

In *z*-table, the value of Z is –1.6666 and corresponds to 0.4525 or 45.24 per cent.
This means that 45.24 per cent pupils scored less than 65 marks.
Similarly, the Z score for 62 marks:

$$Z = (62 - 70) / 3 = -8 / 3 = -2.6666$$

For Z score of –2.6666, the corresponding value in *z*-table is 0.4961 or 49.61 per cent.
This means that 49.61 per cent pupils scored less than 62 marks.
Thus, the percentage of pupils who scored between 62 and 65 marks is:

$$= 49.61 - 45.24 = 4.37\%$$

Exercise

The times of the finishers in the Marathon of Rome City 10-km run are normally distributed with mean 61 minutes and standard deviation 9 minutes.

Determine the percentage of finishers with times between 50 and 70 minutes.

Solution:

We have

$$Z_1 = 50 - 61 / 9 = -1.22$$
$$Z_2 = 70 - 61 / 9 = -1.0$$

Answer

Therefore, the probability is 0.3888 + 0.3413 = 0.7301 = 73.01%.

Exercise

The number of calories in a salad on the lunch menu is normally distributed with mean $\mu = 200$ and standard deviation $\sigma = 5$. Find the probability that the salad you select will contain between 190 and 200 calories.

Solution

We have

$$Z = 190 - 200 / 5 = -2.0$$

Therefore, the probability is = 0.4772 or 47.72%

Key Concepts in Hypothesis Testing

Before we discuss the hypothesis testing procedures for means and proportions, it is necessary to understand the following key concepts commonly used in the process of hypothesis testing. These concepts are:

- The standard error of the mean (σ_M)
- The central limit theorem (CLT)
- The critical value of Z
- The critical value of t

The Standard Error of the Mean (σ_M)

> 'The standard error of the mean is the estimated standard deviation or measure of variability in the sampling distribution of a statistic'.

The standard error decreases as the sample size increases and approaches the size of the population.

Given a population with a mean of μ and a standard deviation of σ, the sampling distribution of the mean has a mean of μ and a standard deviation of:

Standard error of the mean $$\sigma_M = \sigma / \sqrt{n} \tag{5.2}$$

Where
σ_M = standard error of the mean
σ = standard deviation of the population
n = the sample size

In Equation 5.2, the standard deviation of the sampling distribution of the mean is called the standard error of the mean. It indicates the likely accuracy of the sample mean as compared with the population mean. In principle, Equation 5.2 does not assume a normal distribution. However, many of the uses of the formula do assume a normal distribution.

The formula shows that the larger the sample size, the smaller the standard error of the mean (the 'principle of large numbers'), that is, the spread of the sampling distribution of the mean decreases as the sample size increases (Figure 5.9).

We can specify the sampling distribution of the mean whenever two conditions are met:

- The population is normally distributed, or the sample size is sufficiently large.
- The population standard deviation σ is known.

Example

Assume that a school district has 10,000 Grade 6 students. In this district, the average weight of a Grade 6 student is 80 pounds, with a standard deviation of 20 pounds. Suppose you draw a random sample of 50 students. What is the probability that the average weight of a sampled student will be less than 75 pounds?

Solution

To solve this problem, we need to define the sampling distribution of the mean. Because our sample size is greater than 40, the CLT tells us that the sampling distribution will be normally distributed.

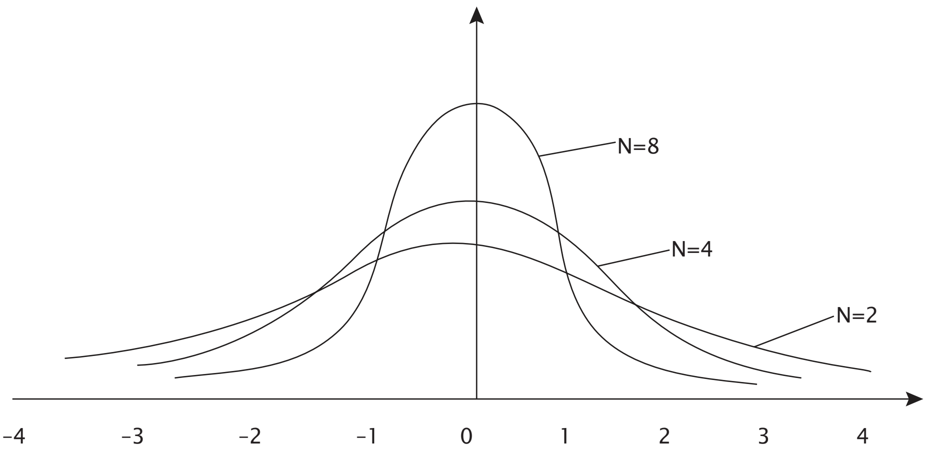


FIGURE 5.9 Sample Size and the Spread of the Sampling Distribution of the Mean

Source: Authors.

To define our normal distribution, we need to know both the mean of the sampling distribution and the standard deviation. Finding the mean of the sampling distribution is easy, since it is equal to the mean of the population. Thus, the mean of the sampling distribution is equal to 80.

The standard deviation of the sampling distribution can be computed by using the following equation:

$$\sigma_M = 20 / \sqrt{50}$$

$$\sigma_M = 20 / 7.07 = 2.82$$

Let us review what we know and what we want to know. We know that the sampling distribution of the mean is normally distributed with a mean of 80 and a standard deviation of 2.82. We want to know the probability that a sample mean is less than or equal to 75 pounds. To solve the problem, we plug these inputs into the normal probability calculator: mean = 80, standard deviation = 2.82 and value = 75. The calculator tells us that the probability that the average weight of a sampled student is less than 75 pounds and is equal to 0.038.

The Central Limit Theorem

The underlying premise of sampling distribution is the CLT. We have seen above that normal distribution is used to help measure the accuracy of many statistics, including the sample mean. CLT enables you to measure how much the means of various samples will vary, without having to take any other sample means to compare it with (Rumsey 2016). You can use this variability together with your data to answer questions about a population, such as 'What is the mean household income for the whole population in your country?' or 'This report said 75 per cent of all gift cards go unused; is that really true?' These two particular analyses are made possible by applications of the CLT called confidence intervals and hypothesis tests, respectively.

> 'The CLT is a statistical theory that states that given a sufficiently large sample size from a population with a finite level of variance, the mean of all samples from the same population will be approximately equal to the mean of the population. Furthermore, all of the samples will follow an approximate normal distribution pattern, with all variances being approximately equal to the variance of the population divided by each sample's size'.

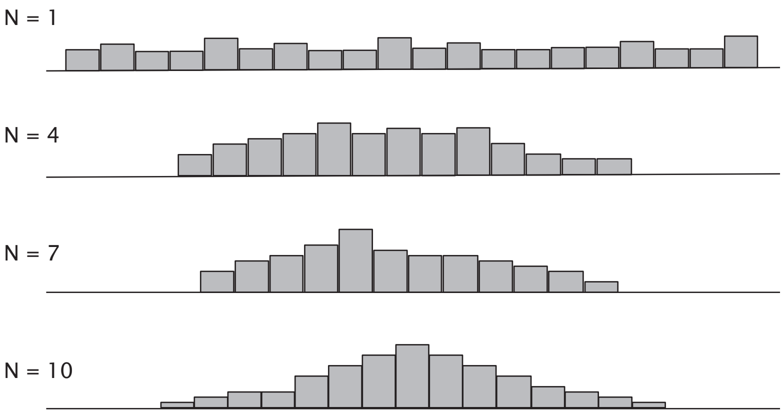


FIGURE 5.10 Central Limit Theorem and Frequency Distribution

Source: Authors.

The following shows the results of a simulation exercise to demonstrate the CLT. The computer sampled N scores from a uniform distribution and computed the mean. This procedure was performed 500 times for each of the sample sizes 1, 4, 7 and 10.

Figure 5.10 shows the resulting frequency distributions each based on 500 means. For N = 4, 4 scores were sampled from a uniform distribution 500 times and the mean computed each time. The same method was followed with means of 7 scores for N = 7 and 10 scores for N = 10.

Two things should be noted about the effect of increasing N:

- The distributions become more and more normal.
- The spread of the distributions decreases.

The CLT states that the sampling distribution of any statistic will be normal or nearly normal, if the sample size is large enough.

The Critical Value of Z

In statistics, the critical values are used to determine what probability any particular variable will have. It is a number that must be achieved so as to show statistical significance. A critical value of Z is used when the sampling distribution is normal or close to normal. If the critical value is achieved, then the null hypothesis is rejected. Usually, the critical value is written as Z_a, where the alpha level (α) is termed as the area in the tail. For example, $Z_{0.8} = 0.8416$.

It is important to note that a two-tailed test means that the answer should be applicable to both halves of the bell curve, and in a two-tailed test the answer must be expressed with both a '+' and '–'. However, for one-tailed test it can be on either side of the mean.

Thus, for a given alpha value 'α', the critical value in a two-tailed test is determined by:

$$\text{Probability (p): } p = (1 - \alpha)/2$$

and then looking up the result on a *z*-table.

Example

What is the critical Z value of $\alpha = 0.03$?

$$Z = (1 - 0.03)/2 = 0.485$$

We should note that all two-tailed tests must follow this step.

We then look up for the result on a z-table. If the number itself is not present, we have to choose the closest match in the table. For this example, 0.485 happens to be displayed on a z-table. First, we search in the z-table the corresponding number of the row in the far left column, which in this case is 2.1. Then we find the corresponding number of the column at the top row, which in this case is 0.07.

The critical value is the sum of 2.1 + 0.07 = 2.17

We should remember that our answer should be expressed in terms of either '+', '–' or both. The proper answer for the above case is ± 2.17.

You can use the online statistics calculator to find the Z score or normal score critical value with the given p value for both one-tailed and two-tailed normal distribution.

The Critical Value of t

The t-distribution (also known as aka, student's t-distribution) is a probability distribution that is used to estimate population parameters when the sample size is small and/or when the population variance is unknown.

For larger samples (at least 30), we typically use critical values on the Z distribution to calculate the margin of error (Appendix 5.1). However, for sample size of less than 30 and/or when the population standard deviation is unknown, we use the t-distribution to find critical values (Appendix 5.2). The distribution of the t statistic is called the t-distribution or the student t-distribution. The t-distribution allows us to conduct statistical analyses on certain data sets that are not appropriate for analysis, using the normal distribution.

Degrees of freedom (*df*) determine the shape of the student's t-distribution. With each increase in *df*, its shape changes. To help you find critical values for the t-distribution, you can use the last row of the t-table which lists common confidence levels, such as 90 per cent, 95 per cent, 98 per cent and so on. To find a critical value, look up your confidence level in the bottom row of the table; this tells you which column of the t-table you need. Intersect this column with the row for your *df*. The number you see is the critical value (or the t*-value) for your confidence interval. For instance, to determine the 0.05 critical value from the t-distribution with 6 *df*, look in the 0.05 column at the 6 row: $t_{(0.05,\ 6)} = 1.94318$.

Across the top row of the t-table, you, in fact, see the right-tail probabilities for the t-distribution. But confidence intervals demand both left- and right-tail probabilities (because you add and subtract the margin of error). It means that half of the probability left from the confidence interval goes into each tail. You need to take that into account. For example, a t*-value for a 90 per cent confidence interval has 5 per cent for its greater-than probability and 5 per cent for its less-than probability (taking 100 per cent minus 90 per cent and dividing by 2). Using the top row of the t-table, you would have to look for 0.05 (rather than 10 per cent, as you might be inclined to do). But using the bottom row of the table, you just look for 90 per cent (the result you get using either method ends up being in the same column).

When looking for t*-values for confidence intervals, use the bottom row of the t-table as your guide, rather than the headings at the top of the table.

The critical values of *t*-distribution are calculated according to the probabilities of two alpha values and the *df*. The alpha (α) values 0.05 one-tailed and 0.1 two-tailed are the two columns to be compared with the *df* in the row of the table.

Statistical Inference Using Hypothesis Testing Methods

The procedure for hypothesis testing follows six steps described above. Specifically, we have, so far, discussed how we can set up competing hypotheses, select a random sample from the population of interest and compute summary statistic. It is on the basis of these summary data, we then determine whether the sample data supports the null or alternative hypotheses. In this section, we present some cases (examples) as to how we can test the hypothesis for the population that is normally distributed.

A Hypothesis Testing for μ

Case 1:
When the population standard deviation 'σ' is known, we use the standardised test statistic Z.

Consider the following problem:

'The Director of the Secondary School Board Examination claims that the waiting time of announcement of the examination result is normally distributed with a mean of 3.5 weeks and standard deviation of 1.3 weeks. The school inspectorate found in a sample of 40 students in a given secondary school that the mean waiting time was 4.8 weeks. How far the claim of the Inspectorate on the mean waiting time is acceptable at 0.05 significance level?'

We know that Z-value in a normal distribution is:

$$Z = (X - \mu) / \sigma$$

Where
X = score from the original normal distribution
μ = mean of the original normal distribution or the hypothesised population mean (claimed mean)
σ = standard deviation of the original normal distribution

Since there is no direction implied in the problem with regard to the parameter of interest which is the population mean (μ) waiting time, thus a two-tailed test will have to be conducted for this particular problem.

Thus, given: $\mu_0 = 3.5$ weeks, $\sigma = 1.3$ weeks, n = 40, x = 4.8 minutes, $\alpha = 0.05$, the first thing we have to do is to compute the standard error of the mean (σ_M) as follows (Equation 5.2):

$$\sigma_M = \sigma / \sqrt{n}$$

Using the formula:

$$\sigma_M = \sigma / \sqrt{n} = 1.3 \text{ weeks} / \sqrt{40} = 0.2055 \text{ week}$$

- **Two-tailed test on μ**

Step 1
The first thing that we have to do for this case is to formulate our hypothesis. This could be written as H_0: $\mu = 3.5$ weeks. The true or population mean waiting time of students for the announcement of their results is 3.5 weeks.

Step 2
Thus, our alternative hypothesis in this case would be H_a: $\mu \neq 3.5$ weeks. The true mean waiting time of students is not 3.5 weeks.

Step 3
$\alpha = 0.05$ is our level of significance at which we will test our null hypothesis.

Step 4
Critical value is a term used in statistics that represents the number that must be achieved in order to demonstrate statistical significance. If the critical value is achieved, then the null hypothesis is rejected. A two-tailed test means that the answer should be applicable to both halves of the bell curve, and in a two-tailed test the answer must be expressed with both a '+' and '–' sign. For a given alpha value 'α', the critical value in a two-tailed test is determined by running the formula $(1 - \alpha)/2$ and then looking up the result on a z-table. In this case:

$$Cr\ z = \pm z_{a/2} = \pm z_{0.025} = \pm 1.96$$

Step 5
Compute

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{\bar{X}}}} = \frac{4.8 \text{ weeks} - 3.5 \text{ weeks}}{0.2055} = +\, 6.3246$$

Step 6
We reject the null hypothesis (H_0) because the sample evidence shows that the true mean waiting time of students in the school is not 3.5 minutes at $\alpha = 0.05$.

This is shown in Figure 5.11.

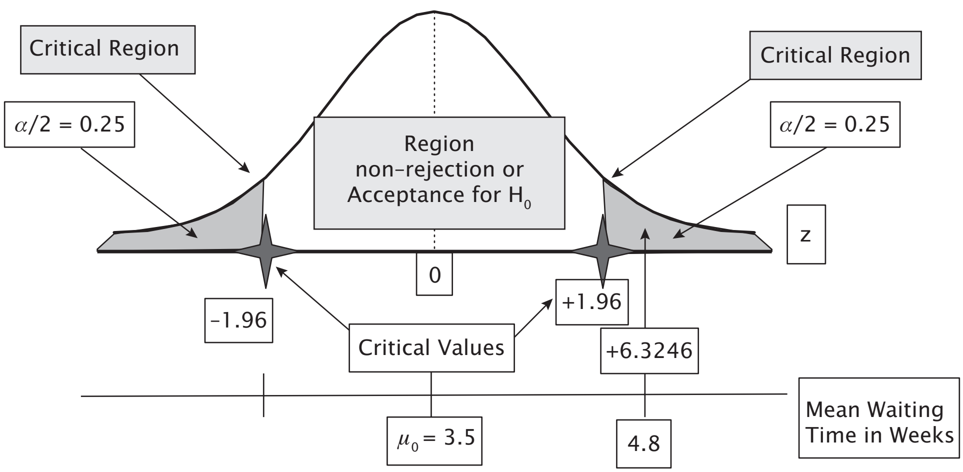


FIGURE 5.11 Hypothesis Testing for μ

Source: Authors.

Case 2

When the sample standard deviation 's' is known or could be known, we should use the test statistic *t* for testing our hypothesis.

Consider the following problem.

'To determine the degree to which the public primary schools in an imaginary country Ardania are dependent on textbook publishers, a survey of 12 publishers was conducted. The study revealed the following data on the total number of publishers dealing with primary schools: 4, 3, 2, 2, 3, 5, 2, 3, 3, 4, 4, 6. Test the hypothesis that the true mean number of publishers engaged by schools exceeds 3, using a 0.01 significance level'.

Given $\mu_0 = 3$ publishers, n = 12, $\bar{x} = 3.4167$ publishers, s = 1.2401 publishers

$$s_{\bar{x}} = s / \sqrt{n} = 0.3589 \text{ publishers}$$

The $\bar{x}$ and s values represent mean and standard deviation of the sample. You can calculate them by using excel. The existence of the word 'exceeds' referring to $\mu_0 = 3$ suggests a one-tailed or directional test.

- **One-tailed test on μ**

Step 1

Null Hypothesis $H_0: \mu = 3$

Step 2

Alternative Hypothesis $H_a: \mu > 3$

The true mean number of publishers exceeds 3. The symbol '>' is determined by comparing the observed value of the mean $\bar{x}$ to the hypothesised mean value μ_0.

Step 3

$$\alpha = 0.01$$

Step 4

$$C_r = +t_{df,a} = t_{11,0.01} = +\,2.7181$$

There is only one critical value since this is a one-tailed test. The sign of Cr t is + because the critical region is singly located at the upper tail area of the *t*-distribution. The *df* is given by (n – 1).

Step 5

$$\text{Computed } t = \frac{\bar{x} - \mu_0}{S_{\bar{x}}} = \frac{3.4167}{0.3580} = 1.1640$$

Step 6

Decision: Accept H_0.

Therefore, sample evidence shows that the true mean number of publishers engaged by school is 3 at $\alpha = 0.01$. The population average number of suppliers does not exceed 3 (Figure 5.12).

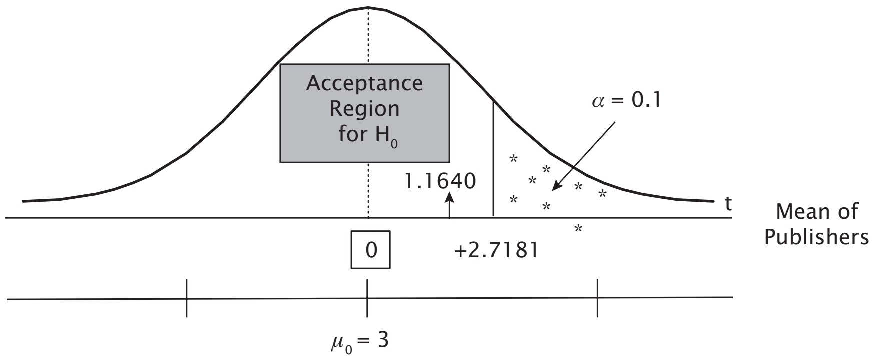


FIGURE 5.12 One-tailed Test on μ

Source: Authors.

Hypothesis Testing on the Difference of Two Population Means, μ_1-μ_2 for Independent Samples

Case 1

When the population standard deviation 'σ_1 and σ_2' are known, we use the standardised test statistic Z.

'The Dean Faculty of Education wishes to compare the score in the master's degree course in statistics of foreign students with those of national students in the university. University records show that the standard deviations for scores of national students and foreign students are 8 and 10, respectively. The Dean takes a sample of 32 national students with a mean score in statistics of 50 marks and a sample of 38 foreign students with a mean score of 40 marks. Test to determine which group is better performing at $\alpha = 0.01$'.

Given the following:

National students	*Foreign students*
$n_1 = 32$	$n_2 = 38$
$\bar{x}_1 = 50$ Marks	$\bar{x}_1 = 40$ Marks
$\sigma_1 = 8$ Marks	$\sigma_2 = 10$ Marks

Since the problem contains the words 'better performing', we have to undertake the one-tailed test.

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} = \sqrt{\frac{8^2}{32}+\frac{10^2}{38}} = 2.1521$$

- **One-tailed test on $\mu_1 - \mu_2$**

Step 1

Our null hypothesis in this case is: H_0: $\mu_1 - \mu_2 = 0$ or H_0: $\mu_1 = \mu_2$ meaning thereby no difference between the true means scores of foreign and national students.

Step 2
Similarly our alternative hypothesis can be formulated as: $H_a: \mu_1 - \mu_2 > 0$ or $H_a: \mu_1 > \mu_2 = 0$. This means that the true mean score of national students is higher than that of foreign students.

Step 3

$$\alpha = 0.01$$

Step 4

$$C_{rz} = +z_a = z_{0.01} = +\,2.33$$

Step 5

$$\text{The calculated } z = \frac{\bar{x}_1 - \bar{x}_2}{\sigma_{x_1 - x_2}} = \frac{50 - 40}{2.1521} = 4.6466$$

Step 6

Decision: Reject H_0.

Therefore, the sample evidence shows that the true mean score of the national students is higher than that of the foreign students at $\alpha = 0.01$. Hence, the national students are performing better than the foreign students (Figure 5.13).

Case 2
If the sample standard deviations, S_1 and S_2, are known and if the samples come from populations having equal deviations or variances, we have to use *t* values.

Let us consider the following case:

'As a result of recent advances in educational telecommunications, an open university is utilising instruction by interactive television for "distance" education. For example, each semester, the university televises six income-generating courses to adults at remote off-campus sites. To compare the performance of the off-campus adults at the university (which takes the televised classes) to the on-campus adults opting for the income-generating courses (who have a "live" professor), a test devised by the university was administered to a sample of both groups of students. The test scores (50 points maximum) are

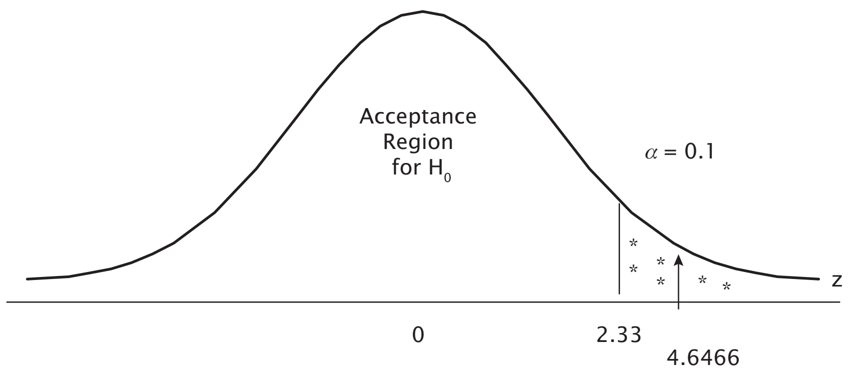


FIGURE 5.13 One-tailed test on $\mu_1 - \mu_2$

Source: Authors.

summarised in the following table. Based on these results, the researcher reports: 'there was no significant difference between the two groups of students'.

	Sample Size	*Mean*	*Standard Deviation*
On-campus Adults	$n_1 = 40$	$\bar{x}_1 = 41.93$	$\sigma_1 = 2.86$
Off-campus Adults	$n_2 = 40$	$\bar{x}_2 = 44.56$	$\sigma_2 = 1.42$

Do you agree with the researcher's findings at $\alpha = 0.05$?

$$S_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2}} \; x \; \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \tag{5.3}$$

$$\sqrt{\frac{(40-1)(2.86)^2 + (40-1)(1.42)^2}{40+40-2}}\left[\sqrt{\frac{1}{40} + \frac{1}{40}}\right] = 0.5049$$

- **Two-tailed test on $\mu_1 - \mu_2$**

Step 1
Our null hypothesis is: H_0: $\mu_1 = \mu_2$.
The true mean scores of the on-campus adults and off-campus TV adults

Step 2
Our alternative hypothesis is: Ha: $\mu_1 \neq \mu_2$
The true mean scores of the on-campus adults and off-campus TV adults are not equal.

Step 3

$$\alpha = 0.05$$

Step 4

$$C_{rt} = \pm t_{df, \frac{a}{2}} = \pm t_{78, 0.025} = \pm 1.9908$$

The *df* for 2 samples is $(n_1 + n_2 - 2)$. Hence, $df = 40 + 40 - 2 = 78$

Step 5
Computed

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_{\bar{x}_1 - \bar{x}_2}} = \frac{41.93 - 44.56}{0.5049} = 5.2090$$

Step 6
Decision: Reject H_0.
The sample evidence shows that the true mean scores of the on-campus students and off-campus TV students are not equal at $\alpha = 0.05$. Or there is a significant difference between the true mean scores of the two groups (Figure 5.14).

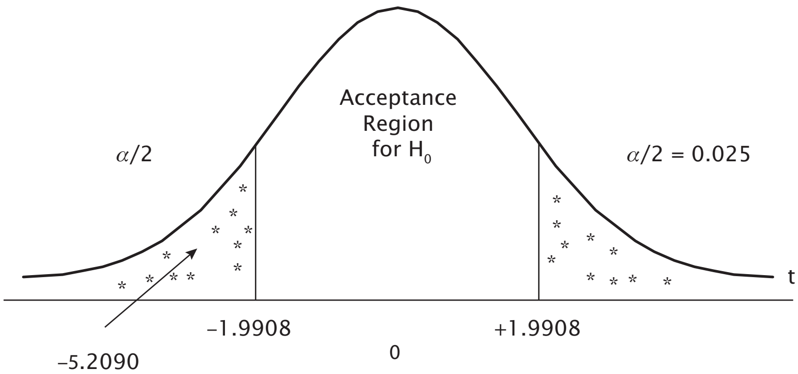


FIGURE 5.14 Two-tailed test on $\mu_1 - \mu_2$

Source: Authors.

Hypothesis Testing for One Population Proportion 'π'

In one-sample tests for a dichotomous outcome, we first formulate our hypotheses against an appropriate comparator then we select a sample and compute descriptive statistics on the sample data. Specifically, we compute the sample size (n) and the sample proportion which is computed by taking the ratio of the number of successes to the sample size.

Let us consider the following case.

Case 3:

'The inspectorate of primary school in an imaginary country Ardania assesses primary school teachers' performance each year. The inspector visits primary schools in a district and samples teachers. If the sample of teachers convinces the inspector that at least 80 per cent of the total primary school teachers as highly performing teachers, the inspector endorses the same for all teachers in the district. Would the inspector recommend the teachers' performance if a random sample of 600 teachers comprises 468 high-quality teachers at $\alpha = 0.05$'.

Let us note down the information that we have in this case:

$$\pi = 0.80, \qquad x = 468, \qquad n = 600$$

Proportion of success in the sample $$p = x / n \tag{5.4}$$

$$p = x / n = 468 / 600 = 0.78$$

Standard error of the proportion $$\sigma_p = \sqrt{\frac{\pi_0(1-\pi_0)}{n}} \tag{5.5}$$

$$\sigma_p = \sqrt{\frac{\pi_0(1-\pi_0)}{n}} = \sqrt{\frac{(0.80)(0.20)}{600}} = 0.0163$$

Where

x: number of successes in the sample

n: sample size

π: hypothesised population proportion or percentage

p: proportion of successes in the sample
σ: standard error of the proportion

Since this case contains the phrase 'at least 80 per cent', a one-tailed test on π will be conducted.

- **One-tailed test on** $\mu_1 - \mu_2$

Step 1
Our null hypothesis is: H_0: $\pi \geq 0.80$
At least 80 per cent of teachers are high-quality teachers.

Step 2
Our alternative hypothesis is: H_a: $\pi < 0.80$
Less than 80 per cent of the total teachers are high-quality teachers.

Step 3

$$\alpha = 0.050$$

Step 4

$$C_r Z = -Z_{0.05} = -1.65$$

Step 5
Computed

$$Z = p - \pi_0 / \sigma_p = 0.78 - 0.80 / 0.0163 = 1.2247$$

Step 6: Decision: Accept H_0.
The sample evidence shows that at least 80 per cent of the total teachers are of high-performing standard at $\alpha = 0.05$. Hence, the inspector would recommend high performance of all primary school teachers (Figure 5.15).

Hypothesis Testing for the Difference of Two Population Proportions, π_1 and π_2

This applies to two sample cases which have binomial distributions and have large (n ≥ 30) sample sizes.
Consider the following case:

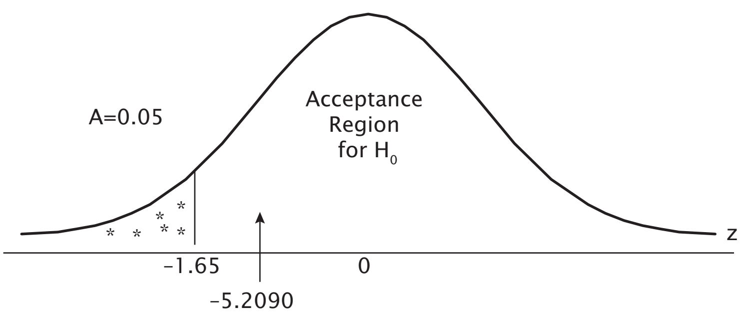


FIGURE 5.15 Hypothesis Testing for One Population Proportion π

Source: Authors.

Case 4:
A random sample of 200 male students was selected from Delhi University of which 40 male students were found to be in favour of the divorce bill. A random sample of 250 female students selected from the University at the same time revealed that 85 were in favour of such a new bill. Is the proportion of female students favouring the divorce bill higher than that of men in the University, given that $\alpha = 0.10$?

Let us first calculate the values of p_1 and p_2

$$p_1 = 40/200 = 0.2 \qquad \text{and} \qquad p_2 = 85/250 = 0.34$$

Thus,

Male Students	*Female Students*
$x_1 = 40$	$x_2 = 85$
$n_1 = 200$	$n_2 = 250$
$p_1 = 0.20$	$p_2 = 0.34$

In this case, we have to compute first the combined estimate of the two proportions.

$$\pi = \frac{x_1 + x_2}{n_1 + n_2} = \frac{40 + 85}{200 + 250} = 0.2778$$

Then we have to compute, $\sigma_{p_1 - p_2}$, the standard error of the difference of two sample proportions. Thus, The standard error of the difference of two sample proportions

$$\sigma_{p_1 - p_2} = \sqrt{\pi_0(1 - \pi_0)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]} \tag{5.6}$$

$$= \sqrt{0.2778(1 - 0.2778)\left[\frac{1}{200} + \frac{1}{250}\right]}$$

$$= 0.0425$$

This is one-tailed case, as determined from the directional words: 'higher than'.

- **One-tailed test on $\pi_1 - \pi_2$**

Step 1
In this case, the null hypothesis would be:

$$H_0: \pi_1 = \pi_2$$

That is, the proportions of male and female students in Delhi University who are in favour of the divorce bill are the same.

Step 2
Our alternative hypothesis would be:

$$H_a: \pi_1 < \pi_2$$

The proportion of male students in Delhi University who are in favour of the divorce law is less than that of women. The symbol < in H_a is determined by comparing p_1 with p_2.

Step 3

$$\alpha = 0.10$$

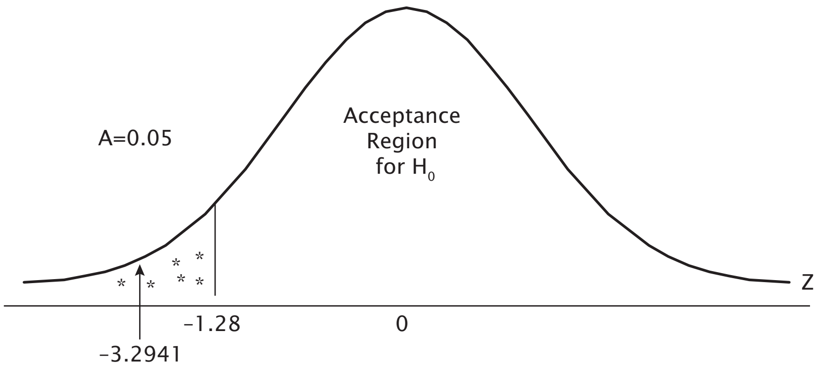


FIGURE 5.16 Hypothesis Testing for the Difference of Two Population Proportions, π_1 and π_2

Source: Authors.

Step 4

$$Cr\ Z = -z_a = -z_{0.10} = -1.28$$

Step 5
Computed

$$Z = \frac{p_1 - p_2}{\sigma_{p_1 - p_2}} = \frac{0.2 - 0.34}{0.0425} = 3.2941$$

Step 6
Decision: Reject H_0

The sample evidence (Figure 5.16) shows that the proportion of male students in Delhi University who are in favour of the divorce bill is lower than that of women at $\alpha = 0.10$.

Statistical Inference Using *f*-distribution

The *f*-distribution is the probability distribution associated with the *f*-statistic. Here we show you the method for computing an *f*-statistic and how to find probabilities associated with specific *f*-statistic values.

~~The *f* statistic~~

The *f*-statistic, also known as an *f*-value, is a random variable that has an *f*-distribution.

Here are the steps required to compute an *f*-statistic:

- Select a random sample of size n_1 from a normal population, having a standard deviation equal to σ_1.
- Select an independent random sample of size n_2 from a normal population, having a standard deviation equal to σ_2.
- The *f*-statistic is the ratio of s_1^2/σ_1^2 and s_2^2/σ_2^2.

The following equivalent equations are commonly used to compute an *f*-statistic:

$$f = \frac{(s_1^2 / \sigma_1^2)}{(s_2^2 / \sigma_2^2)} \tag{5.6}$$

$$f = \frac{(s_1^2 \times \sigma_2^2)}{(s_2^2 \times \sigma_1^2)} \tag{5.7}$$

$$f = \frac{(X^2_1 / v_1)}{(X^2_2 / v_2)} \tag{5.8}$$

$$f = \frac{(X^2_1 \times v_2)}{(X^2_2 \times v_1)} \tag{5.9}$$

Where
σ_1 = standard deviation of population 1
s_1 = standard deviation of the sample drawn from population 1
σ_2 = standard deviation of population 2
s_2 = standard deviation of the sample drawn from population 2
X_1^2 = Chi-square statistic for the sample drawn from population 1
v_1 = *df* for X_1^2
X_2^2 = chi-square statistic for the sample drawn from population 2
v_2 = *df* for X_2^2
df $v_1 = n_1 - 1$ and *df* $v_2 = n_2 - 1$

The distribution of all possible values of the *f*-statistic is called an *f*-distribution, with $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$ *df*.

The curve of the *f*-distribution depends on the *df*, v_1 and v_2. When describing an *f*-distribution, the number of *df* associated with the standard deviation in the numerator of the *f*-statistic is always stated first. Thus, *f* (5, 9) would refer to an *f*-distribution with $v_1 = 5$ and $v_2 = 9$ *df*, whereas *f* (9, 5) would refer to an *f*-distribution with $v_1 = 9$ and $v_2 = 5$ *df*. Note that the curve represented by *f* (5, 9) would differ from the curve represented by *f* (9, 5).

Cumulative Probability and the f-distribution

Every *f*-statistic can be associated with a unique cumulative probability (Appendix 5.3). This cumulative probability represents the likelihood that the *f*-statistic is less than or equal to a specified value.

Statisticians use f_α to represent the value of an *f*-statistic having a cumulative probability of $(1 - \alpha)$. For example, suppose we were interested in the *f*-statistic having a cumulative probability of 0.95. We would refer to that *f*-statistic as $f_{0.05}$, since $(1 - 0.95) = 0.05$.

Of course, to find the value of f_α, we would need to know the *df*, v_1 and v_2. It is important to note that the *df* appear in parentheses as follows: $f_\alpha\ (v_1, v_2)$. Thus, $f_{0.05}$ (5, 7) refers to value of the *f*-statistic having a cumulative probability of 0.95, $v_1 = 5$ *df*, and $v_2 = 7$ *df*.

The easiest way to find the value of a particular *f*-statistic is to use the *f*-distribution calculator. For example, the value of $f_{0.05}$ (5, 7) is 3.97. The use of the *f*-distribution calculator is illustrated in the examples below.

Example

Suppose you randomly select seven women from a population of women, and 12 men from a population of men. The table below shows the standard deviation in each sample and in each population.

Population	*Population Standard Deviation*	*Sample Standard Deviation*
Women	30	35
Men	50	45

Compute the *f*-statistic.

Solution A: The *f*-statistic can be computed from the population and sample standard deviations, using the following equation:

$$f = \left[s_1^2 / \sigma_1^2\right] / \left[s_2^2 / \sigma_2^2\right]$$

where σ_1 is the standard deviation of population 1, s_1 is the standard deviation of the sample drawn from population 1, σ_2 is the standard deviation of population 2 and s_2 is the standard deviation of the sample drawn from population 2.

As you can see from the equation, there are actually two ways to compute an *f*-statistic from these data. If the women's data appear in the numerator, we can calculate an *f*-statistic as follows:

$$f = (35^2 / 30^2) / (45^2 / 50^2) = (1{,}225 / 900) / (2{,}025 / 2{,}500) = 1.361 / 0.81 = 1.68$$

For this calculation, the numerator *df* v_1 are 7 − 1 or 6; and the denominator *df* v_2 are 12 − 1 or 11.

On the other hand, if the men's data appear in the numerator, we can calculate an *f*-statistic as follows:

$$f = (45^2 / 50^2) / (35^2 / 30^2) = (2{,}025 / 2{,}500) / (1{,}225 / 900) = 0.81 / 1.361 = 0.595$$

For this calculation, the numerator *df* v_1 are 12 − 1 or 11; and the denominator *df* v_2 are 7 − 1 or 6.

When you are trying to find the cumulative probability associated with an *f*-statistic, you need to know v_1 and v_2. This point is illustrated in the next example.

Example

Find the cumulative probability associated with each of the *f*-statistics from the aforementioned example.

To solve this problem, we need to find the *df* for each sample. Then, we will use the *f*-distribution calculator to find the probabilities.

- The *df* for the sample of women is equal to $n - 1 = 7 - 1 = 6$.
- The *df* for the sample of men is equal to $n - 1 = 12 - 1 = 11$.

Therefore, when the women's data appear in the numerator, the numerator *df* v_1 is equal to 6; and the denominator *df* v_2 is equal to 11. And, based on the computations shown in the previous example, the *f*-statistic is equal to 1.68. We plug these values into the *f*-distribution calculator and find that the cumulative probability is 0.78.

On the other hand, when the men's data appear in the numerator, the numerator *df* v_1 is equal to 11; and the denominator *df* v_2 is equal to 6. And, based on the computations shown in the previous example, the *f*-statistic is equal to 0.595. We plug these values into the *f*-distribution calculator and find that the cumulative probability is 0.22.

Statistical Inference Using Chi-square Distribution

The distribution of the chi-square statistic is called the chi-square distribution (Appendix 5.4). In this section, we will learn to compute the chi-square statistic and find the probability associated with the statistic.

Suppose we conduct the following statistical experiment. We select a random sample of size *n* from a normal population, having a standard deviation equal to σ. We find that the standard deviation in our sample is equal to *s*. Given these data, we can define a statistic, called 'chi-square', using the Equation 5.10:

$$X^2 = (n-1) \times s^2 / \sigma^2 \tag{5.10}$$

If we repeat this experiment an infinite number of times, we could obtain a sampling distribution for the chi-square statistic. The 'chi-square distribution' is defined by the following probability density function.

$$Y = Y_0 \times (X^2)^{(v/2-1) \times e^{x2/2}} \tag{5.11}$$

where

Y_0 = a constant that depends on the number of *df*
X^2 = Chi-square statistic
$V = n - 1$ number of units of *df*
e = a constant equal to the base of natural logarithm (approximately 2.7183)

In Figure 5.17, the red curve shows the distribution of chi-square values computed from all possible samples of size 3, where *df* is $n - 1 = 3 - 1 = 2$. Similarly, the green curve shows the distribution for samples of size 5 (*df* equal to 4); and the blue curve, for samples of size 11 (*df* equal to 10).

The chi-square distribution has the following properties:

- The mean of the distribution is equal to the number of *df*: $\mu = v$.
- The variance is equal to two times the number of *df*: $\sigma^2 = 2 \times v$.

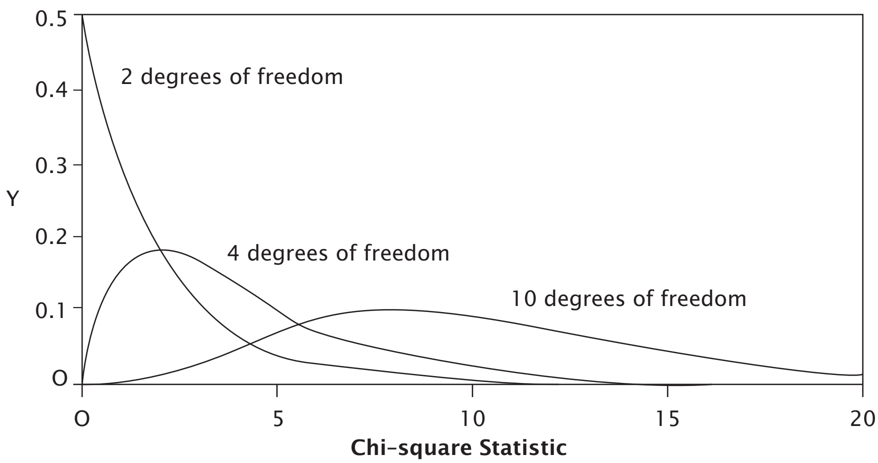


FIGURE 5.17 Distribution of Chi-square Values for Different Sample Sizes

Source: Authors.

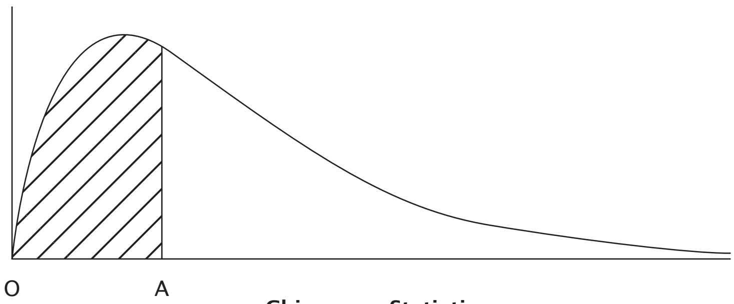


FIGURE 5.18 Cumulative Probability and Chi-square Distribution

Source: Authors.

- When the *df* are greater than or equal to 2, the maximum value for Y occurs when $X^2 = \nu - 2$.
- As the *df* increase, the chi-square curve approaches a normal distribution.

The chi-square distribution (Appendix 5.4) is constructed so that the total area under the curve is equal to 1. The area under the curve between 0 and a particular chi-square value is a cumulative probability associated with that chi-square value. For example, in Figure 5.18, the shaded area represents a cumulative probability associated with a chi-square statistic equal to *A*, that is, it is the probability that the value of a chi-square statistic will fall between 0 and *A*.

Fortunately, we do not have to compute the area under the curve to find the probability. The easiest way to find the cumulative probability associated with a particular chi-square statistic is to use the chi-square distribution calculator.

Example

The ABC Battery Company has developed a new cell phone battery. On average, the battery lasts 60 minutes on a single charge. The standard deviation is 4 minutes.

Suppose the manufacturing department runs a quality control test. They randomly select 7 batteries. The standard deviation of the selected batteries is 6 minutes. What would be the chi-square statistic represented by this test?

We know the following:

- The standard deviation of the population is 4 minutes.
- The standard deviation of the sample is 6 minutes.
- The number of sample observations is 7.

To compute the chi-square statistic, we plug these data in the chi-square equation, as shown below.

$$X^2 = [(n - 1) \times s^2] / \sigma^2$$

$$X^2 = [(7 - 1) \times 6^2] / 4^2 = 13.5$$

where

X^2 = Chi-square statistic

n = sample size

s = standard deviation of the sample

σ = standard deviation of the population

Example

Let us revisit the problem presented above. The manufacturing department ran a quality control test, using 7 randomly selected batteries. In their test, the standard deviation was 6 minutes, which equated to a chi-square statistic of 13.5.

Suppose they repeated the test with a new random sample of 7 batteries. What is the probability that the standard deviation in the new test would be greater than 6 minutes?

We know the following:

- The sample size n is equal to 7
- The *df* are equal to $n - 1 = 7 - 1 = 6$
- The chi-square statistic is equal to 13.5 (see example above)

Given the *df*, we can determine the cumulative probability that the chi-square statistic will fall between 0 and any positive value. To find the cumulative probability that a chi-square statistic falls between 0 and 13.5, we enter the *df* (6) and the chi-square statistic (13.5) into the chi-square distribution calculator. The calculator displays the cumulative probability: 0.96.

This tells us that the probability that a standard deviation would be less than or equal to 6 minutes is 0.96. This means (by the subtraction rule) that the probability that the standard deviation would be 'greater than' 6 minutes is 1 – 0.96 or 0.04.

Example

The Acme Widget Company claims that their widgets last 5 years, with a standard deviation of 1 year. Assume that their claims are true.

If you test a random sample of nine Acme widgets, what is the probability that the standard deviation in your sample will be less than 0.95 years?

We know the following:

- The population standard deviation is equal to 1
- The sample standard deviation is equal to 0.95
- The sample size is equal to 9
- The *df* is equal to 8 (because sample size minus one = 9 – 1 = 8)

Given these data, we compute the chi-square statistic:

$$X^2 = [(n - 1) \times s^2] / \sigma^2$$

$$X^2 = [(9 - 1) \times (0.95)^2] / (1.0)^2 = 7.22$$

where σ is the standard deviation of the population, s is the standard deviation of the sample and n is the sample size.

Now, using the chi-square distributor, we can determine the cumulative probability for the chi-square statistic. We enter the *df* (8) and the chi-square statistic (7.22) into the calculator and hit the calculate button. The calculator reports that the cumulative probability is 0.49. Therefore, there is a 49 per cent chance that the sample standard deviation will be no more than 0.95.

Example

Find the chi-square critical value, if the cumulative probability is 0.75 and the sample size is 25. We know the following:

- The cumulative probability is 0.75.

- The sample size is 25.
- The *df* is equal to 24 (because sample size minus one = 25 − 1 = 24).

Given these data, we compute the chi-square statistic, using the chi-square distribution calculator. We enter the *df* (24) and the cumulative probability (0.75) into the calculator and hit the calculate button. The calculator reports that the chi-square critical value is 28.2.

This means that if you select a random sample of 25 observations, there is a 75 per cent chance that the chi-square statistic from that sample will be less than or equal to 28.2.

Summary

In this chapter, we discussed how we use inferential statistics to measure behaviour in samples to learn more about the behaviour in populations that are often too large or inaccessible. We used samples because we know how they are related to populations. In behavioural research, we select samples to learn more about populations of interest to us. In terms of the mean, we measure a sample mean to learn more about the mean in a population. Therefore, we use the sample mean to describe the population mean.

Hypothesis testing begins with the drawing of a sample and calculating its characteristics (aka, 'statistics'). A statistical test (a specific form of a hypothesis test) is an inferential process, based on probability, and is used to draw conclusions about the population parameters.

The chapter explained the process of hypothesis testing which compares a population to which no treatment has been applied (control group) whose parameters (such as the mean and standard deviation) are already available with the population that is the same as the first, except that some treatment has been applied (experimental group) and the researcher does not know the parameters of this population. The researcher draws sample from this latter population which serves as the estimates of the unknown population parameters. The chapter explains in some details this situation in which researchers apply hypothesis testing. The chapter finally argues that hypothesis testing is not necessary if the entire population is small and accessible. But this is almost never the case.

Chapter 6 describes the various sampling techniques that the researchers use to conduct investigations.

Self-test Exercises

Exercise 5.1:
5.1.1. When deciding on a topic of your research, list the key things that you will need to consider.

Exercise 5.2:
5.2.1. A normal distribution has a mean of 75 and a standard deviation of 10. What percentage of the area is within 1.96 standard deviations of the mean?

Exercise 5.3:
Suppose a study was conducted on the effectiveness of a class on 'How to take tests'. The SAT scores of an experimental group and a control group were compared (there were 100 students in each group). The mean score of the experimental group was 503 and the mean score of the control group was 499. The difference between means was found to be significant, $p = 0.037$.

5.3.1. What would you conclude about the effectiveness of the class?

Exercise 5.4:

Ms Helen has been running a small-time gambling operation on her campus for several months. She sells each of the numbers 1 through 5 for $1.00 (collecting a total of $5.00) for each spin of a wheel. Then she spins the wheel of destiny. The person who holds the number where the spinner comes to rest gets $4.75 (Dian keeps 25¢ per spin for running the game).

She just purchased a new spinner, the critical piece of equipment for the game. Before she begins using this spinner, she wants to make certain that it is, in fact, fair—that is, she does not want some numbers to come up too often and others, not often enough (given the nature of the game, Helen has no incentive to cheat and she wants the game to be as fair as possible).

The results of spins are as follows:

Number	Number of Observations	Expected Number of Observations	Difference Between Observed and Expected	Difference Squared	Sq.Diff/exp. Number of Observations
1	10				
2	12				
3	6				
4	7				
5	15				
Total	50				

5.4.1. She comes to you, her statistical guru, and asks you to verify that the new spinner is fit to use. Describe a procedure for deciding whether the spinner is fair.

5.4.2. Complete the blanks in the above table.

5.4.3. What is the measure of the magnitude of the discrepancy between what she actually observed and what she could anticipate from a fair spinner?

5.4.4. Calculate chi-square from the data.

5.4.5. Explain in common sense terms the conditions under which the calculated X^2 (chi-square) value would be large and the condition under which it would be small.

5.4.6. What is the range of possible values? Smallest?____ Largest?_____

5.4.7. Is the calculated value of X^2 large or small?

5.4.8. Is the spinner fair or not?

References

Christopher, L.H. 2017. 'Developing the Hypothesis'. In *Research Methods*. Psych Central's Virtual Psychology Classroom. Available at https://allpsych.com/researchmethods/developingthehypothesis/ (accessed on 17 November 2017).

Creswell, J.W. 1994. *Research Design: Qualitative and Quantitative Approaches*. Thousand Oaks, CA: SAGE Publications.

Kowalczyk, D. 2015. *Writing Research Questions: Purpose and Examples*. Available at: https://study.com/academy/lesson/writing-research-questions-purpose-examples.html (accessed on 29 January 2018).

Rumsey, D.J. 2016. *Statistics for Dummies*, 2nd ed. Hobokey, NJ: Wiley.

USAISR. 2010. Available at: http://science.dodlive.mil/2010/10/04/defining-the-beginning-importance-of-research-questions-hypotheses/ (accessed on 29 January 2018).

Vasquez, E. 2003, March–April. 'Statistical Inference Using Hypothesis-testing Methods'. *Notes on Business Education* 6 (1).

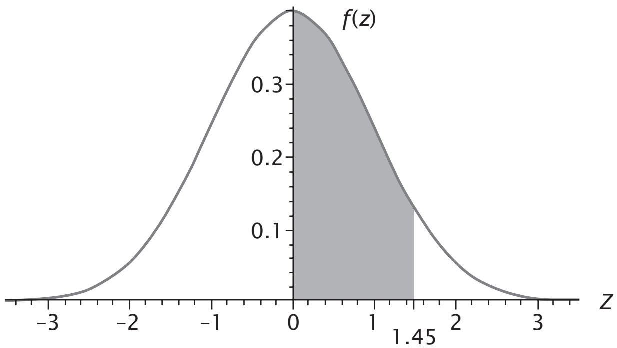
Appendix 5.1. *z*-Table

The *z*-Table

The *z*-table on this page indicates the area to the right of the vertical centre line of the *z*-curve (or the standard normal curve) for different SDs.

This table is very useful for finding probabilities when the event in question follows a normal distribution.

Example



In this standard normal curve, on the right, the mean is 0 and the SD is 1. The green shaded area represents the area that is within 1.45 SDs above the mean. The area of this shaded portion is 0.4265 (or 42.65 per cent of the total area under the curve).

To get this area of 0.4265, we read down the left side of the table for the SD's first two digits (the whole number and the first number after the decimal point; in this case, 1.4), then we read across the table for the '0.05' part (the top row represents the second decimal place of the SD that we are interested in).

z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	**0.4265**	0.4279

We have:

(left column) 1.4 + (top row) 0.05 = 1.45 SDs

The area represented by 1.45 SDs to the right of the mean is shaded in green in the standard normal curve.

You can see how to find the value of 0.4265 in the full *z*-table.

Follow the '1.4' row across and the '0.05' column down until they meet at 0.4265.

z-Table

z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3304	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952

z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999

Source: Bourne, M. 2016. The z-Table. Interactive Mathematics-Learn Math while you play with it. Available at http://www.intmath.com/counting-probability/z-table.php (accessed on 1 August 2017).

Appendix 5.2. Student *t*-Distribution Table

The critical values of *t*-distribution are calculated according to the probabilities of two alpha values and the *df*. The alpha (α) values 0.05 one-tailed and 0.1 two-tailed are the two columns to be compared with the *df* in the row of the table.

***t*-Table with Right-tailed Probabilities**								
Numbers in each row of the table are values on a* t*-distribution with* df *for selected right-tail (greater than) probabilities								
df/p	***0.4***	***0.25***	***0.1***	***0.05***	***0.025***	***0.01***	***0.005***	***0.0005***
1	0.32492	1	3.077684	6.313752	12.7062	31.82052	63.65674	636.6192
2	0.288675	0.816497	1.885618	2.919986	4.30265	6.96456	9.92484	31.5991
3	0.276671	0.764892	1.637744	2.353363	3.18245	4.5407	5.84091	12.924
4	0.270722	0.740697	1.533206	2.131847	2.77645	3.74695	4.60409	8.6103
5	0.267181	0.726687	1.475884	2.015048	2.57058	3.36493	4.03214	6.8688
6	0.264835	0.717558	1.439756	1.94318	2.44691	3.14267	3.70743	5.9588

(Continued)

(Continued)

t-Table with Right-tailed Probabilities								
Numbers in each row of the table are values on a t*-distribution with* df *for selected right-tail (greater than) probabilities*								
df/p	*0.4*	*0.25*	*0.1*	*0.05*	*0.025*	*0.01*	*0.005*	*0.0005*
7	0.263167	0.711142	1.414924	1.894579	2.36462	2.99795	3.49948	5.4079
8	0.261921	0.706387	1.396815	1.859548	2.306	2.89646	3.35539	5.0413
9	0.260955	0.702722	1.383029	1.833113	2.26216	2.82144	3.24984	4.7809
10	0.260185	0.699812	1.372184	1.812461	2.22814	2.76377	3.16927	4.5869
11	0.259556	0.697445	1.36343	1.795885	2.20099	2.71808	3.10581	4.437
12	0.259033	0.695483	1.356217	1.782288	2.17881	2.681	3.05454	4.3178
13	0.258591	0.693829	1.350171	1.770933	2.16037	2.65031	3.01228	4.2208
14	0.258213	0.692417	1.34503	1.76131	2.14479	2.62449	2.97684	4.1405
15	0.257885	0.691197	1.340606	1.75305	2.13145	2.60248	2.94671	4.0728
16	0.257599	0.690132	1.336757	1.745884	2.11991	2.58349	2.92078	4.015
17	0.257347	0.689195	1.333379	1.739607	2.10982	2.56693	2.89823	3.9651
18	0.257123	0.688364	1.330391	1.734064	2.10092	2.55238	2.87844	3.9216
19	0.256923	0.687621	1.327728	1.729133	2.09302	2.53948	2.86093	3.8834
20	0.256743	0.686954	1.325341	1.724718	2.08596	2.52798	2.84534	3.8495
21	0.25658	0.686352	1.323188	1.720743	2.07961	2.51765	2.83136	3.8193
22	0.256432	0.685805	1.321237	1.717144	2.07387	2.50832	2.81876	3.7921
23	0.256297	0.685306	1.31946	1.713872	2.06866	2.49987	2.80734	3.7676
24	0.256173	0.68485	1.317836	1.710882	2.0639	2.49216	2.79694	3.7454
25	0.25606	0.68443	1.316345	1.708141	2.05954	2.48511	2.78744	3.7251
26	0.255955	0.684043	1.314972	1.705618	2.05553	2.47863	2.77871	3.7066
27	0.255858	0.683685	1.313703	1.703288	2.05183	2.47266	2.77068	3.6896
28	0.255768	0.683353	1.312527	1.701131	2.04841	2.46714	2.76326	3.6739
29	0.255684	0.683044	1.311434	1.699127	2.04523	2.46202	2.75639	3.6594
30	0.255605	0.682756	1.310415	1.697261	2.04227	2.45726	2.75	3.646
Inf.	0.253347	0.67449	1.281552	1.644854	1.95996	2.32635	2.57583	3.2905
CL	-	-	**90.0%**	**95.0%**	**98.0%**	**99.0%**	**99.8%**	**99.9%**

Source: https://mat.iitm.ac.in/home/vetri/public_html/statistics/t-table.pdf (accessed on 1 August 2017).

Appendix 5.3. *f*-Distribution Table

Critical Values (Percentiles) for the *f*-Distribution

df_2 ↓	Numerator df df_1 →																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66

(Continued)

(Continued)

df_2 ↓	Numerator df df_1 → 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	4.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

Critical Values (Percentiles) for the *f*-Distribution
Upper One-sided 0.05 Significance Levels; Two-sided 0.10 Significance Levels, 95.0 Per Cent Percentiles

	Numerator df																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	6.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81

(Continued)

(Continued)

	Numerator df																		
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Critical Values (Percentiles) for the *f*-Distribution

Upper One-sided 0.025 Significance Levels; Two-sided 0.05 Significance Levels, 97.5 Per Cent Percentiles

	Numerator df																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	976.7	984.9	993.1	997.2	1001	1006	1010	1014	1018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.47	39.48	39.49	39.50	39.52
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.46	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11	2.04

(Continued)

(Continued)

	Numerator df																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57	2.47	2.36	2.25	2.19	2.13	2.07	2.00	1.93	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43	1.31
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.64	1.57	1.48	1.39	1.27	1.00

Critical Values (Percentiles) for the *f*-Distribution

Upper One-sided 0.01 Significance Levels; Two-sided 0.02 Significance Levels, 99.0 Per Cent Percentiles

	Numerator df																		
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*
1	4,052	5,000	5,403	5,625	5,764	5,859	5,928	5,982	6,022	6,056	6,106	6,157	6,209	6,235	6,261	6,287	6,313	6,339	6,366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36

(Continued)

(Continued)

	Numerator df																		
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

Critical Values (Percentiles) for the *f*-Distribution

Upper One-sided 0.005 Significance Levels; Two-sided 0.01 Significance Levels, 99.5 Per Cent Percentiles

	Numerator df																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	16,211	20,000	21,615	22,500	23,056	23,437	23,715	23,925	24,091	24,224	24,426	24,630	24,836	24,940	25,044	25,148	25,253	25,359	25,465
2	198.5	199.0	199.2	199.2	199.3	199.3	199.4	199.4	199.4	199.4	199.4	199.4	199.4	199.5	199.5	199.5	199.5	199.5	199.5
3	55.55	49.80	47.47	46.19	45.39	44.84	44.43	44.13	43.88	43.69	43.39	43.08	42.78	42.62	42.47	42.31	42.15	41.99	41.83
4	31.33	26.28	24.26	23.15	22.46	21.97	21.62	21.35	21.14	20.97	20.70	20.44	20.17	20.03	19.89	19.75	19.61	19.47	19.32
5	22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77	13.62	13.38	13.15	12.90	12.78	12.66	12.53	12.40	12.27	12.14
6	18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39	10.25	10.03	9.81	9.59	9.47	9.36	9.24	9.12	9.00	8.88
7	16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51	8.38	8.18	7.97	7.75	7.65	7.53	7.42	7.31	7.19	7.08
8	14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34	7.21	7.01	6.81	6.61	6.50	6.40	6.29	6.18	6.06	5.95
9	13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54	6.42	6.23	6.03	5.83	5.73	5.62	5.52	5.41	5.30	5.19
10	12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97	5.85	5.66	5.47	5.27	5.17	5.07	4.97	4.86	4.75	4.64
11	12.23	8.91	7.60	6.88	6.42	6.10	5.86	5.68	5.54	5.42	5.24	5.05	4.86	4.76	4.65	4.55	4.44	4.34	4.23
12	11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20	5.09	4.91	4.72	4.53	4.43	4.33	4.23	4.12	4.01	3.90
13	11.37	8.19	6.93	6.23	5.79	5.48	5.25	5.08	4.94	4.82	4.64	4.46	4.27	4.17	4.07	3.97	3.87	3.76	3.65
14	11.06	7.92	6.68	6.00	5.56	5.26	5.03	4.86	4.72	4.60	4.43	4.25	4.06	3.96	3.86	3.76	3.66	3.55	3.44
15	10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54	4.42	4.25	4.07	3.88	3.79	3.69	3.58	3.48	3.37	3.26
16	10.58	7.51	6.30	5.64	5.21	4.91	4.69	4.52	4.38	4.27	4.10	3.92	3.73	3.64	3.54	3.44	3.33	3.22	3.11
17	10.38	7.35	6.16	5.50	5.07	4.78	4.56	4.39	4.25	4.14	3.97	3.79	3.61	3.51	3.41	3.31	3.21	3.10	2.98
18	10.22	7.21	6.03	5.37	4.96	4.66	4.44	4.28	4.14	4.03	3.86	3.68	3.50	3.40	3.30	3.20	3.10	2.99	2.87
19	10.07	7.09	5.92	5.27	4.85	4.56	4.34	4.18	4.04	3.93	3.76	3.59	3.40	3.31	3.21	3.11	3.00	2.89	2.78
20	9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96	3.85	3.68	3.50	3.32	3.22	3.12	3.02	2.92	2.81	2.69
21	9.83	6.89	5.73	5.09	4.68	4.39	4.18	4.01	3.88	3.77	3.60	3.43	3.24	3.15	3.05	2.95	2.84	2.73	2.61
22	9.73	6.81	5.65	5.02	4.61	4.32	4.11	3.94	3.81	3.70	3.54	3.36	3.18	3.08	2.98	2.88	2.77	2.66	2.55

(Continued)

(Continued)

	Numerator df																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
23	9.63	6.73	5.58	4.95	4.54	4.26	4.05	3.88	3.75	3.64	3.47	3.30	3.12	3.02	2.92	2.82	2.71	2.60	2.48
24	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.69	3.59	3.42	3.25	3.06	2.97	2.87	2.77	2.66	2.55	2.43
25	9.48	6.60	5.46	4.84	4.43	4.15	3.94	3.78	3.64	3.54	3.37	3.20	3.01	2.92	2.82	2.72	2.61	2.50	2.38
26	9.41	6.54	5.41	4.79	4.38	4.10	3.89	3.73	3.60	3.49	3.33	3.15	2.97	2.87	2.77	2.67	2.56	2.45	2.33
27	9.34	6.49	5.36	4.74	4.34	4.06	3.85	3.69	3.56	3.45	3.28	3.11	2.93	2.83	2.73	2.63	2.52	2.41	2.25
28	9.28	6.44	5.32	4.70	4.30	4.02	3.81	3.65	3.52	3.41	3.25	3.07	2.89	2.79	2.69	2.59	2.48	2.37	2.29
29	9.23	6.40	5.28	4.66	4.26	3.98	3.77	3.61	3.48	3.38	3.21	3.04	2.86	2.76	2.66	2.56	2.45	2.33	2.24
30	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45	3.34	3.18	3.01	2.82	2.73	2.63	2.52	2.42	2.30	2.18
40	8.83	6.07	4.98	4.37	3.99	3.71	3.51	3.35	3.22	3.12	2.95	2.78	2.60	2.50	2.40	2.30	2.18	2.06	1.93
60	8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01	2.90	2.74	2.57	2.39	2.29	2.19	2.08	1.96	1.83	1.69
120	8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81	2.71	2.54	2.37	2.19	2.09	1.98	1.87	1.75	1.61	1.43
∞	7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62	2.52	2.36	2.19	2.00	1.90	1.79	1.67	1.53	1.36	1.00

Critical Values (Percentiles) for the *f*-Distribution
Upper One-sided 0.001 Significance Levels; Two-sided 0.002 Significance Levels, 99.9 Per Cent Percentiles

	Numerator df																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	4,053	5,000	5,404	5,625	5,764	5,859	5,929	5,981	6,023	6,056	6,107	6,158	6,209	6,235	6,261	6,287	6,313	6,340	6,366
2	998.5	999.0	999.2	999.2	999.3	999.3	999.4	999.4	999.4	999.4	999.4	999.4	999.4	999.5	999.5	999.5	999.5	999.5	999.5
3	167.0	148.5	141.1	137.1	134.6	132.8	131.6	130.6	129.9	129.2	128.3	127.4	126.4	125.9	125.0	125.0	124.5	124.0	123.5
4	74.14	61.25	56.18	53.44	51.71	50.53	49.66	49.00	48.47	48.05	47.41	46.76	46.10	45.77	45.43	45.09	44.75	44.40	44.05
5	47.18	37.12	33.20	31.09	29.75	28.84	28.16	27.64	27.24	26.92	26.42	25.91	25.39	25.14	24.87	24.60	24.33	24.06	23.79
6	35.51	27.00	23.70	21.92	20.81	20.03	19.46	19.03	18.69	18.41	17.99	17.56	17.12	16.89	16.67	16.44	16.21	15.99	15.75
7	29.25	21.69	18.77	17.19	16.21	15.52	15.02	14.63	14.33	14.08	13.71	13.32	12.93	12.73	12.53	12.33	12.12	11.91	11.70
8	25.42	18.49	15.83	14.39	13.49	12.86	12.40	12.04	11.77	11.54	11.19	10.84	10.48	10.30	10.11	9.92	9.73	9.53	9.33
9	22.86	16.39	13,90	12.56	11.71	11.13	10.70	10.37	10.11	9.89	9.57	9.24	8.90	8.72	8.55	8.37	8.19	8.00	7.81
10	21.04	14.91	12.55	11.28	10.48	9.92	9.52	9.20	8.96	8.75	8.45	8.13	7.80	7.64	7.47	7.30	7.12	6.94	6.76
11	19.69	13.81	11.56	10.35	9.58	9.05	8.66	8.35	8.12	7.92	7.63	7.32	7.01	6.85	6.68	6.62	6.35	6.17	6.00
12	18.64	12.97	10.80	9.63	8.89	8.38	8.00	7.71	7.48	7.29	7.00	6.71	6.40	6.25	6.09	5.93	5.76	5.59	5.42
13	17.81	12.31	10.21	9.07	8.35	7.86	7.49	7.21	6.98	6.80	6.52	6.23	5.93	5.78	5.63	5.47	5.30	5.14	4.97
14	17.14	11.78	9.73	8.62	7.92	7.43	7.08	6.80	6.58	6.40	6.13	5.85	5.56	5.41	5.25	5.10	4.94	4.77	4.60
15	16.59	11.34	9.34	8.25	7.57	7.09	6.74	6.47	6.26	6.08	5.81	5.54	5.25	5.10	4.95	4.80	4.64	4.47	4.31
16	16.12	10.97	9.00	7.94	7.27	6.81	6.46	6.19	5.98	5.81	5.55	5.27	4.99	4.85	4.70	4.54	4.39	4.23	4.06
17	15.72	10.66	8.73	7.68	7.02	6.56	6.22	5.96	5.75	5.58	5.32	5.05	4.78	4.63	4.48	4.33	4.18	4.02	3.85
18	15.38	10.39	8.49	7.46	6.81	6.35	6.02	5.76	5.56	5.39	5.13	4.87	4.59	4.45	4.30	4.15	4.00	3.84	3.67
19	15.08	10.16	8.28	7.26	6.62	6.18	5.85	5.59	5.39	5.22	4.97	4.70	4.43	4.29	4.14	3.99	3.84	3.68	3.51
20	14.82	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.24	5.08	4.82	4.56	4.29	4.15	4.00	3.86	3.70	3.54	3.38
21	14.59	9.77	7.94	6.95	6.32	5.88	5.56	5.31	5.11	4.95	4.70	4.44	4.17	4.03	3.88	3.74	3.58	3.42	3.26

(Continued)

(Continued)

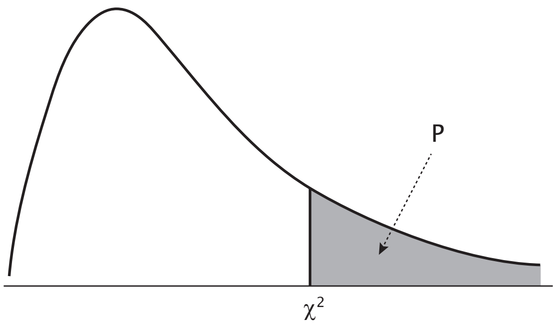
22	14.38	9.61	7.80	6.81	6.19	5.76	5.44	5.19	4.99	4.83	4.58	4.33	4.06	3.92	3.78	3.63	3.48	3.32	3.15
23	14.19	9.47	7.67	6.69	6.08	5.65	5.33	5.09	4.89	4.73	4.48	4.23	3.96	3.82	3.68	3.53	3.38	3.22	3.05
24	14.03	9.34	7.55	6.59	5.98	5.55	5.23	4.99	4.80	4.64	4.39	4.14	3.87	3.74	3.59	3.45	3.29	3.14	2.97
25	13.88	9.22	7.45	6.49	5.88	5.46	5.15	4.91	4.71	4.56	4.31	4.06	3.79	3.66	3.52	3.37	3.22	3.06	2.89
26	13.74	9.12	7.36	6.41	5.80	5.38	5.07	4.83	4.64	4.48	4.24	3.99	3.72	3.59	3.44	3.30	3.15	2.99	2.82
27	13.61	9.02	7.27	6.33	5.73	5.31	5.00	4.76	4.57	4.41	4.17	3.92	3.66	3.52	3.38	3.23	3.08	2.92	2.75
28	13.50	8.93	7.19	6.25	5.66	5.24	4.93	4.69	4.50	4.35	4.11	3.86	3.60	3.46	3.32	3.18	3.02	2.86	2.69
29	13.39	8.85	7.12	6.19	5.59	5.18	4.87	4.64	4.45	4.29	4.05	3.80	3.54	3.41	3.27	3.12	2.97	2.81	2.64
30	13.29	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39	4.24	4.00	3.75	3.49	3.36	3.22	3.07	2.92	2.76	2.59
40	12.61	8.25	6.60	5.70	5.13	4.73	4.44	4.21	4.02	3.87	3.64	3.40	3.15	3.01	2.87	2.73	2.57	2.41	2.23
60	11.97	7.76	6.17	5.31	4.76	4.37	4.09	3.87	3.69	3.54	3.31	3.08	2.83	2.69	2.55	2.41	2.25	2.08	1.89
120	11.38	7.32	5.79	4.95	4.42	4.04	3.77	3.55	3.38	3.24	3.02	2.78	2.53	2.40	2.26	2.11	1.95	1.76	1.54
∞	10.83	6.91	5.42	4.62	4.10	3.74	3.47	3.27	3.10	2.96	2.74	2.51	2.27	2.13	1.99	1.84	1.66	1.45	1.00

Source: http://www.socr.ucla.edu/Applets.dir/F_Table.html (accessed on 12 December 2017).

Appendix 5.4. Chi-squared Distribution Table

Values of the Chi-squared Distribution

The areas given across the top are the areas to the right of the critical value. To look up an area on the left, subtract it from one, and then look it up (i.e., 0.05 on the left is 0.95 on the right).



df	*0.995*	*0.99*	*0.975*	*0.95*	*0.90*	*0.10*	*0.05*	*0.025*	*0.01*	*0.005*
1	---	---	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582

(Continued)

(Continued)

df	*0.995*	*0.99*	*0.975*	*0.95*	*0.90*	*0.10*	*0.05*	*0.025*	*0.01*	*0.005*
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

Source: https://people.richland.edu/james/lecture/m170/tbl-chi.html (accessed on 1 August 2017).

6

CHAPTER

Selecting a Sample

Introduction

In educational research, we use sampling to allow the detailed study of part, rather than the whole, of a population. We use the information obtained from the sample (the portion of a population selected for analysis) to draw and/or develop meaningful generalisation about the universe (the whole population). It is a shortcut method for investigating a whole population. Researchers collect data on a small part of the whole parent population or sampling frame and use it to inform what the whole picture looks like.

It is nearly impossible to investigate and study an entire population (e.g., all college students, every individual in a country, every household, every geographic area and so on); researchers typically use sampling technique to gain a section of the population to perform an experiment or observational study.

> 'Investigation of certain queries to less than 100% of the population (group of all items that the researcher is trying to observe and analyse) is known as sampling'.

Suppose your MOE wants to undertake a survey about the smoking habits of high school students in your country. The population of high school student in the country is about 2 million. The Ministry cannot even think of running the survey for the entire student population, that is, asking each and every student to obtain the relevant data for reasons of considerable amount of time, money and other resources required for the study. The survey of the entire population of high school students would be too expensive to justify the effort. To solve these types of problem, the researcher uses extensively sampling. In this <fl, suppose the researcher chooses 2,000 students among 2 million students, then:

- 2 million students is the population and
- The size of the sample is 2,000 students

The key feature or importance of sampling is that the researcher can select a reasonable number of respondents (students in this case) from the target population to determine the size of the sample for his or her study. However, the number of respondents included in the sample should be adequate to warrant generalisation of the findings to the target population.

This chapter explains:

- The need for sampling in educational research
- Probability and non-probability sampling techniques
- Techniques for the selection of appropriate sampling techniques for different types of research topics
- Methods for assessing the representativeness of respondents

Why Do We Need Sampling?

A critical question in any research enquiry is: Who are our 'people' or what is our population of interest? In other words:

- To whom would we like the results to be generalised?
- How do we select our sample to generalise the results of this particular sample?
- What do we wish to generalise to this sample?
- What is an appropriate sample size?

We all know that we cannot observe directly each and every individual in the population we are studying. Alternately, we collect data from a subset of individuals (a sample) and use those observations to make inferences about the entire population.

It should be noted that when a researcher wants to know what proportion of the population has certain typical characteristics such as a particular opinion or a demographic feature, the correspondence between the sample and the larger population is extremely important (Henry 1990). A good example could be the public opinion polls which try to describe what percentage of the population is likely to vote, say, for the presidential candidate. The investigator does not ask the opinion of each and every voter to answer this question. What he or she requires is a sample that is highly representative of the population.

Sampling is used when:

- Large population can be conveniently covered
- We do not want to use up the cases (for example, when testing electric bulbs to see how long they last, we take a bulb and leave it on until it burns out. We cannot test all the bulbs this way, because our whole objective is to 'sell' the bulbs and not 'burn' them out).
- It is not necessary to survey all cases (for most purposes, taking a sample yields estimates that are accurate enough).
- We have limited time, money and energy.
- Units of area are homogenous.
- Per cent accuracy is not acquired.
- The data are unlimited.

What are the Advantages and Disadvantages of Sampling?

There are several advantages and disadvantages of sampling.

Advantages

- Economical: significantly less costly than the entire population
- Increased speed: Less time-consuming than the population to collect, analyse and interpret data
- Greater scope: easier data handling and management, comprehensive scope and flexibility
- Accuracy: accurate and complete studies, authentic results of the analysis, possibility of drawing valid inferences or generalisations
- Practical method: very practical method when the population is infinite
- Rapport: establishes adequate rapport with the respondents

Disadvantages

- Biased: possibility of biased selection resulting in erroneous conclusions
- Difficulty in selecting truly a representative sample: difficult to select a truly representative sample for complex topics
- Need for subject-specific knowledge: possibility of serious mistakes due to limited knowledge, training and experience of researcher in sampling theory
- Changeability of sampling units: unscientific method for heterogeneous units of population
- Impossibility of sampling: impossible to select a representative sample in case of small or too heterogeneous population

Sampling Terminology

Sampling constitutes an important component of any piece of research because of the significant impact that it can have on the quality of your results/findings. If you are not familiar with the sampling techniques and are planning to undertake a research enquiry, it is essential for you to know a number of key terms and basic tenets that act as a foundation to the subject. This section of the chapter explains these key terms and basic principles. Rather than describing comprehensively the sampling techniques, the section presents the sampling basics that you would need to know if you were an undergraduate or master's level student prior to preparing a dissertation (or similar piece of research). Some of the important sampling terms you will come across frequently include population, units, sample, sample size, sampling frame, sampling techniques, sampling bias and so on.

Population (Target Population)

The word 'population' is used differently in statistics than the way people normally consider it. In day-to-day parlance, it refers to the population of a country (or its geographical divisions/regions) such as China or India. But in research (particularly in the theory of sampling), the word 'population' refers to a different meaning. It means the units that researchers wish to investigate and study. These units could be people, cases and pieces of data. Some examples of each of these types of population are presented below:

People

- Number of students registered in a university (e.g., Beijing Normal University) or studying a particular course (e.g., statistics or research methodology and so on)
- Secondary school principals in a country having master's degree and two-year teachers training diploma
- University students using Facebook or Twitter
- Inspectors and supervisors of primary schools

Cases (i.e., organisations, institutions, countries and so on)

- Employment bureaus/recruitment agencies in Bangkok, Thailand
- United Nations Educational, Scientific and Cultural Organization (UNESCO)
- Signatories of Education for All declaration.

Pieces of data

- Average attendance rates of girls in primary education during 2014 and 2015 school years in Mali
- University applications in the People Republic of China in 2015
- Households with Internet subscriptions in Manila, Philippines

> 'Population is the universe of cases. It is the entire aggregation of items from which samples can be drawn. It is the group that you ultimately want to say something about. In statistics, *N* represents the "size of a population"'.

Suppose, the Director of the Non-formal Adult Education of Morocco wishes to generalise the impact of income-generation activities organised by the Directorate of Non-formal Adult Education for rural illiterate females in the age group of 30–50 years old. If that is the population of his or her interest, he or she is likely to have a very hard time developing a reasonable sampling plan. He or she is probably not going to find an accurate listing of this population, and even if he or she did, he or she would almost certainly not be able to mount a national sample across hundreds of rural areas. So he or she probably should make a distinction between the population he or she would like to generalise to and the population that will be accessible to him or her. The former is called 'theoretical population' (target population) and the latter 'accessible population' (survey population). In this particular example, the accessible population might be illiterate females between the ages of 30 and 50 in six selected rural areas across Morocco. Figure 6.1 shows the distinction between population, sample and case (element).

Census

It is a study in which the researcher obtains data from every member of a population. It is a complete study of all the elements present in the population. There are several studies where a census is not practical, because of the cost and time required. The general idea that a census generates more accurate data than sampling is not always true. Limitations include failure in generating a complete and accurate

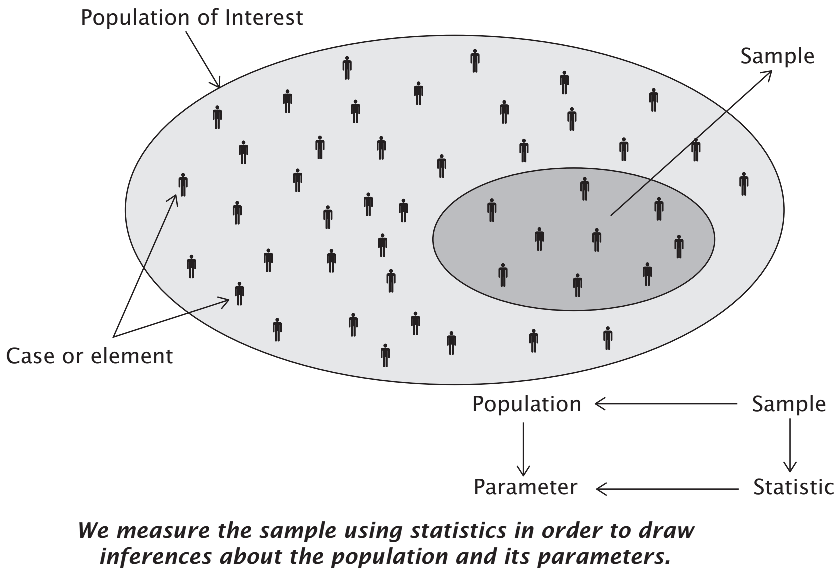


FIGURE 6.1 Population Versus Sample

Source: Authors.

list of all the members of the population and refusal of the elements (case) to provide information (see below 'response rate'). The national population census is an example of census survey.

Elements

Elements (cases) are the individuals in the population (usually, persons). Each object in a 'set' is an 'element of the set'. An element provides the basis of our research analysis and is the unit about which information is collected. In survey research, elements are people or certain types of people.

Statistic

Statistic is measurable characteristic of a sample. A statistic (singular) is a single measure of some attribute of a sample (e.g., its arithmetic mean value). It is calculated by applying a function (statistical algorithm) to the values of the items of the sample, which are known together as a set of data.

A statistic is a measure of a sample and not that of a population. Generally, a statistic is used to estimate the value of a population parameter. Unlike a parameter (a characteristic of a population), a statistic is a characteristic or measure of a sample and not that of a population.

Parameter

A parameter is a measurable characteristic of a population, such as a 'mean' or a 'standard deviation'. If you measure the entire population and calculate a value like a mean or an average, we do not refer this as statistic; we call it a 'parameter of the population'. It is the summary description of a given variable in a sample.

It follows from the above distinction that the mean of a population is denoted by the symbol 'μ'; but the mean of a sample is denoted by the symbol $\overline{X}$.

When we look across the responses that we get from an entire sample, we use a statistic. There are a wide variety of statistic we can use—mean, median, mode and so on.

Suppose we select a random sample of 50 students from a school with a total enrolment of 500 students. The average height of students included in the sample would be an example of a statistic. Thus, any measurable characteristic of the sample would be an example of a statistic.

Sampling Frame

Prior to selecting a sample, you have to define a sampling frame. A sample frame is a list of all the units of the population of interest (e.g., names of individuals, telephone numbers, residential addresses, census tracts and so on). If you were doing a phone survey and selecting names from a telephone directory, then that telephone directory would be your sampling frame. Similarly, if your research question or objective is concerned with registered children in primary schools in a local area, your sampling frame will be the school class lists of all registered children in primary schools in this area. Obtaining a sampling frame is therefore very important. You should be aware of the possible problems when using existing databases. These problems are likely to take any of the following form as revealed by research evidence:

- Individual databases are often incomplete
- The information held about organisations in databases is sometimes inaccurate
- The information held in databases soon becomes obsolete or out of date

In view of these issues, you have to make sure that your sampling frame is as complete, accurate and up to date as possible. If it is not, then it means that you have missed or excluded some cases. This is shown in Figure 6.2.

If this is the case, then your sample may not be representative of the total population and your research may be criticised for this reason. Similarly, you can draw inferences from the outcomes of the sample only to your population from which the sample has been drawn and not for other similar population groups or organisations. Saunders, Lewis and Thornhill (2009) suggest (Box 6.1) the checklist against which you can check your sampling frame.

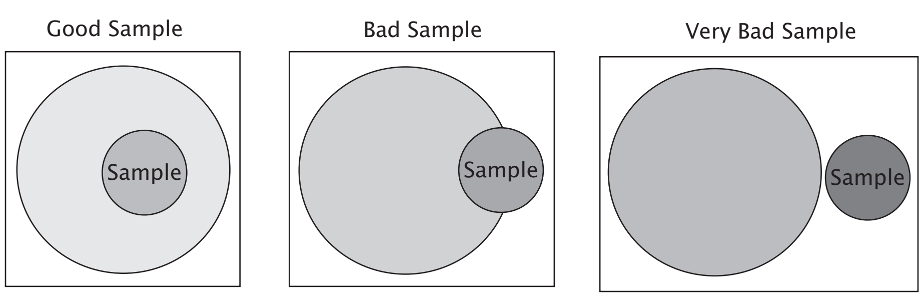


FIGURE 6.2 Good, Bad and Very Bad Samples

Source: Authors.

BOX 6.1: Checklist for Selecting the Sampling Frame

- ✓ Are cases listed in the sampling frame relevant to your research topic, in other words, will they enable you to answer your research question and meet your objectives?
- ✓ How recently was the sampling frame compiled? In particular, is it up to date?
- ✓ Does the sampling frame include all cases? In other words, is it complete?
- ✓ Does the sampling frame contain the correct information? In other words, is it accurate?
- ✓ Does the sampling frame exclude irrelevant cases? In other words, is it precise?
- ✓ (For purchased lists) can you establish and control precisely how the sample will be selected?

Source: Saunders, Lewis and Thornhill (2009).

Replacement

Sampling with replacement means the selection of a population element more than one time. It means that after you draw a number or name out of the box and record it, you put the name back so that it can be chosen again. Sampling without replacement means that once you draw the name out, it is no longer available to be chosen again.

Response

When you sample, there is possibility that some of the people of your sample supply you with one or more responses. In this sense, a response is a specific measurement value that a sampling unit supplies (refer below to 'response rate').

Sampling Design

It refers to how we define who the customer is (population), how we can contact them (sample frame) and the actual sample method to be used. The sample design encompasses all aspects of how to group units on the frame, determine the sample size, allocate the sample to the various classifications of frame units and finally select the sample. Figure 6.3 shows the five distinct steps to be followed in determining a sampling design process.

How Can You Select a Sample Unit?

The choice of sample unit must be appropriate. A sample unit is the actual unit that we include in our sample. Usually this unit refers to an individual person, but it could be an organisation, a company, a school or a neighbourhood, depending on what you are measuring and how you are measuring it. Suppose, you want to organise a research to study the attendance pattern or job satisfaction among primary school teachers, then all primary school teachers comprise the 'universe'. The 'elements of universe' may comprise primary school teachers in rural areas and in urban areas. The 'sample unit' may include male and female teachers, or junior and senior teachers as shown in Figure 6.4.

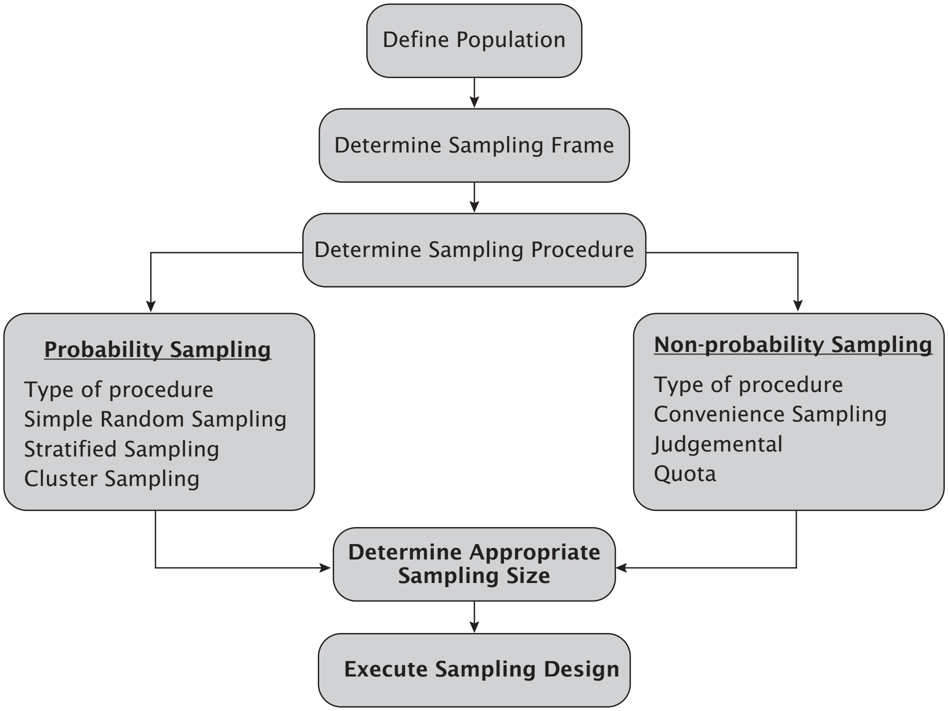


FIGURE 6.3 Sampling Design Process

Source: Authors.

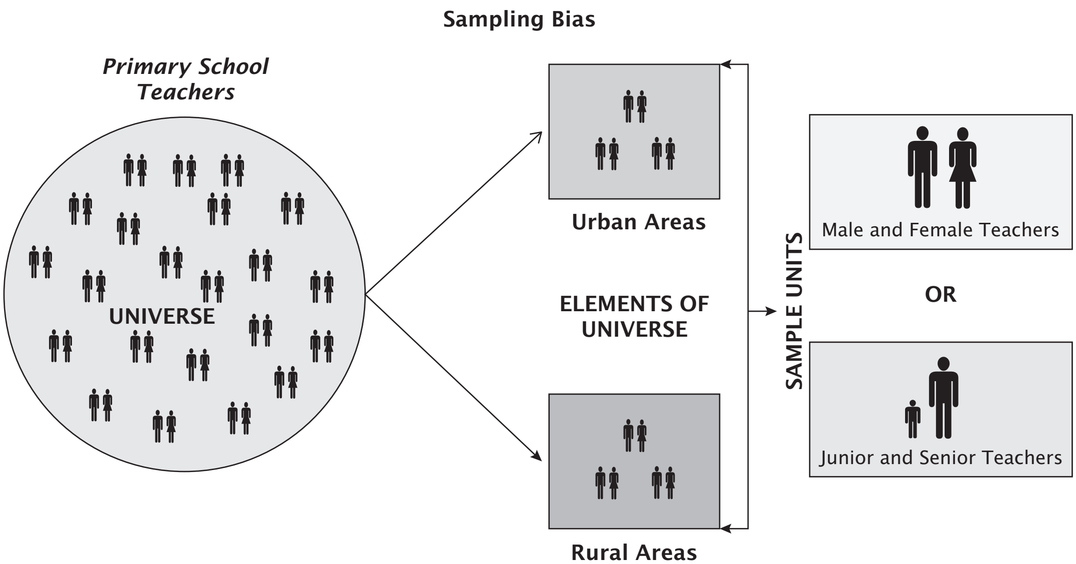


FIGURE 6.4 Universe, Elements and Sample Unit

Source: Authors.

Sampling bias occurs when the units that are selected from the population for inclusion in our sample do not reflect the population. Put in simple terms, it is any systematic failure of a sampling method to represent its population.

Suppose you want to calculate how often university students go to university library in a given month of the academic year. Standing at the entrance gate of the library and asking students as they pass by how often they went to library in a given month would not make sense because a higher proportion of those passing by are likely to have just come out of library. The sample would, therefore, be biased.

For this reason, you have to think very carefully the types of sampling techniques that you will use for selecting units for your sample. There are some sampling techniques, such as convenience sampling which is a type of non-probability sampling (the university library example above), which are vulnerable to greater bias than probability sampling techniques. We discuss sampling techniques further next.

What are Different Sampling Techniques?

The purpose of sampling techniques is to help the researcher select units (sample elements from population) to be included in his or her sample. Broadly speaking, there are two groups of sampling technique:

- Probability or representative sampling
- Non-probability or judgmental sampling

Probability Samples

Probability samples are those samples in which every unit in the population has an equal chance or chance greater than zero of being selected in the sample. To achieve this, the researcher utilises some form of random selection.

A typical feature of a probability sample is the absence of both 'systematic' and 'sampling bias'. If random selection has been done properly, your sample would be representative of the entire population. It means you will be able to answer your research questions to realise your research objectives that require you to estimate statistically the characteristics of the population from which you have drawn the sample. For example, you want to estimate the total income of adults living in your street. You will visit each household in that street, identify all adults living there and randomly select one adult from each household. You will then interview the selected persons and find their income.

Non-probability Samples

Non-probability samples are samples which a researcher using his or her subjective judgement selects units from the population to be included in the sample. These are samples where some elements of the population have 'no' chance of selection or where the probability of selection cannot be accurately determined. It involves the selection of elements based on assumptions regarding the population of interest. Since the selection of elements is non-random, non-probability sampling does not allow the estimation of sampling errors. As the information about the relationship between sample and population is limited, non-probability samples pose difficulties when the researchers intend to extrapolate from the sample to the population.

For some of the different types of non-probability sampling technique, experts have developed clearly defined procedures for selecting units to be included in the sample. However, in other sampling designs (e.g., purposive sampling), the units from the sample are selected on the basis of subjective judgement, which involves a combination of theory, experience and insight from the research process. This makes the selection of units more complicated. The most commonly used non-probability sampling techniques include quota sampling, purposive sampling, convenience sampling, snowball sampling and self-section sampling which are explained next in more details.

It is important here to point out that non-probability samples do not answer at all research questions or address objectives that require researchers to make statistical inferences about the characteristics of the population. Even then, researchers do generalise from non-probability samples about the population, but their generalisations are not grounded on statistical methods.

How Can You Select a Sampling Technique and the Sample?

Once you have decided the suitable sampling frame and the actual sample size needed for your study, the next step you have to follow is to select the most appropriate sampling technique or method to obtain a representative sample.

Researchers use several methods of sampling when doing a research study. This section of the chapter attempts to explain most of them so that you can pick up the most appropriate and suitable method of sampling for a given research problem. Indeed, 'simple random sampling' is the ideal, but researchers seldom have adequate resources (time and money) to approach the whole population, so many compromises often have to be made.

Probability Sampling

Among all the methods, probability sampling is the most preferred and the best overall group of methods of sampling. This is because the method enables researchers:

- Make statistical inferences
- Draw a representative sample
- Minimise sampling bias
- Choose units using probabilistic methods
- Fulfil the criteria for probability sampling

Henry (1990) advises that for doing a research inquiry comprising a total population of less than 50 cases, the researchers should avoid totally the use of the probability sampling methods. Instead, he or she should collect data for the entire population, 'as the influence of a single extreme case on subsequent statistical analyses is more pronounced for smaller samples than for larger samples'.

Your choice of selecting a probability sampling technique will be determined by your research question(s) and the objectives of your study. Subsequently, your need for face-to-face contact with respondents, the geographical area over which the population is spread and the nature of your sampling frame will further influence your choice of probability sampling technique. The structure of the sampling frame and the sample size you need will also influence your decision.

The following five key probability sampling methods are the most widely used techniques in social sciences research studies.

- Simple random
- Systematic
- Stratified random
- Cluster
- Multi-stage

Simple Random Sampling or Lottery Method of Sampling

Among all the sampling techniques, random sampling (also known as the 'lottery method of sampling') is the most popular and extensively used sampling method. It is the easiest form of probability sampling. In this method, each and every unit of the population has an equal chance of being selected or included in the sample. What you have to do here is to make sure that all the members (case or elements) of the population are included in your list for selecting randomly the desired number of subjects.

In random sampling, the researcher considers the complete list of the universe but he or she selects cases from this list 'at random'. However, he or she follows some uniform system for the selection of sample. This method of sampling is an appropriate and useful method for the conduct of telephone or mail survey.

Random selection does not mean haphazard selection. What it means is that the process of selecting a sample is independent of human judgment. In fact, the mathematics of random sampling is very complex.

Within this method, the most commonly used methods are:

- Lottery method
- Random numbers table

Lottery Method

In lottery method, each member or item of the population at hand is given a unique number. The numbers are then thoroughly mixed, like if you put them in a bowl or jar and shake it. Then, without looking these numbers, you select '*n*' numbers. The population members or items that are given that number are then included in the sample (Figure 6.5). In Figure 6.5, the population has 18 cases of which cases with numbers 2, 3, 5, 7, 10, 13, 14 and 18 are randomly selected.

Random Numbers Table

The other method is the use of 'random numbers table'. Most statistics and research methods books contain a table of random numbers as a part of the appendices. In a random table, there are 10,000 random digits between 0 and 9 arranged in groups of 5 and displayed in rows. In the table, all digits are equally probable and the probability of any given digit is unaffected by the digits that precede it (Appendix 6.1).

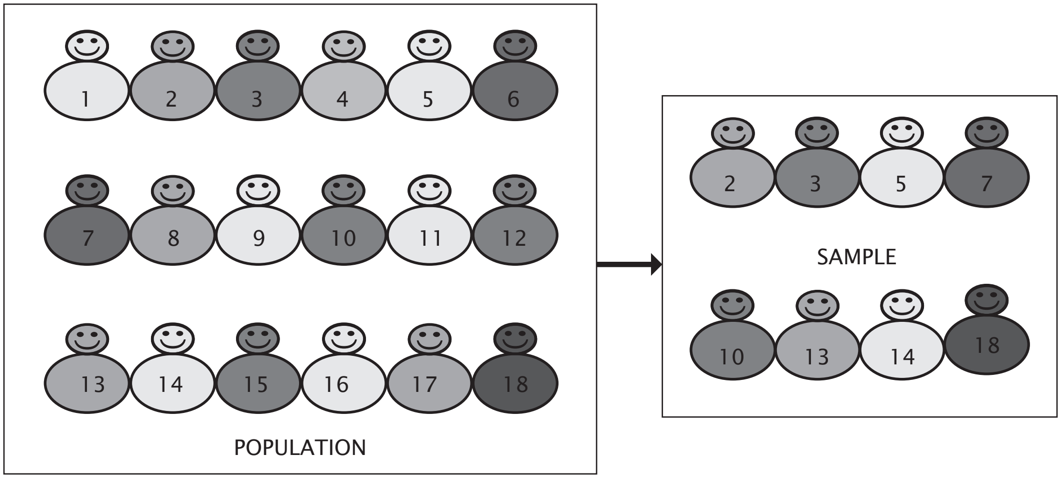


FIGURE 6.5 Selecting Simple Random Sample

Source: Authors.

The recommended way of obtaining a study group sample that best reflects or explains the main characteristics of the total population is to use simple random sampling. The following steps are used to randomly select a sample:

- Identify the total population.
- Determine the appropriate sample size.
- List the population and assign a number to each member of the population.
- Use the 'random number table' to select each member of the sample.

How can you obtain your study group sample using this method? The process is simple. First, the members of a given group or population that are accessible to you should be given an equal and independent chance of being selected into the study group. This means that no individual should have an effect on the selection of other individuals in the group. Then you place the names in a hat and draw the sample using simple random sampling. You use tables of random numbers or computerised randomiser in the simple random sampling process.

Each member of the population should be given a number from 1 to N. For determining the population size and the sample size, you select a starting point on the random number table. This is done by closing your eyes and pointing randomly on the page. Whichever number your finger touches is the number you start with. You must then choose a direction to read (up to down, left to right or right to left) to select other numbers. Next, you select the first *n* numbers (no matter how many numbers there are in your sample) whose last X digits are between 0 and N. Suppose, if N were a 3-digit number, then X would be 3. Put another way, if your population contained 600 people, you would use numbers from the table whose last three digits were between 0 and 600. If the number on the table is 85,941, then this number will not be selected because the last three digits (941) are greater than 600. You would skip this number and move to the next one. If the number is 61,149, you would use it and you would select the person in the population who is assigned the number 149. Continue this way through the table until you have selected your entire sample.

Example

Population of 600 Grade 6 students are assigned numbers 1 through 600. Sample of 10% or 60 students is needed for the study.

Randomly select a starting point from the random number table. This table with 5-digit numbers allows for populations up to 100,000.

Move down columns selecting the appropriate numbers. Use the last three digits in this case. Continue until 60 students are selected.

Using Table of Random Numbers

Selected numbers underlined and in **bold**

54 **463**	22 662	65 905
15 **389**	85 **205**	18 850
85 941	40 756	82 **414**
61 **149**	69 **440**	11 **268**
05 **219**	81 619	10 651
41 **417**	98 **326**	87 719
28 **357**	94 **070**	20 652
17 783	00 **015**	10 806
40 950	84 820	29 881
82 995	64 **157**	66 **164**

Sampling with Replacement

Sampling with replacement means selection of members or items of the population more than once for inclusion in the sample. For example, we have 100 names, each written on a piece of paper. You put all of those pieces of paper in a bowl and mixed up. Then you pick a name from the bowl, record the information to include that person in the sample, then put the name back in the bowl, mix up the names and select another piece of paper. The person who has been sampled just now has the same chance of being selected again. This is known as sampling with replacement.

Sampling Without Replacement

Sampling without replacement means selection of members or items only one time for inclusion in the sample. In the case of above example, suppose you put the 100 pieces of paper in a bowl, mix them up and randomly select one name to include in the sample. This time, however, you record the information to include that person in the sample and then set that piece of paper aside rather than putting it back into the bowl. Here, each element of the population can only be selected once.

Example

Let us say you have a population of 500 people and you wish to choose a simple random sample (SRS) of 50 people. First, each person is numbered 1 through 500. Then, you generate a list of 50 random numbers (typically with a computer programme) and those individuals assigned those numbers are the ones you include in your sample.

Advantages of Simple Random Sampling

A typical and significant feature of simple random sampling is the ease with which you can assemble your sample. The method is a fair method of selecting a sample from a given population since every member is assigned an equal opportunity of being selected.

Another key feature of simple random sampling is its representativeness of the population. Theoretically, the only thing that can compromise its representativeness is luck. If the sample is not representative of the population, the random variation is called 'sampling error'.

It is important to mention here that for drawing conclusions from the results of your research study, both an unbiased random selection and a representative sample are vital. While taking a sample from the population, the goal of your research to produce significant conclusions about the population from the results of the sample must always be in your mind. Due to the representativeness of a sample obtained by simple random sampling, it is reasonable to make generalisations from the results of the sample back to the population.

Disadvantages of Simple Random Sampling

Despite a wider use of simple random sampling methods, these methods are expensive to use, extremely time-consuming and difficult to manage/organise. They need a complete and an up-to-date list of all members of the target population. This list is usually not available for large populations. In such cases, it is wiser to use other sampling techniques.

Systematic Random Sampling

Because simple random sampling can be inefficient and time-consuming, statisticians turn to other methods, such as systematic sampling. Selection of a sample through a systematic approach can be done quickly.

Systematic sampling is another random sampling technique. Researchers use this technique for its simplicity and its periodic quality. In this sampling method, the researcher first randomly picks the first case from the population. After that, he or she selects each 'nth case' from the list. The interval between first case 'n' and the 'nth case' is called the sampling interval and is computed by dividing the population size by the desired sample size, that is:

$$\text{Sampling interval} = \frac{\text{total population}}{\text{sample size}}$$

The method is random because of two reasons. First, the periodic interval is determined beforehand. Second, the case 'n' (the starting point) is random. The procedure involved in systematic random sampling is very easy and can be done manually (Figure 6.6). The results are representative of the population. The process the systematic sampling technique follows is much like an arithmetic progression. The following steps are used to select a systematic sample using every 'kth' name.

Step 1: Identify the total population.
Step 2: Determine the appropriate sample size.
Step 3: List the population using the names of the members.
Step 4: Divide this sample number into the total population producing *n*th number.
Step 5: Select a random starting point on the population list within the first *k*th number. For example, suppose the total population is 1,000 people and we need a sample of 50 people. The *k*th number is 20. Select a starting point (some number between 1 and 20).
Step 6: Choose the next number by adding 20 to the random starting point.

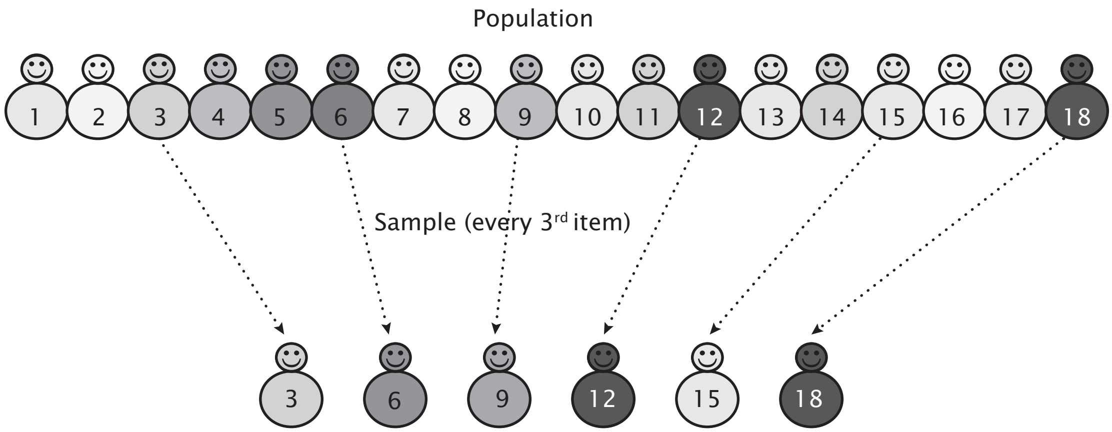


FIGURE 6.6 Selecting a Random Sample Using the Systematic Sampling Technique

Source: Authors.

Example

Suppose your total population 'N' is 200 and you want to select a sample 'n' of 50 units. It means the sampling interval for drawing a systematic random sample would be N / n = 200 / 50 = 4, that is, each fourth unit after the random selection of first unit. Randomly select a number from 1 to 4. Suppose you first pick your starting number 3. Then you add 4 (the interval number). The members of your sample will be individuals 7, 11, 15, 19, 23, 27 and so on until you complete the requisite size of 50.

Some researchers use a modified systematic random sampling technique. Here they first identify the needed sample size. They then divide the total number of the population with the sample size to obtain the sampling fraction. The sampling fraction is then used as the constant difference between subjects.

Example

Systematic Sampling

Selected numbers underlined and in **bold**

Obtain a listing of the entire population (200 students).

Select the number needed for the sample (20% or 40 students).

Divide sample size into population. In this scenario, the result would be 5.

Randomly select a number from 1 to 5. Suppose that we select number 2.

Start with the second person (Rita) and pick every fifth person on the list (Qutub).

1. Marry	14. Thomas	**27. <u>Vicky</u>**
2. <u>Rita</u>	15. Aftab	28. Pinky
3. Carlos	16. Yamin	29. Lata
4. Myriam	**17. <u>Edward</u>**	30. Qamar
5. Claire	18. Betu	31. Shams
6. Daniele	19. Huma	**32. <u>Najam</u>**
7. <u>Qutub</u>	20. Najma	33. Gudiya
8. Dian	21. Badar	34. Samina
9. Chao	**22. <u>Ren</u>**	35. Renu
10. Tania	23. Wang	36. Rahim
11. Ariba	24. Lisa	**37. <u>Naeema</u>**
12. <u>Salman</u>	25. Sonu	38. Ismail
13. Sami	26. Hasan	39. Anjum

One risk that researchers must take into account when conducting systematic sampling involves how the list used with the sampling interval is organised. If the population placed on the list is organised in a cyclical pattern that matches the sampling interval, the selected sample may be biased.

Example

The Human Resources Department of UNESCO wants to pick a sample of its professional staff and ask how they feel about the organisation's policies. The professionals are grouped in teams of 20, with each team headed by a director. If the list used to pick the sample size is organised with teams clustered together, the researcher risks picking only directors (or no directors at all) depending on the sampling interval.

Advantages of Systematic Sampling

Compared to random sampling, systematic sampling technique is a simple technique. With this technique, the researcher has the possibility to add a degree of system or process into the random selection of subjects.

The other advantage of systematic random sampling is an evenly sampled population. There exists a chance in simple random sampling that allows a clustered selection of subjects. This is systematically eliminated in systematic sampling.

Disadvantage of Systematic Sampling

Systematic sampling is more biased, as not all members or points have an equal chance of being selected. It may, therefore, lead to over- or under-representation of a particular pattern.

Stratified Random Sampling

Imagine that MOE in an imaginary country Ardania wants to organise a research to know more about the career goals of primary school teachers Atoll (province) A. Let us say that there are currently 2,000 primary school teachers at this level of education. Thus, these 2,000 teachers become our 'population (*N*)'. In order to select a 'sample (*n*)' of 400 teachers from this population of 2,000 teachers, we could use a 'SRS' or a 'systematic random sample'. However, the Ministry is interested in understanding the career goals of particular 'strata (groups)' within the population. Therefore, the stratified random sample involves dividing the population into two or more strata (groups). These strata are expressed as *H*.

Example

Consider Figure 6.7. It shows how you can use the stratified random sampling technique for selecting your sample of 400 teachers. We know that the total primary teachers are 2,000 of which 600 (30%), 1,000 (50%) and 400 (20%) are in urban, suburban and rural schools (strata), respectively. Using proportional sampling of 20 per cent in each area, you will have 120, 200 and 80 teachers from these three strata for your sample—a total of 400 teachers. You can further divide strata into two substrata, namely, merit pay group and non-merit pay group using 50 per cent. The number then will turn out to be 60 teachers each in merit pay group and non-merit pay group in urban school, 100 teachers each in merit and non-merit pay group in suburban schools and 40 teachers each in merit and non-merit group in rural schools.

Stratified sampling is a probability sampling technique. In this technique, the researcher divides the entire population into different subgroups. Each subgroup is called 'strata'. He then randomly selects the final subjects proportionally from the different strata (Figure 6.7). Age, gender, location (urban/rural),

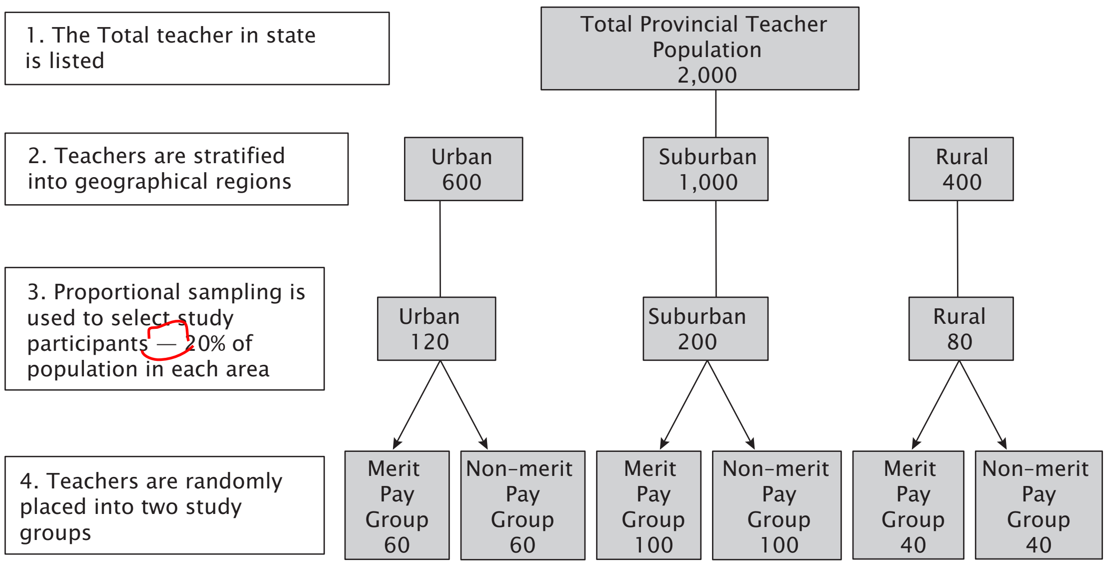


FIGURE 6.7 Selecting a Random Sample Using the Stratified Sampling Technique

Source: Authors.

socio-economic background, religion, nationality and educational attainment are the most common strata used in stratified random sampling.

Stratified sampling can be proportional or non-proportional. In proportional sampling, the participants are chosen in proportion to the number in each subgroup. Non-proportional sampling occurs when the response weight of the subgroup is not a factor.

Procedure for Creating a Stratified Random Sample

To create a stratified random sample, there are seven steps:

Step 1

In our above example, the total primary teachers are 2,000 in The Maldives. We express this population as *N*. Since we are interested in all of these primary school teachers, we can safely say that 2,000 teachers is our sampling frame. If our interest were only in urban teachers, for instance, we would exclude all urban teachers in creating our sampling frame, which in this case would be 600 teachers.

Step 2

If our aim is to understand the differences among urban, suburban and rural teachers, this would mean choosing location as the stratification, but it could similarly involve choosing teachers from different pay groups (e.g., merit pay group and non-merit pay group) or some other variable(s). For the purpose of this example, we will use school location as (urban, suburban and rural) as our strata.

Step 3

We need to identify all 2,000 primary teachers. If we were actually carrying out this research, we would most likely have to seek permission from teachers' records department (or another department in the ministry responsible for overall teacher recruitment) to view a list of all primary school teachers. You can read about this later in the part of this section under disadvantages (limitations) of stratified random sampling.

Step 4
As with simple random sampling and systematic random sampling techniques, we need to assign a consecutive number from '1 to NK' to each of the students in each stratum. As a result, we would end up with three lists. Lists 1, 2 and 3 would provide details of all urban, semi-urban and rural primary school teachers.

Step 5
Let us imagine that we choose a sample size of 400 teachers. The sample is expressed as *n*. We have decided this number because it takes into account the limit of our budget and the time we have to distribute our questionnaires to teachers. However, we could have also determined the sample size we needed using a sample size calculation, which is a particularly useful statistical tool discussed as well in this chapter.

Step 6
Imagine that of the 2,000 teachers, 30 per cent, 50 per cent and 20 per cent are urban, suburban and rural teachers, respectively. We need to ensure that the number of units selected for the sample from each stratum is proportionate to the number of males and females in the population. To achieve this, we first divide the desired sample size (*n*) by the proportion of units in each stratum. Therefore, to calculate the number of teachers in urban schools required in our sample, we divide 600 by 0.20 (i.e., 0.20 = 20% of the population of urban teachers), which gives us a total of 120 urban teachers. If we do the same for teachers in suburban school, we get 200 teachers (i.e., 20% of teachers are from suburban schools, where 1,000 ÷ 0.20 = 200) and so on. This means that we need to select 120, 200 and 80 teachers from urban, semi-urban and rural primary schools, respectively.

Step 7
Now that we have chosen to sample 120, 200 and 80 teachers from three strata, we still need to select these teachers from our three lists of urban, semi-urban and rural schools. We do this using either simple random sampling or systematic random sampling.

In fact, stratified random sampling is a modification of random sampling in which you divide the population into two or more relevant and significant strata based on one or a number of attributes. In effect, your sampling frame is divided into a number of subsets. A random sample is then drawn from each of the strata. Consequently, stratified sampling shares many of the advantages and disadvantages of simple random or systematic sampling.

A stratified sampling approach is, therefore, most effective when the following three conditions are met:

- Variability within strata is minimised
- Variability between strata is maximised
- The variables upon which the population is stratified are strongly correlated with the desired dependent variable

When to Use Stratified Random Sampling

The stratified random sampling method is best suited when there are specific subgroups in the population to be investigated (e.g., demographic groupings)

- When you want to highlight a specific subgroup within the population
- When you want to observe existing relationships between two or more subgroups

Types of Stratified Sampling

There are two types of stratified random sampling: proportionate and disproportionate.

Proportionate Stratified Random Sampling

In proportionate stratified random sampling, the sample size of each stratum is proportionate to the population size of the stratum when viewed against the entire population. This means that the each stratum has the same sampling fraction.

Example

You have three strata with 50, 100 and 150 population sizes, respectively. And the sample fraction that you chose is ½. Then, you must randomly sample 25, 50 and 75 subjects from each stratum, respectively. The important thing to remember in this technique is to use the same sampling fraction for each stratum regardless of the differences in population size of the strata.

Stratum	***A***	***B***	***C***
Population Size	50	100	150
Sampling Fraction	1 / 2	1 / 2	1 / 2
Final Sample Size	**25**	**50**	**75**

Disproportionate Stratified Random Sampling

The only difference between proportionate and disproportionate stratified random sampling is their sampling fractions. With disproportionate sampling, the different strata have different sampling fractions.

The precision of this design is highly dependent on your allocation of the sampling fraction. If you make mistakes in allotting sampling fractions, a stratum may be either over-represented or under-represented which will result in skewed results.

The advantages and disadvantages of the stratified random sampling are as follows.

Advantages

- Focuses on important sub-populations and ignores irrelevant ones
- Allows use of different sampling techniques for different sub-populations
- Improves the accuracy/efficiency of estimation
- Permits greater balancing of statistical power of tests of differences between strata by sampling equal numbers from strata varying widely in size

Disadvantages

- It requires selection of relevant stratification variables which can be difficult to find.
- It is not useful when there are no homogeneous subgroups.
- It can be expensive to implement.

Cluster Sampling

Cluster sampling refers to selection of respondents in groups (‘clusters’). In this method, instead of selecting all the subjects from the entire population right off, we take several steps in gathering our

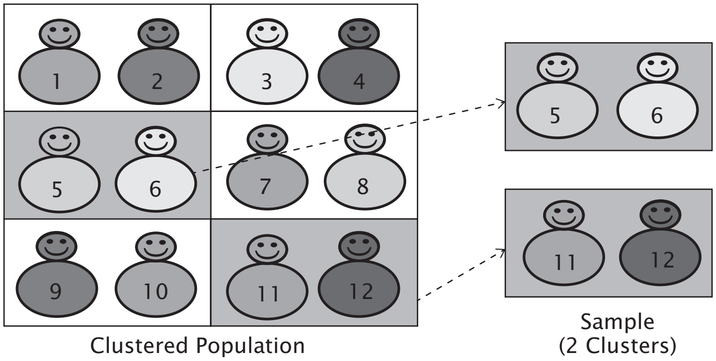


FIGURE 6.8 Selecting a Random Sample Using the Cluster Sampling Technique

Source: Authors.

sample population. We start with the selection of groups or clusters, and then from each cluster, we select the individual subjects by either 'simple random' or 'systematic random sampling'. For instance, we want to do a household survey within a city. We might select 100 city blocks and then interview every household within the selected blocks. Figure 6.8 shows the cluster sampling method.

Cluster sampling is often done on the basis of geographical cluster.

Example

Suppose you want to survey academic performance of high school students in the Philippines.

We divide the entire population of the Philippines into different clusters (cities). Out of these different clusters, we then select through simple random or systematic sampling a number of clusters (cities) depending on our research questions or hypothesis. From the randomly selected cities, either we can include all the high school students as subjects or we select a number of subjects from each cluster (city).

The important thing to remember about this sampling technique is to give all the clusters equal chances of being selected.

Clustering can reduce travel and administrative costs. In the example above, the interviewer can visit several households in one block in a single trip, rather than having to drive to a different block for each household.

In cluster sampling techniques, we do not need a sampling frame listing all elements in the target population. Instead, clusters can be chosen from a cluster-level frame. In the example above, we only need for our sample a block-level city map for initial selections, and then a household-level map of the 100 selected blocks. We do not need a household-level map of the whole city.

Horner and O'Kelly (2005) argue that the researcher can modify the random sampling process by using the cluster sampling method. This is possible because the cluster sampling utilises the convenience of naturally occurring groups. The following is an example where cluster sampling would be appropriately used.

Example

A large suburban school district wants to test the effect of a newly integrated reading programme on sixth graders. There are 3,000 Grade 6 students based in 100 classrooms in the school division. The researcher would prepare a list of all 3,000 students with the help of the table of random numbers to select the study participants. A possible problem with the process could be the probability of emergence of a situation where every one of the

100 classes would have a few students represented in the sample. What are some of the problems random sampling would create?

The answer to the above question is:

- The researcher can find the cluster sampling relatively difficult to administer because each class would have only a small number of students in the sample.
- It is difficult for the interviewer to set up a control and experimental group because some students would be in the same class.
- The technique increases cost and time to train the participants in all 100 classrooms.

For this example, the interviewer will follow the four steps (see also Figure 6.9).

Step 1
Determine what cluster to use. The logical cluster to use for this example would be each of the 100 individual classrooms.

Step 2
Number the 100 classrooms and determine the number of subjects (classrooms in this case) needed. In this case, we have chosen 30 classrooms.

Step 3
Number the 30 chosen classrooms. Using random selection, determine which 15 cases would be chosen for each of the experimental and control groups.

Step 4
Apply the treatment or independent variable to the experimental classrooms.

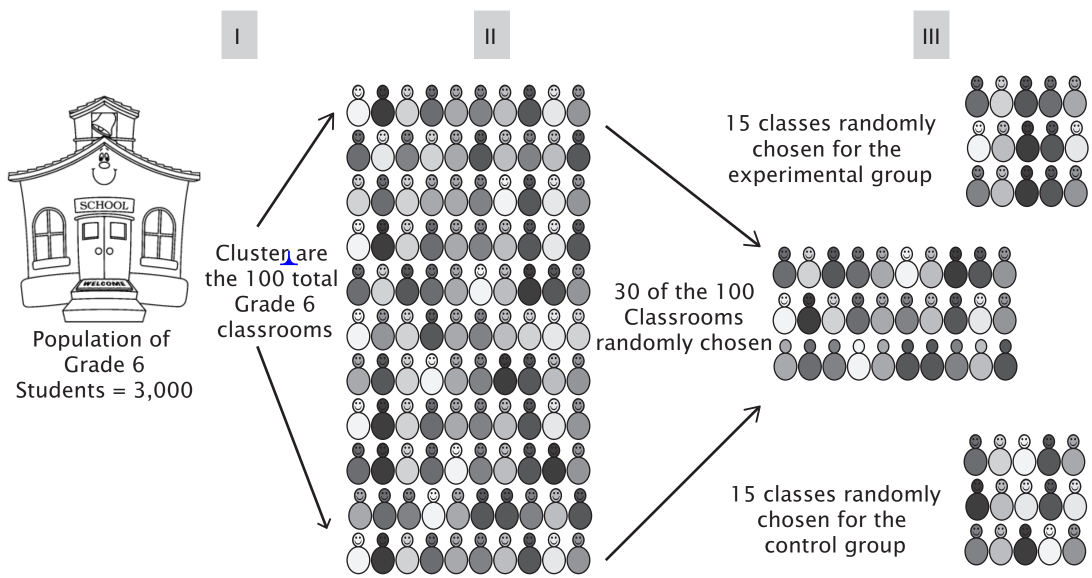


FIGURE 6.9 Schematic of Cluster Sampling

Source: Authors.

Difference Between Cluster Sampling and Stratified Sampling

The main difference between cluster sampling and stratified sampling lies with the inclusion of the cluster or strata.

In stratified random sampling, all the strata of the population is sampled, while in cluster sampling, we only randomly select a number of clusters from the collection of clusters of the entire population. Therefore, only a number of clusters are sampled, and all the other clusters are left unrepresented.

Advantages and Disadvantages of Cluster Sampling

Advantages

- Cluster sampling is cheap, quick and easy. The researcher does not sample the whole country in this technique. He randomly selects few clusters or areas when using cluster samples.
- The researcher can also increase his sample size with this technique.

Disadvantages

- From all the different type of probability sampling, this technique is the least representative of the population. The tendency of individuals within a cluster is to have similar characteristics and with a cluster sample, there is a chance that we can have an over-represented or under-represented cluster which can skew the results of the study.
- This is also a probability sampling technique with a possibility of high sampling error.

The cluster sampling method is most suitable when population groups are separated and access to all is difficult, for example, in many distant cities.

Multi-stage Sampling

Sometimes it is possible that the population you wish to investigate is too large and scattered. In situations like this, it is nearly impractical for you to make a list of the entire population from which to draw a SRS. For instance, when an Election Commission samples voters in a given country, it does not do a SRS. Since districts compile voter lists, the Commission might first do a sample of the districts and then sample within the selected districts. This illustrates two stages (strata). In some instances, the Commission might use even more stages (strata). At each stage, it might do a stratified random sample on sex, race, income level or any other useful variable on which the Commission could get information before sampling.

Multi-stage sampling is a sampling technique that breaks the sampling process into several steps. To understand this method of sampling, let us look the following example.

Example

A country may be considered as divided into a 'number of districts'; each district into a 'number of villages' and each village into a 'number of farms'. In multi-stage sampling, a number of districts are selected in the first stage, within each such selected district a number of villages are selected in the second stage and from each selected village, a number of farms are selected at the third stage for enquiry. All the farms which are final units of sampling are surveyed.

The advantages of multi-stage sampling are convenience, economy and efficiency. Multi-stage sampling does not require a complete list of members in the target population, which greatly reduces sample preparation cost. The list of members is required only for those clusters used in the final stage.

The main disadvantage of multi-stage sampling is the same as for cluster sampling: lower accuracy due to higher sampling error.

Non-probability Sampling

Non-probability sampling, a valuable group of sampling technique, is commonly used in qualitative, mixed methods and even quantitative research designs. In this sampling, the probability of any particular element of the population being chosen is unknown and the basis of selection of units is quite arbitrary because researchers rely heavily on personal judgement.

It should be noted that the researcher cannot measure random sampling error in non-probability sampling, as there are no appropriate statistical techniques available in this method. Thus, projecting the data beyond the sample is statistically inappropriate. In other words, in non-probability sampling designs, the elements in the population do not have any probabilities attached to their being chosen as sample subjects. This means that the researcher cannot confidently generalise the findings of the study of the sample to the population. Nevertheless, there are occasions when non-probability samples are best suited for the researcher's purpose. These possibilities of using non-probability sampling techniques have been discussed further.

When to Use Non-probability Sampling

The method can be used when

- The aim is to show that a particular trait exists in the population
- The aim is to do a qualitative, pilot or exploratory study
- It is impossible to do randomisation due to infinite size of the population
- It is unlikely to generate results that will be used to create generalisations pertaining to the entire population
- The resources (budget, time and workforce) available are limited
- The aim is to significantly diminish the potential for researchers to study certain types of population, such as those populations that are hidden or hard to reach (e.g., drug addicts, prostitutes), where a list of the population simply does not exist. Here, 'snowball' sampling, a type of non-probability sampling technique, provides a solution (Benz 2008)

Types of Non-probability Sampling

There are five types of non-probability sampling techniques that you may use when doing your research:

- Convenience sampling
- Purposive or judgemental sampling
- Quota sampling

- Self-selection sampling
- Snowball sampling

Convenience Sampling

Convenience sampling (also called haphazard or accidental sampling) is a non-probability sampling technique. The sample is chosen on the basis of the convenience of the investigator. Often the respondents are selected because they are at the right place at the right time. A convenience sample is simply one where the units that are selected for inclusion in the sample have the easiest access.

In view of convenience and economical, this method of sampling, for example, is best suited to sample employees in companies in a nearby area, sample from a pool of friends and neighbours in clinical research. The person-on-the street type of interview conducted by TV programmes is another example. TV interviewers visit the street with camera and microphone, and talk to those people who are convenient and easily accessible for interview. They select people who look 'normal' to them and avoid people who are unattractive, poor, very old or inarticulate.

Another example of convenience sample is that of a newspaper that asks the readers to cut a questionnaire from the newspaper, fill it and mail. Based on our daily observations, we know that not everyone reads the newspaper, has an interest in the topic or will take the time to cut out the questionnaire and mail it. Those who will do so, their number may appear large, but the researcher cannot use this sample for purpose of an accurate generalisation to the population. It means that convenience samples are least reliable but normally the cheapest and easiest to conduct. The method is often used during the exploratory phase of a research project. It is possibly the best way of getting speedily and efficiently some basic information. Often such sample is taken to test ideas or even to get ideas about a subject of interest.

Example

Suppose we have 10,000 university students, and we are only interested in achieving a sample size of, say, 100 students. We may simply stand at one of the main entrances to campus, where it would be easy to invite the many students that pass by to take part in the research.

Advantages and Disadvantages of Convenience Sampling

Advantages

- It is very easy to handle as the method is based on limited number of rules that govern the sample selection.
- Cost and time involved are relatively small as compared to other sampling techniques. This enables the researcher to achieve the sample size he or she wants in a fast and inexpensive way.
- The method may help the researcher in gathering useful data and information that would not have been possible using probability sampling techniques, which require more formal access to lists of populations.

Disadvantages

- The method can lead to the under-representation or over-representation of particular groups within the sample.
- The sample is unlikely to be representative of the population being studied. This undermines the researcher's ability in generalising the findings to the population he or she is studying.

Purposive Sampling

Purposive sampling is more commonly known as judgemental sampling. In this type of sampling, the researcher chooses subjects (e.g., people, cases/organisations, events, pieces of data) as part of the sample with a specific purpose in mind. He or she believes that some subjects are more suitable (fit) for the research compared to other individuals. In other words, the researcher chooses purposively the participants for his or her study.

A purposive sample is non-representative subset of some larger population. It is developed to serve a very specific need or purpose. It normally targets a particular group of people.

The researcher lays down the criteria for the subjects to be chosen for the sample which in turn depends upon the type of topic of study. Whosoever fulfils that criterion could be the subject of the sample. The selection of the subjects can be done by the researcher himself or herself, or he or she might provide the criteria to somebody else and leave it to his or her judgement for the actual selection of the subjects. That is why such a sample is also called as judgemental or expert opinion sample. Thus, a purposive sampling starts with a purpose in mind and the sample is thus selected to include people of interest and excludes those who do not suit the purpose.

Example

You want to get opinions from non-working mothers. You go around an area knocking on doors during the day when children are usually to be at school. You speak to the 'woman of the house'. You start asking them about whether there are children and whether the woman has a day job. For this study of working women, you may lay down the subject selection criteria such as the following: the lady is married, has two children, one of her children is school going age and is living in a joint family.

Example

A researcher is interested in studying students on the basis of the following criteria: (a) students are enrolled in a course on research methods, (b) they are highly regular, (c) they are frequent participants in the class discussions and (d) they often come to class with new ideas. He or she may do this job himself or herself, or may ask another teacher of this class to select the students following the above criteria. In the latter situation, the researcher is giving it to the judgement of the teacher to select the subjects.

You can use this method when you do not have access to sufficient people who possess the criteria you are seeking for your sampling. The method is popular with newspapers and magazines seeking to make a particular point. It is equally useful for marketing researchers who are looking for support for their product(s).

Advantages and Disadvantages of Purposive Sampling

Advantages

Purposive Sampling Technique:

- is both less costly and less time-consuming
- ensures adequate presentation
- prevents entry of unnecessary and irrelevant items into the sample per chance
- allows an in-depth analysis of selected items
- offers better results if the researcher is unbiased, a keen observer and has the ability for making judgement

Disadvantages

- no equal chance for selection of all subjects
- it is 'non-probability' sampling techniques and hence can be subject to bias and error
- indeterminate in suggesting the degree of accuracy of the investigation

Quota Sampling

Quota sampling ensures that certain characteristics of a population sample will be represented to the exact extent that the researcher desires. This technique demands the researcher first to identify relevant categories of people (e.g., male and female; or under the age of 30; ages 30 to 60; over 60 and so on) only then decide how many subjects to be included in each category. Thus, the number of people in various categories of the sample remains fixed.

Example

For example, the researcher decides to select 10 males and 10 females underage 30; 15 males and 15 females aged 30 to 60 and 5 males and 5 females overage 60 for a sample of 60 persons. This is quota sampling.

Example

In a sample of 100, the researcher wishes to have 40 per cent men and 60 per cent women in the sample. He or she would stop when 40 men are recruited, that is, 'quota' for men is over.

Quota sampling is a form of proportionate stratified sampling. It samples a predetermined proportion of people from different groups, but on a convenience basis. The main purpose is to select a representative sample and/or to allow subgroup analyses.

Apparently, stratified and quota sampling look similar but in reality there are certain differences between them. In stratification method, subjects are selected on the basis of simple random sampling techniques. It uses callbacks for the identification of particular subject. Stratified sampling without callbacks may not, in practice, be much different from quota sampling. In quota sampling, interviewers select first available subject who meets the inclusion criteria and hence it is considered as convenience sampling. Further, for stratified sampling, a sampling frame is a prerequisite but not for the quota sampling. A significant difference is that the stratified sampling uses probability sampling and hence allows estimation of sampling error, which is not possible with quota samples.

The major advantages of quota sampling are speed of data collection, lower costs and convenience. Quota sampling becomes necessary when a subset of a population is under-represented and may not get any representation if equal opportunity is provided to each. There are many problems with quota sampling, yet a careful supervision of the data collection may provide a representative sample of the various subgroups within the population.

In quota sampling, the aim of the researcher is to represent the major characteristics of the population by sampling a proportional amount of each. Let us say, for example, that you want to obtain a proportional quota sample of 100 people based on sex. First, you will have to find out the proportion of the population, that is, men, and the proportion, that is, women. If you found out that 40 per cent were women and 60 per cent were men, then you would need a sample of 40 women and 60 men for a total of 100 respondents. You would start sampling and continue until you got those proportions and then you would stop. So if you have already got 40 women for the sample, but not 60 men, you would continue to sample men and discard any legitimate women respondents that came along. You do not need them because you have already 'met your women quota'. The difficulty here is that you have to decide in advance the specific characteristics on which you will base the quota. Will it be by gender, age, education, race, religion and so on?

To create a quota sample, there are three steps:

- Selecting the relevant stratification and dividing the population accordingly
- Computing a quota for each stratum
- Continuing to invite and include cases until the quota for each stratum is met

Advantages

- It is much quicker and easier to carry out because it does not require a sampling frame and the strict use of random sampling techniques. This makes it popular in undergraduate and master's level dissertations where there is a need to divide the population being studied into strata (groups).
- It improves the representation of particular strata (groups) within the population and allows us to more easily compare these groups (strata).
- Costs are moderate.
- It can be extensively used, understood and does not need a list of population elements.

Disadvantages

- It is impossible to determine the possible sampling error.
- It is not possible to make statistical inferences from the sample to the population. This can lead to problems of generalisation.
- Extending the sampling requirements can increase costs and time to carry out the research.

Example

Suppose we want to understand more about the career goals of students at University X. In particular, we want to look at the differences in career goals among fresh students, second year students, junior students and senior students to examine how career goals might change over the course of a college education. University has 20,000 students, which is our population. Next, we need to find out how our population of 20,000 students is distributed among the four class categories that we are interested in. If we discover that there are 6,000 fresh students (30%), 5,000 first-year students (25%), 5,000 second-year students (25%) and 4,000 final-year students (20%), this means that our sample must also meet these proportions. If we want to sample 1,000 students, this means that we must survey 300 fresh students, 250 first-year students, 250 second-year students and 200 final-year senior students. We would then continue to randomly select these students for our final sample.

Self-selection Sampling

This technique is self-explanatory. It is respondents who themselves decide whether to take part in the survey or not. Self-selection sampling is appropriate when the researcher wants to allow units or cases, whether individuals or organisations, to choose to take part in research on their own accord.

As a sample strategy, self-selection sampling can be used with a wide range of research designs and research methods. For example, survey researchers may put a questionnaire online and subsequently invite anyone within a particular organisation to take part. Similarly, scientists that conduct experiments using human subjects may advertise the need for volunteers to take part in drug trials or research on physical activity. The key element is that research participants (or organisations) volunteer to take part in the research on their own accord. The researcher does not approach them directly.

The self-selection sample involves two simple steps:

- Advertising your need for units (or cases).
- Verifying and checking the relevance of units (or cases) and either inviting or rejecting them.

You need to let potential participants or organisations know about your study. For this, you will have to do some kind of advertising (publicity) or promotion. You can use print media, the radio, an online notice board or some other medium. The invitation should contain certain ethical guidelines clearly explaining the purpose of the study and also more practical information on items such as the types of applicants that are required (e.g., age, gender or some other more subject-specific criteria).

Step 1
Imagine that a researcher wants to undertake an investigation to understand more about the career goals of students at a particular university. Let us say that the university has roughly 10,000 students. These 10,000 students are our 'population (*N*)'. Each of the 10,000 students is known as a unit. For selecting a 'sample (*n*)' of students from this population of 10,000 students, we can choose self-selection sampling technique. Let us imagine that because we have a small budget and limited time, we choose a sample size of 100 students. However, it is important to see that each of the students is in their final year of university. Therefore, you would need to ensure that only final-year students take part in the research. To publicise this, the study could run an advertisement on student radio, the student newspaper or on physical and online notice boards accessed by students at the university.

Step 2
It is not necessary that all applicants be relevant for the study. Some of them may have not read or understood the purpose of the study. Furthermore, some of them may not be the type of applicants the researcher is looking for. For instance, students that are not in their final year at the university may still choose to apply to take part in the study. Issues such as these require your particular attention and a complete check before any particular unit or case, whether an individual or organisation, is invited to become part of the sample.

Advantages

- This can reduce the amount of time necessary to search for appropriate units (or cases), that is, those individuals or organisations that meet the selection criteria needed for your sample.
- The potential units or cases (individuals or organisations) are likely to be committed to take part in the study, which can help in improving attendance (where necessary) and greater willingness to provide more insight into the phenomenon being studied (e.g., a respondent may be more willing to spend the time filling in qualitative, open-ended questions in an online survey, where others may leave them blank; Acharya et al. 2013).

Disadvantages

- There is likely to be a degree of self-selection bias. For example, the decision to participate in the study may reflect some inherent bias in the characteristics/traits of the participants (e.g., an employee with a 'chip of his shoulder' wanting to give an opinion).
- This can either lead to the sample not being representative of the population being studied or exaggerating some particular findings from the study.

Snowball Sampling

Snowball sampling (also called network, chain referral or reputational sampling) is usually done when there is a very small population size or when the population you are interested in is hidden and/or hard to reach. You can use this technique when you do not have access to sufficient people with the characteristics you are seeking. These include populations such as drug addicts, homeless people, individuals with HIV/AIDS, prostitutes, unemployed people, minority ethnic residents and so on.

In this type of sampling, the researcher initially contacts a few potential respondents, interviews them and then asks if they know of anybody else with the same characteristics he or she is looking for. A serious limitation of using a snowball sample is that it is hardly representative of the population.

Example

A researcher is studying traits of environmental engineers but can only find five. He asks these five engineers if they know any more. They give him or her several further references who in turn provide additional contacts. In this way, he or she manages to contact sufficient engineers.

Example

A researcher is examining friendship networks among teenagers in a community. He or she begins with three teenagers who do not know each other. Each teen names four close friends. The researcher then goes to the four friends and asks each to name four close friends, then goes to those four and does the same thing again and so on. Before long, a large number of people are involved. Each person in the sample is directly or indirectly tied to the original teenagers, and several people may have named the same person. The researcher eventually stops, either because no new names are given, indicating a closed network, or because the network is so large that it is at the limit of what he or she can study.

The 'snowball' effect takes place with each multiplication of referrals. For example, if the researcher got two referrals from each person, then he or she starts with two people then he or she gets four more, then eight, sixteen and so on.

The way that the sample is chosen by target people makes it liable to various forms of bias. People tend to associate not only with people with the same study selection characteristics but also with other characteristics. This enhances the chance of correlations being found in the study that do not apply to the generalised wider population.

The need to get the person who can give you a reference also implies that the researcher has to form a relationship with the person and be nice to them. This can change the study results as affective biases in both the researcher and the target person change how they think and behave.

How Can You Determine the Sample Size?

When we ask 'how many survey respondents do we need?', what we are really asking is, 'how big does our sample need to be in order to accurately estimate our population or generalise about populations from data collected using any probability sample?' This means: 'how can we determine the sample size for our study?' The answer to this question involves a complete knowledge of statistical methods commonly used in sampling theory.

Before explaining the calculation process to determine the sample size, it is essential to understand the key factors which govern the choice of the sample size. The choice is governed by:

- The confidence we need to have in your data—that is, the level of certainty that the data collected will need to represent the characteristics of the total population;
- The margin of error that we can tolerate—that is, the accuracy we require for any estimates made from our sample;
- The types of analyses we are going to undertake—in particular, the number of categories into which we wish to subdivide our data, as many statistical techniques have a minimum threshold of data cases for each cell (e.g., chi-square) and to a lesser extent: the size of the total population from which our sample is being drawn.

Researchers normally work to a 95 per cent level of certainty (confidence level). This means that if our sample was selected 100 times, at least 95 of these samples would be certain to represent the characteristics of the population. The confidence level states the precision of our estimates of the population.

This being said and knowing the fact that the determination of an ideal size of sample from the population is a complex process involving several concepts and mathematical equations, we explain below a reasonably simple way to help you calculate your ideal sample size. The calculation process has been explained by considering the following five steps to make sure that your sample accurately estimates your population.

Step 1: Knowing your population
We have noticed earlier that by population we mean the entire set of people whom we want to understand. Our sample will be drawn from this population that ends up actually taking our survey.

So, for example, a publisher wants to know how to sell a new chemistry textbook to students in secondary schools of Samoa, our population would be all secondary school students learning chemistry in Samoa. If we were trying to understand how many holidays the secondary teachers who in Samoa would like to have, our population would be all secondary school teachers in Samoa.

The vital first step for you is to figure out what population you are trying to understand. Once you know what your population is, figure out how many people (roughly) make-up that population. For example, roughly 190,000 people live in Samoa and we are guessing that the education system of Samoa has fewer secondary school teachers than that.

Step 2: Degree of Accuracy
It means how much of a risk you are willing to tolerate for the responses you will get for your survey due to the fact that you are not surveying your entire population. So here are your two questions to answer:

Question 1: How sure do you need to be that the answers reflect the views of your population?

This means the limit of your 'margin of error'. Suppose that 90 per cent of your sample likes KFC chicken burger. A 5 per cent margin of error would mean +5 per cent on one side of 90 per cent (90% + 5% = 95%) and –5 per cent on other side of 90 per cent (90% – 5% = 85%), meaning that actually 85–95 per cent of your sample likes KFC chicken burger. The most commonly used margin of error is 5 per cent. You may select margin of error anywhere from 1–10 per cent depending on your survey. However, increasing your margin of error beyond 10 per cent is not recommended.

Question 2: How sure do you need to be that the sample accurately samples your population?

This question relates directly to your 'confidence level'. A confidence level is the likelihood that the sample you picked mattered in the results you got. This is how it is calculated. If you picked 30 more samples randomly from your population, how often would the results you got in your one sample significantly differ from those other 30 samples? A 95 per cent confidence level means that you

TABLE 6.1 Sample Size

	Margin of Error			*Confidence Level*		
Population	*10%*	*5%*	*1%*	*90%*	*95%*	*99%*
100	50	80	99	74	80	88
500	81	218	476	176	218	286
1,000	88	278	906	215	278	400
10,000	96	370	4,900	264	370	623
100,000	96	383	8,763	270	383	660
1,000,000+	97	384	9,513	271	384	664

Source: Authors.
Note: These are intended as rough guidelines only. Also, for populations of more than 1 million you might want to round up slightly to the nearest hundred.

would get the same results 95 per cent of the time. The most commonly used confidence level is 95 per cent. You may use a 90 per cent or 99 per cent confidence level depending on your survey. But a level of confidence below 90 per cent is not recommended.

Step 3: How big a sample do you need?
Table 6.1 will help you to determine the sample size of your study for different levels of margin of error or confidence level. Suppose that the total size of population under consideration is 100. For the margin of error of the order of 10 per cent, 5 per cent and 1 per cent, the sample size would be 50, 80 and 99, respectively. The corresponding sample sizes at confidence level of 90 per cent, 95 per cent and 99 per cent would be 74, 80 and 88, respectively. First thing that you have to do is to select your approximate target population and then choose your margin of error to estimate the number of completed surveys you will require. The computation method is described in Box 6.2.

BOX 6.2

Minimum Sample Size

$$n = a \times b \times \left(\frac{c}{d}\right)^2$$

Where
n = is the minimum sample size
a = people who responded to survey request
b = people who did not respond survey request
c = z score, i.e., level of confidence
d = margin of error (that can be tolerated)

and

Adjusted Sample Size

$$n' = \frac{n}{1 + \left(\frac{n}{N}\right)}$$

Where
n' = is the adjusted minimum sample size
n = is the minimum sample size
N = is the total (target) population

Example
To answer a research question, Wang needs to estimate the proportion of a total population of 4,000 teachers who were contacted by their principals at least once a month. Wang decides that he needs to be 95 per cent certain (confident) that his 'estimate' is accurate (the level of confidence in the estimate); this corresponded to z score of 1.96. He also decides that his 'estimate' needs to be accurate within plus or minus 5 per cent of the true percentage (the margin of error that can be tolerated).

In order to calculate the minimum sample size, Wang still needs to estimate the proportion of respondents (teachers) who were contacted by their principals at least once a month. From his pilot survey, he discovers that 12 out of the 30 teachers were contacted at least once a week—in other words, that 40 per cent belongs to this specified category. This meant that 60 per cent did not respond to survey request.

Wang substitutes these figures into the formula:

$$\begin{aligned} \text{n} &= 40 \times 60 \times (1.96/5)^2 \\ &= 2,400 \times (0.392)^2 \\ &= 2,400 \times 0.154 \\ &= 369.6 \text{ or } 370 \end{aligned}$$

His minimum sample size therefore would be 370 returns.

As the total population of teachers was 4,000, Wang could now calculate the 'adjusted minimum sample' size.

$$n' = \frac{369.6}{1 + \left(\frac{369.6}{4,000}\right)}$$

$$= \frac{369.6}{1 + 0.092}$$

$$= 369.6 / 1.092$$

$$= 338.46$$

Because of the small total population, Wang needs a minimum sample size of only 339 teachers. However, this assumes that he has a response rate of 100 per cent.

Where our population is less than 10,000, a smaller sample size can be used without affecting the accuracy. This is called 'adjusted sample size'. It is calculated using the following formula:

$$n' = \frac{n}{1 + \left(\frac{n}{N}\right)}$$

Source: Authors.

Using these equations, the sample sizes of various levels of population have been calculated. These are given in Table 6.2.

Table 6.2 shows the sample size for population ranging from 10 to 300,000,000 at 95 and 99 per cent level of confidence with different levels of margin of errors. Here Equation 2.1 has been used in calculation of these samples.

TABLE 6.2 Population and Required Sample Size

	Confidence = 95%				*Confidence = 99%*			
	Margin of Error				*Margin of Error*			
Population Size	*5.0%*	*3.5%*	*2.5%*	*1.0%*	*5.0%*	*3.5%*	*2.5%*	*1.0%*
10	10	10	10	10	10	10	10	10
20	19	20	20	20	19	20	20	20
30	28	29	29	30	29	29	30	30
50	44	47	48	50	47	48	49	50
75	63	69	72	74	67	71	73	75
100	80	89	94	99	87	93	96	99
150	108	126	137	148	122	135	142	149
200	132	160	177	196	154	174	186	198
250	152	190	215	244	182	211	229	246
300	169	217	251	291	207	246	270	295
400	217	265	318	384	250	309	348	391
500	234	306	377	475	285	365	421	485
600	248	340	432	565	315	416	490	579
700	260	370	481	653	341	462	554	672
800	278	396	526	739	363	503	615	763
1,000	291	440	606	906	399	575	727	943
1,200	306	474	674	1,067	427	636	827	1,119
1,500	322	515	759	1,297	460	712	959	1,376
2,000	333	563	869	1,655	498	808	1,141	1,785
2,500	346	597	952	1,984	524	879	1,288	2,173
3,500	357	641	1,068	2,565	558	977	1,510	2,890
5,000	365	678	1,176	3,288	586	1,066	1,734	3,842
7,500	370	710	1,275	4,211	610	1,147	1,960	5,165
10,000	378	727	1,332	4,899	622	1,193	2,098	6,239

(Continued)

(Continued)

Population Size	Confidence = 95% Margin of Error 5.0%	3.5%	2.5%	1.0%	Confidence = 99% Margin of Error 5.0%	3.5%	2.5%	1.0%
25,000	381	760	1,448	6,939	646	1,285	2,399	9,972
50,000	382	772	1,491	8,056	655	1,318	2,520	12,455
75,000	383	776	1,506	8,514	658	1,330	2,563	13,583
100,000	384	778	1,513	8,762	659	1,336	2,585	14,227
250,000	384	782	1,527	9,248	662	1,347	2,626	15,555
500,000	384	783	1,532	9,423	663	1,350	2,640	16,055
1,000,000	384	783	1,534	9,512	663	1,352	2,647	16,317
2,500,000	384	784	1,536	9,567	663	1,353	2,651	16,478
10,000,000	384	784	1,536	9,594	663	1,354	2,653	16,560
100,000,000	384	784	1,537	9,603	663	1,354	2,654	16,584
300,000,000	384	784	1,537	9,603	663	1,354	2,654	16,586

Source: Authors.

Step 4: How responsive will people be?

We know that everyone to whom you sent the survey questionnaire may not respond to your request for the collection of information. The percentage of people who actually respond to your survey is considered the response rate. The response rate helps you determine the total number of surveys you will need to send out to obtain the required number of completed surveys. A high survey response rate helps to ensure that the survey results are representative of the survey population.

Response rates vary widely and depend on factors such as your relationship with the target audience, duration of survey and complexity, incentives and topic of your survey. If you were conducting online surveys where you do not have any prior relationship with the respondents, an acceptable response rate would be between 20–30 per cent. A response rate of 10–15 per cent is a more conservative and a safer guess if you have not surveyed your population before.

Acceptable response rates vary by how you administer the survey. Researchers have fixed the following norms:

- Mail: 50 per cent adequate, 60 per cent good, 70 per cent very good
- Phone: 80 per cent good
- E-mail: 40 per cent average, 50 per cent good, 60 per cent very good
- Online: 30 per cent average
- Classroom paper: Greater than 50 per cent good
- Face-to-face: 80–85 per cent good

The calculation method of response rate for your survey is given in Box 6.3.

BOX 6.3: Computation of Response Rate

The most common reason for non-response is the respondent's refusal to answer all the questions without giving a reason. You can minimise such non-response by selecting carefully the methods of data collection. Alternatively, some of your selected respondents may not meet your research requirements and thus will be ineligible to respond. Non-location and non-contact create further problems; the fact that these respondents are unreachable means they will not be represented in the data you collect.

Saunders, Lewis and Thornhill (2009) suggest that when you calculate this you should include all eligible respondents. This he calls the total response rate.

Total Response Rate

$$\text{Total response rate} = \frac{\text{total number of responses}}{\text{total number in sample} - \text{ineligible cases}}$$

A more common way of doing this excludes ineligible respondents and those who, despite repeated attempts, were unreachable. This is known as the active response rate:

Active Response Rate

$$\text{Active response rate} = \frac{\text{total number of responses}}{\text{total number in sample} - (\text{ineligible} + \text{unreachable})}$$

Example

MS Lin, Director Education, Nanjing Province, had decided to administer a telephone questionnaire to teachers who had left primary schools over the past five years. From her records, she found a list of 1,034 primary school teachers who had left over this period (the total population) and selected a 50 per cent sample (i.e., 517 teachers) randomly. Unfortunately, she could obtain current telephone numbers for only 311 of the 517 ex-teachers who made up her total sample. Of these 311 ex-teachers who were potentially reachable, she obtained a response from 147. In addition, her list of teachers who had left primary schools was inaccurate, and nine of those she contacted were ineligible to respond, having left primary schools over five years earlier.

$$\text{Her total response rate} = 147/517 - 9 = 147/508 = 28.9\%$$
$$\text{Her active response rate} = 147/311 - 9 = 147/302 = 48.7\%$$

Calculation of Actual Sample Size

Actual Sample Size

$$n^a = n \times 100 / r^e\%$$

where
n^a = the actual sample size required
n = the minimum (or adjusted minimum) sample size
r^e = the estimated response rate expressed as a percentage

Example

Kumar was a part-time student employed by Nike. He had decided to send a questionnaire to Nike's customers and calculated that an adjusted minimum sample size of 439 was required. Kumar estimated the response rate would be 30 per cent. From this, he could calculate his actual sample size as follows:

$$
\begin{aligned}
n^a &= 439 \times 100 / 30 \\
&= 43,900 / 30 \\
&= 1,463
\end{aligned}
$$

Kumar's actual sample, therefore, needed to be 1,463 customers. The likelihood of 70 per cent non-response means that Kumar needed to include a means of checking that his sample was representative when designing his questionnaire.

Source: Adapted from Saunders, Lewis and Thornhill (2009).

Step 5: So how many people do I send it to?

Just divide the number you got from Step 3 by the number you got from Step 4. That is your sample size—the magic number.

Now that you have your numbers from Step 1 and Step 2, check out from Tables 6.1 and 6.2 above and figure out how big a sample you will need.

Example

Suppose you want a sample of 100 girls whose performance in mathematics is below the district average to fill out your survey. You think that approximately 10 per cent of these low performing girls that you send your survey will actually fill the survey. Then you need to send the survey to 100/10 per cent girls, that is, 1,000 girls.

This is how you can calculate the size of your sample. However, you should keep in mind that there is no specific sample size for a research study. How big or small the size of the sample would be depends on the type of study being conducted and the population being studied. The following is a list of examples to determine sample sizes for different kinds of research studies.

Experimental/Causal Research

Most researchers recommend that the sample size for each group in an experimental study be at least 30 participants. In some cases, the group could be as small as 15 if tight controls are used in establishing the research groups.

Correlational Studies

The recommended number for relationship studies is also 30. Smaller numbers make it difficult to obtain statistical significance.

Descriptive Studies

The number of participants in a descriptive study can vary significantly. Usually the size of the population to be studied has more of an effect on the sample size than any general sampling rule. Small populations require a large percentage of the population to be included in the sample.

The general rules you have to follow while determining the sample size are as follows:

- A smaller percentage is required for a larger population.
- Studies with a population of 100 or less should use the entire population.
- Population of approximately 300 requires a sample size of 50 per cent.
- Population of 100,000 and over would require only 384 in the sample.

Summary

We have noted in this chapter that the topic of sampling techniques and sampling design is a very involved one where there are few singly right solutions. Your decision to select a particular technique for designing field studies will be governed by several factors such as the overall assessment/monitoring objectives, sampling objectives, the population being sampled, underlying variability and so on. The best approach to sample design is to couple a thorough understanding of the system you are sampling with a good foundation in basic sampling concepts and different sampling design methods. This chapter described various ways samples are gathered and discussed the importance of reducing the effects of bias in sampling as well as non-random sample techniques.

So far, we have explained the several types of research methodologies educational researchers and social scientists use to study a variety of issues, no matter whether the research is very informal, simple process or it is a formal, somewhat a sophisticated process. We have also explained the several sampling techniques in order to enable you determine an acceptable sample size for your study. Regardless of the type of process, all research begins with a generalised idea in the form of a research question or a hypothesis. Chapter 6 describes the standard procedures to help you find out a research topic, identify a research question and formulate the hypothesis.

Self-test Exercise

Exercise 6.1: Suppose there are $N = 850$ students in a school from which a sample of $n = 10$ students is to be taken.

6.1.1. How will you select a sample of 10 students out of 850 students using random number table?

Exercise 6.2: Imagine that District A of Country X employs a total of 2,700 primary school teachers. The District Education Officer (DEO) decides to ask the teachers for suggestions on how to improve their workplace. It would take too long to survey everyone, so the DEO chooses to systematically sample 300 of the teachers.

6.2.1. What would be the sampling interval?

6.2.2. If the number 8 is your first randomly drawn number, what would be the first 5 numbers of your sample?

Exercise 6.3: A researcher who wanted to determine the benefits of using a new beginning algebra study technique obtained permission from a school district to select 50 high school students. The researcher selected 50 beginning algebra students at random. The researcher selected 25 of these 50 students to participate in the new study programme. The researcher gave a training session on traditional study techniques to the other 25 students and asked them to use these methods.

6.3.1. Which is the most likely target population in this study?

1. algebra students in the district
2. all students in the district
3. all algebra students
4. the 25 students who learned the new study techniques

Exercise 6.4: When researchers have a limited amount of time to collect data, they use convenience sampling because it is often:

1. Unbiased
2. Nearby
3. Inexpensive
4. Fast

Exercise 6.4: How do cluster sampling and stratified random sampling differ?

1. Stratified random sampling requires selection from each group proportional to their share of the population, but cluster sampling does not.
2. Stratified random sampling allows groups to mix together, but cluster sampling does not.
3. Cluster sampling allows groups to mix together, but stratified random sampling does not.
4. In cluster sampling, the sample must match the proportion of the groups in the population, but in stratified random sampling, this is not necessary.

Exercise 6.5: You hypothesise that students in brotherhood/sisterhood (fraternities/sororities) on your university campus are more extroverted than other students. To answer your research question, you obtained a sample of 'Arab' students and a sample of 'independent' students.

6.5.1. For each of these two populations of students, how would you obtain SRSs using the four steps described in this chapter?

6.5.2. Describe how would you obtain samples using stratified random sampling, with stratification based on the sex of the student. Assume that 60 per cent of the Arab students are females.

References

Acharya, A., A. Prakash, P. Saxena and A. Nigum. 2013. 'Sampling: Why and How of it?' *Indian Journal of Medical Specialities* 4 (2).

Benz, C. 2008. *Mixed Methods Research: Exploring the Interactive Continuum*. Carbondale, IL: Southern Illinois University Press.

Henry, G.T. 1990. *Practical Sampling*, vol. 21. Thousand Oaks, CA: SAGE Publications.

Horner, M.W. and M.E. O'Kelly. 2005, July. 'A Combined Cluster and Interactive Model: The Hierarchical Assignment Problem'. *Geographical Analysis* 37 (3): 315–335.

Saunders, N.K., P. Lewis and A. Thornhill. 2009. *Research Methods for Business Students*. Harlow: Pearson Education Limited.

Appendix 6.1. Random Number Table

Random Number Table

13,962	70,992	65,172	28,053	02,190	83,634	66,012	70,305	66,761	88,344
43,905	46,941	72,300	11,641	43,548	30,455	07,686	31,840	03,261	89,139
00,504	48,658	38,051	59,408	16,508	82,979	92,002	63,606	41,078	86,326
61,274	57,238	47,267	35,303	29,066	02,140	60,867	39,847	50,968	96,719
43,753	21,159	16,239	50,595	62,509	61,207	86,816	29,902	23,395	72,640
83,503	51,662	21,636	68,192	84,294	38,754	84,755	34,053	94,582	29,215
36,807	71,420	35,804	44,862	23,577	79,551	42,003	58,684	09,271	68,396
19,110	55,680	18,792	41,487	16,614	83,053	00,812	16,749	45,347	88,199
82,615	86,984	93,290	87,971	60,022	35,415	20,852	02,909	99,476	45,568
05,621	26,584	36,493	63,013	68,181	57,702	49,510	75,304	38,724	15,712
06,936	37,293	55,875	71,213	83,025	46,063	74,665	12,178	10,741	58,362
84,981	60,458	16,194	92,403	80,951	80,068	47,076	23,310	74,899	87,929
66,354	88,441	96,191	04,794	14,714	64,749	43,097	83,976	83,281	72,038
49,602	94,109	36,460	62,353	00,721	66,980	82,554	90,270	12,312	56,299
78,430	72,391	96,973	70,437	97,803	78,683	04,670	70,667	58,912	21,883
33,331	51,803	15,934	75,807	46,561	80,188	78,984	29,317	27,971	16,440
62,843	84,445	56,652	91,797	45,284	25,842	96,246	73,504	21,631	81,223
19,528	15,445	77,764	33,446	41,204	70,067	33,354	70,680	66,664	75,486
16,737	01,887	50,934	43,306	75,190	86,997	56,561	79,018	34,273	25,196
99,389	06,685	45,945	62,000	76,228	60,645	87,750	46,329	46,544	95,665
36,160	38,196	77,705	28,891	12,106	56,281	86,222	66,116	39,626	06,080
05,505	45,420	44,016	79,662	92,069	27,628	50,002	32,540	19,848	27,319
85,962	19,758	92,795	00,458	71,289	05,884	37,963	23,322	73,243	98,185
28,763	04,900	54,460	22,083	89,279	43,492	00,066	40,857	86,568	49,336
42,222	40,446	82,240	79,159	44,168	38,213	46,839	26,598	29,983	67,645
43,626	40,039	51,492	36,488	70,280	24,218	14,596	04,744	89,336	35,630
97,761	43,444	95,895	24,102	07,006	71,923	04,800	32,062	41,425	66,862
49,275	44,270	52,512	03,951	21,651	53,867	73,531	70,073	45,542	22,831
15,797	75,134	39,856	73,527	78,417	36,208	59,510	76,913	22,499	68,467
04,497	24,853	43,879	07,613	26,400	17,180	18,880	66,083	02,196	10,638

(Continued)

(Continued)

95,468	87,411	30,647	88,711	01,765	57,688	60,665	57,636	36,070	37,285
01,420	74,218	71,047	14,401	74,537	14,820	45,248	78,007	65,911	38,583
74,633	40,171	97,092	79,137	30,698	97,915	36,305	42,613	87,251	75,608
46,662	99,688	59,576	04,887	02,310	35,508	69,481	30,300	94,047	57,096
10,853	10,393	03,013	90,372	89,639	65,800	88,532	71,789	59,964	50,681
68,583	01,032	67,938	29,733	71,176	35,699	10,551	15,091	52,947	20,134
75,818	78,982	24,258	93,051	02,081	83,890	66,944	99,856	87,950	13,952
16,395	16,837	00,538	57,133	89,398	78,205	72,122	99,655	25,294	20,941
53,892	15,105	40,963	69,267	85,534	00,533	27,130	90,420	72,584	84,576
66,009	26,869	91,829	65,078	89,616	49,016	14,200	97,469	88,307	92,282
45,292	93,427	92,326	70,206	15,847	14,302	60,043	30,530	57,149	08,642
34,033	45,008	41,621	79,437	98,745	84,455	66,769	94,729	17,975	50,963
13,364	09,937	00,535	88,122	47,278	90,758	23,542	35,273	67,912	97,670
03,343	62,593	93,332	09,921	25,306	57,483	98,115	33,460	55,304	43,572
46,145	24,476	62,507	19,530	41,257	97,919	02,290	40,357	38,408	50,031
37,703	51,658	17,420	30,593	39,637	64,220	45,486	03,698	80,220	12,139
12,622	98,083	17,689	59,677	56,603	93,316	79,858	52,548	67,367	72,416
56,043	00,251	70,085	28,067	78,135	53,000	18,138	40,564	77,086	49,557
43,401	35,924	28,308	55,140	07,515	53,854	23,023	70,268	80,435	24,269
18,053	53,460	32,125	81,357	26,935	67,234	78,460	47,833	20,496	35,645

Source: From The Rand Corporation, A Million Random Digits with 100,000 Normal Deviates (New York: The Free Press, 1955).

7

CHAPTER

Knowing Our Data

Introduction

Key stakeholders of education—school board members, teachers, parents, alumni and community members—all want to know 'what works' and 'what does not' in terms of educational programmes and innovations. They want to know how the education system is managed and how policies that best suit the needs of the school as well as the overall interests of the students are formulated. They also want to know how schools are managed with a view to create an ideal learning environment for their students and nurturing talents. They want to know all of this so as (a) to be sure that the public outlay is being spent on educational programmes that provide an environment that facilitates continuous improvement, (b) provide a positive return in terms of student progress and (c) to stay informed and aware of trends and the direction of educational development so they know that their school districts are keeping up with the latest practices and programmes.

In order to meet these expectations, researchers carry out thorough investigations. But a key challenge they often encounter is: where and how to start their research? A scientific research inquiry involves identification of a pressing key educational problem research topic, sampling design and frame, a research hypothesis, sources of reading materials and, above all, the time management. All these steps take us to an equally important stage of research—data collection. Thus, access to necessary data is a precondition to conduct research. The researcher cannot analyse the most promising educational problem if he or she does not gain access to relevant data. His or her efforts will be fruitless.

Historically, data collection performed in educational research has its roots in methods developed for studies in the field of psychology, a discipline that took what is termed a 'quantitative' approach. This approach entails the use of instruments such as scales, tests and structured observation and interviewing. During the last half of the twentieth century, disciplines such as anthropology and sociology started to influence educational researchers. The scope of data collection techniques enlarged to include what is now termed as 'qualitative' methods, with an emphasis on narratives, participants' perspectives and less-structured observation and interviewing. With an increasing reliance of modern educational researchers on disciplines such as business, political science, economics and medicine, data collection in education has become a multidisciplinary feature.

Data collection is a broad topic. Most general overviews that claim to cover all or most techniques tend to offer introductory treatments. Few texts, however, provide comprehensive coverage of every data collection technique. Instead, some cover techniques appropriate for either quantitative or qualitative research approaches.

This chapter attempts to highlight the significance and centrality of data in the context of educational research designs. It explores data as a vital element of scholarly enquiry by outlining the role of data in educational research and identifying some issues in data classification and their impact on the design of a proposed research. The chapter details the role of data as key element of research in a scholarly inquiry in an educational environment, challenges faced by the researchers with access to collecting data and its different categories and sources. Finally, the chapter describes how meaningfully the data can be presented on successful and innovative educational practices and also for those practices that do not yield maximum effectiveness, impact and influence.

Data and Statistics in Research

Statistics is the science of collecting, analysing and making inference from data. It is one of the useful branches of mathematics that is not only studied theoretically by expert mathematicians but one that is used by researchers in many fields to organise, analyse and summarise data. Researchers often use statistical methods and analyses to communicate research findings and to support hypotheses and give credibility to research methodology and conclusions. It is, therefore, important for all those involved in research to understand statistics so that they can be informed, evaluate the credibility and usefulness of information and make appropriate decisions.

Difference Between 'Data' and 'Statistics'

Although the importance of data and statistics is well accepted, providing an authoritative definition of these two terms is challenging, as definitions are likely to depend on the context in which the question is asked.

In common parlance, the terms 'data' and 'statistics' are often used interchangeably. However, in the world of research and academia there is an important distinction between data and statistics. In the domain of research, data are the 'raw information' from which 'statistics are created'. In other words, statistics furnish an interpretation and summary data.

Statistics comprises:

- Statistical tables, charts and graphs
- Reported numbers and percentages in a research study

If, for instance, you are searching a number, it means you want a statistic (without 's'). Thus, a statistic answers 'how much' or 'how many'. A statistic replicates or repeats pre-defined observation about reality.

Statistics is usually shown in the form of a table or chart. This is what a statistical table looks like:

Total Adult Illiterates (15+ Years) by Major Geographical Regions in Country X in 2015

Region	*Total Adult Illiterates*	*% Share in Total*
Northern	110,000	19.2
Eastern	140,000	24.3
Central	75,000	13.0
Southern	90,000	15.7
Western	160,000	27.8
Total	**575,000**	**100.0**

Data, on the other hand, refer to:

- Data sets
- Machine-readable data files or computer-readable data files for statistical software programmes

For example, if your aim is to understand a phenomenon, it means you are looking for data. You can analyse and interpret data using statistical procedures or methods to answer 'why' or 'how'. We use data to generate new information and to create new knowledge on a subject.

Raw data are the direct outcome of your research that you have conducted as part of your study or survey. It is the primary source of data. It is usually presented in the form of digital data set that you can analyse using software such as Excel, SPSS, SAS and so on. This is what a data set looks like:

	School Identity (SCID)	*Nature of School (NASC)*	*Location of School (LOSC)*	*Enrolment (ENRO)*
1	1	1	1	900
2	2	2	2	220
3	3	2	2	180
4	4	1	1	700
5	5	1	1	880

NASC: 1 = public school; 2 = private school
LOSC: 1 = urban school; 2 = rural school

The term data should not be confused with information. Data are plain facts. When data are processed, organised, structured or presented in a given context so as to make them useful, they are called 'information'.

It is not enough to have data (such as statistics on the national education system). Data in themselves are fairly useless. The moment these data are 'interpreted' and processed to determine their true 'meaning', they become useful and can be called information. Data are the computer's language. Information is our translation of this language.

Researchers often use the term 'research data'. What does this mean? Research data refers to facts, observations, images, computer programme results, recordings, measurements or experiences on which an argument, theory, test or hypothesis, or another research output is based. Research data are usually numerical, descriptive, visual or tactile. This data may be raw, cleaned or processed data and may be kept in any format or media. Research data include: laboratory notebooks, field notebooks, primary research data (including research data in hard copy or in computer-readable form), questionnaires, audio tapes, videotapes, models, photographs, films, test responses and so on.

In this chapter, we define data as

'A purposive collection of facts, observations or other information related to a particular question, problem or perceived facts'.

Now let us describe the meaning of the term 'statistics'. Basically, the word statistics has three different meanings (sense) which are explained next.

Statistics in Singular Sense

In singular sense, the term statistics has been defined as statistical methods which are used in collection, analysis, interpretation and presentation of numerical data. The statistical methods are used to draw conclusion about the population parameter.

Let us consider the following example. Suppose you want to have a study about the distribution of weights of students in a certain college. First of all, you will collect the information on the weights which may be obtained from the records of the college or you may collect from the students directly. The large number of weight figures might confuse you. In this case, you may arrange the weights in groups such as '40 kg to 50 kg', '50 kg to 60 kg' and so on and find the number of students belonging to each group. This step is called a presentation of data. You may still go further and compute the averages and some other measures which may give you a complete description of the original data.

Statistics in Plural Sense

In plural sense, the term statistics refer to collection of numerical facts and figures in a systematic way with a definite purpose in any field of study. It is in this sense that the educational researchers and public usually think of statistics, say figures concerned with enrolment in primary schools or national population over a period of time or the number of teachers leaving the school system in different years (teachers' attrition) and so on. Percentages, averages and coefficients (secondary statistics) derived from numerical facts are also considered in the term statistics in this sense.

Plural of Word 'Statistic'

The word statistics is used as the plural of the word 'statistic' which refers to a numerical quantity such as mean, median, variance and so on, calculated from sample value. For example, if you select 15 students from a class of 80 students, measure their heights and find the average height. This average would be a statistic.

There are three keywords in the aforementioned definitions that deserve special attention.

Purposive: All that can be observed or otherwise sensed is not necessarily 'data'. An observation is data only within the context of its use.
Collection: Data must be comprised of a set of related observations. Trying to draw conclusions from a single observation would be erroneous.
Perceived facts: The researcher should remember that data are not equal to facts, or, by extension, 'truth'. It is impacted by several considerations.

Definition of Statistics

There are series of definitions but most of the definitions given are not complete and lay emphasis only on some of the aspects of the science. For instance, Bowley (1901) defines 'statistics as the science of counting'. At another place, he is of the view that 'Statistics may rightly be called the science of average'. Both these definitions are defective, as the science of statistics does not confine itself either to counting or to averaging alone. These are no doubt important statistical methods but they do not cover the entire field of the science of statistics.

Croxton and Cowden (1946) give the best definition of statistics. They define statistics: 'as the science of collection, organisation, presentation, analysis and interpretation of numerical data'.

This definition comprises the following stages of statistical inquiry.

Stage 1: Collection of data

This is the foundation of statistical analysis. Within this stage, data are collected from various sources—primary (collected directly) or secondary (available from existing published sources). Primary data collection is one of the most difficult and important tasks faced by researchers and investigators.

Stage 2: Organisation of data

Data collected need to be organised by editing, classifying and tabulating. Editing is done to correct omissions, inconsistencies and erroneous calculation in the survey; classification to arrange the data according to some common characteristics possessed by the item and tabulation to arrange the data in columns and rows for complete clarity as far as presentation of data is concerned.

Stage 3: Presentation

Once data are collected and organised, they should be presented systematically and in an orderly manner to facilitate statistical analysis. The classified data should be presented in such a way that even a non-technical person is able to understand it easily. Tables, graphs and diagrams are the most popular techniques used by the researchers for presenting their data.

Stage 4: Analysis of data

It involves application of a large number of statistical measures to analyse the data. The methods most commonly used are measures of central tendency (measures of the first order) and measures of dispersion (measures of the second order). Among the statistical methods of third order, skewness, kurtosis, correlation, regression and interpolation are the most prominent measures. The analysis of facts based on observation is classified as (a) scientific analysis, (b) numerical analysis and (c) empirical analysis.

Stage 5: Interpretation of data

It implies drawing of conclusion(s) on the basis of data analysed. The interpretation of data is not an easy task. If the analysed data are not properly interpreted, the whole object of the inquiry may be erroneous. It is only correct interpretation which may yield reliable conclusions.

Statistics is, thus, a form of mathematical analysis that uses quantified models, representations and synopses for a given set of experimental data or real-life studies. Statistics studies methodologies to gather, review, analyse and draw conclusions from data.

In simple terms, statistics is a set of concepts, rules and procedures that help us to:

- 'Organise' numerical information in the form of tables, graphs and charts;
- 'Understand' statistical techniques underlying decisions that affect our lives and well-being and
- 'Make' informed decisions

Salient Features of Statistics

Within the context of aforementioned definitions, figures to be considered as a quantitative or numerical data should have the following salient features:

Aggregate of Facts

Single and isolated figures do not constitute statistics because such figures are unrelated and cannot be compared. Statistics refers to a series of figures. For example, if your annual income is US$7,000 or you weigh 75 kg, they would not constitute statistics but a series relating to income or weight of a group of persons is called statistics. Thus, all those figures related to the totality of facts are called statistics. Such figures should also be comparable.

Affected to a Marked Extent by Multiplicity of Causes

Education statistics are not influenced alone by a single variable rather they are affected by a multitude of factors. The obvious reason is that they are extensively used in an educational setting. It, therefore, becomes almost difficult for researchers to study and isolate the effects of any single variable on a phenomenon. Most research studies present combined effects of a set of variables. For example, primary school dropout rates are not affected by a single variable but are jointly determined by a multitude of factors such as poor parental care and attention, poverty, child labour, death of parents, pregnancy among girls and early marriage, teachers' morality and ethics, pupil indiscipline and so on.

Numerically Expressed

All statistics are expressed in numbers. Qualitative statements such as 'population of India is increasing rapidly' does not constitute statistics. Similarly, qualitative statements such as young, old, good, bad and so on are not statistics. For all statistics, a numerical value must be attached. Consider another example. The statement like 'there are 930 females per 1,000 males' must contain figures so that they are called numerical statement of facts. Numerical expressions are thus precise, meaningful and convenient form of communication.

Estimated According to Reasonable Standard of Accuracy

Only precise and accurate numerical statement can be measured or enumerated. But for large number of observations, the figures are just estimated. However, the estimated figures cannot be 100 per cent accurate and precise. In fact, the accuracy depends on the purpose for which statistics are collected. And the standards of accuracy cannot be uniform and will vary from one research enquiry to another. On the other hand, 'enumeration' refers to an 'exact count'. For example, a statement describing 30 students in the English language class is a 100 per cent accurate statement. On the other hand, 'estimation' refers to 'round about figure'—2 million people who listened to the Republic Day address of the prime minister of India or 765 million adult illiterates in the world. These numbers can be few hundred more or less. Statistical results are true only on average.

Collected Systematically According to Plan

For ensuring reliability and accuracy, data should be collected systematically and scientifically. If not, the results would be erroneous and misleading.

Collected for Predetermined Purpose

The purpose of collecting data should be predetermined, specific and well defined otherwise the data will be meaningless and its usefulness will be negligible. Thus, the objective of data collection should be concrete, specific and precise. Suppose, if we want to collect data on the salary scales of teachers, then we must be clear whether we have to collect data for primary school teachers, lower secondary school teachers or secondary school teachers. If we want data on primary school teachers only, then we have to see the number of teachers with years of teaching experience to serve the purpose of our enquiry.

Placed in Relation to Each Other

Researchers frequently use data for making periodical or regional comparison, that is, they collect data with the purpose to compare. However, research evidence reveals that comparable figures tend to lose a large part of their significance. Thus, our data collection methods should ensure homogeneity for comparison and not heterogeneity. If the figures are heterogeneous, they cannot be placed in relation to each other. For example, if data on enrolment at primary level of education for India and Pakistan for the year 2016 are available, the researcher can compare these two figures of India and Pakistan for the same year.

Applications of Statistics

We have seen earlier that data are indispensable in almost all walks of human life. Researchers use extensively observations, facts and figures, interviews and experiments to collect information (qualitative and quantitative) with the help of diverse statistical methods in all investigations. Statistics highlights almost all aspects in an enquiry. Its main aim is to simplify the complexity of information collected in an enquiry. Researchers present data in a simplified form so as to make them intelligible. Statisticians analyse data and facilitate the interpretation of conclusions.

Simplification of Complex Facts

One of the most important applications of statistics is that it simplifies huge collection of data. Statistics present facts and figures in a definite and explicit form that builds the statement logical and convincing than mere description. Statistical methods enable the researcher to understand the whole in the short span of time and in a better way. It condenses the whole mass of figures into a single figure which in turn makes the problem intelligible.

Reduces the Complexity of Data

Statistics simplifies the complexity of data. With raw data, we cannot draw any meaningful conclusion. Statistical methods help researchers in interpretation and drawing inferences. Therefore, statistics enables to enlarge the horizon of our knowledge.

Comparison

To determine the efficiency of any measure, comparison is necessary. Statistics helps comparison of data. For example, comparison between past and present trends helps us draw conclusions. They enable us draw and ascertain the reason(s) for changes which have taken place and the possible effect of such changes in future. Several statistical tools and devices such as averages, ratios, coefficients and so on are used for making comparison.

Relationship Between Facts

Statistical methods are used to investigate the cause–effect relationship between two or more variables. For example, the relationship between midday meal programme and attendance rate in school, teachers' absenteeism and discipline, textbooks and student learning achievement and so on can best understood with the help of statistical methods.

Formulation and Testing of Hypothesis

Formulating and testing of hypothesis is one of the key functions of statistics. This helps in developing new theories. So statistics examines the truth and helps in innovating new ideas and theories. For example, by using appropriate statistical tools we can test the hypothesis whether a particular coin is fair or not, whether students score high marks with the new teaching method and so on.

Formulation of Policies

Statistical analysis of data forms the beginning of policy formulation. It helps planners, economists, educational administrators and managers prepare all sorts of plans and programmes, and assist them in monitoring and evaluating programme success and outcomes.

Forecasting

Statistical methods are of great use to forecast the trend and tendencies and to predict the future values and subsequent course of action of the phenomenon. For example, it is only on the basis of statistical techniques that educational planners project and estimate future enrolments and teachers to meet the expected increase in the demand for education. They prepare enrolment projections on the basis of current trends in student flows at a given level of education. Similarly, the planners can forecast the future school-age population and so on, considering the present population trends.

Inferences

The main aim of statistical methods is to derive inferences from an enquiry. Statistics helps to draw inferences regarding population parameters on the basis of sample information.

Magnitude of a Problem

Statistics facilitates the realisation of the magnitude of a problem. Let us consider the following statement: the number of unemployed university graduates in a given country has been increasing rapidly. This statement is vague, as it does not allow us to understand the gravity of the problem. But, if one says that graduate unemployment is increasing by 5.5 per cent per annum, everyone will properly realise the gravity of the problem.

Uncertainty

Quite often in education, business, commerce and economics, it becomes necessary to take decisions in the face of uncertainty and study the chance of occurrence of certain events and their effect on the policies adopted.

Classification of Data

There are several ways in which data are classified. The key categories commonly used in research are discussed next.

Quantitative Data

Quantitative data are measures of values or counts and are expressed as numbers. This type of data comprises numeric variables, for example, 'how many'; 'how much' or 'how often' and can be analysed with various mathematical and statistical procedures.

> 'Quantitative data refer to data that can be quantified and verified, and is amenable to statistical manipulation' (*Business Dictionary*).
>
> 'Quantitative data is data expressing a certain quantity, amount or range' (OECD).

Basically, quantitative data 'defines', whereas qualitative data 'describes' a phenomenon. Let us consider the following case. What is the difference between having seven mangoes and saying that they are delicious? Well, the first part of the question counts or measures the seven mangoes. But for the second part of the question we cannot assign a number to show how delicious they are. Those mangoes might be delicious to someone and completely sour to another person. This example tells us precisely what quantitative data are and what are not.

There are two types of quantitative data: (a) data that can be 'counted' and (b) data that can be 'measured'. Data which can be counted are 'discrete data'. We can easily express them in numerical numbers (7 mangoes). We can also physically count them (1, 2, 3, 4, 5, 6 and 7, and so on).

Quantitative data can also be measured. In math terminology, these are called 'continuous data'. The weight of seven mangoes is continuous data because you can measure it. Similarly, the heights of Grade I primary school girls and length of the school playground are other examples of continuous data because you can express them in numerical figures.

Quantitative data are classified in four categories on the basis of 'precision' criteria. As each type of data set uses a different range of statistical tests, it is important to understand and identify the type of quantitative data available. Quantitative data based on the precision criteria can be (a) nominal, (b) ordinal, (c) interval and (d) ratio.

It is important for you to learn about these four data levels of measurement. Knowing the level of measurement of your variable is important for two reasons. First, each of the levels of measurement provides a different level of detail. Nominal data provides the least amount of detail, ordinal provides the next highest amount of detail, and interval and ratio provide the most amount of detail.

The second reason why knowledge of levels of measurement is important to you is because different statistical tests are appropriate for variables with different levels of measurement. For example, chi-square tests of independence are most appropriate for nominal level data, the Mann–Whitney U test for an ordinal level dependent variable and a nominal level independent variable. Similarly, an ANOVA is most appropriate for a continuous level dependent variable and a nominal level independent variable.

Nominal Data or Variable

Nominal data are also known as 'categorical data'. These type of data classify the data into two or more groupings that have no meaningful order (Gay, Mills and Airasian 2006). Examples of nominal data are name of your university or name of a book, gender, ethnic background and job level (executive, managers, administrators, entrepreneurs, other) and political affiliation. Group members are assigned a label in that group without any hierarchical ordering. Purely nominal data are one that simply allows you to assign categories but you cannot clearly order the variables. Key DS associated with nominal data are frequencies and percentages. Since nominal data sound like name, it is easy to remember.

In a nominal variable, values assigned serve only as labels, even if those values are numbers. For example, if you want to categorise male and female respondents in your study, you could use a number of 1 for male and 2 for female. Though 1 and 2 both are numbers, in this case they do not represent any meaningful order or carry any mathematical meaning. These numbers are simply used to denote labels. You cannot use nominal data to perform many statistical computations, such as mean and SD, because such statistics do not have any meaning when used with nominal variables.

However, nominal variables can be used to do cross tabulations. The chi-square test can be performed on a cross tabulation of nominal data.

Ordinal Data

Ordinal data are similar to nominal data. However, the difference between the two is that there is a meaningful and clear ordering of the data (variables). Ordinal data present quantities that have a natural ordering. In other words, ordinal data classify the instances of the phenomena being observed into rank order.

Let us consider this example. Education level of your participants in the study (with possible values of high school, undergraduate degree and graduate degree) would be an ordinal variable. The ordering of the three categories is a definitive (i.e., graduate is higher than undergraduate, and undergraduate is higher than high school), but we cannot make any other arithmetic assumptions beyond that. For instance, we cannot assume that the difference in education level between undergraduate and high school is the same as the difference between graduate and undergraduate. Even though these categories of education level can be ranked from lowest to the highest, the spacing between the values may not be the same across the levels of all types of educational experiences. The difference between category graduate and undergraduate may be probably much bigger than the difference between categories undergraduate and high school. Ordinal data sound like order. With ordinal data, you cannot state with certainty whether the intervals between each value are equal. If these categories were equally spaced, then the data would be interval data.

You can use frequencies, percentages and certain non-parametric statistics with ordinal data. However, means, SDs and parametric statistical tests are generally not appropriate to be used with ordinal data.

Interval Data

Interval data are like ordinal except we can say the intervals between each value are equally spaced or split. Unlike ordinal data, the difference between values in interval data is also meaningful. In interval data, the differences between two categories are assumed equal. For example, the difference between the score 80 and the score 90 is the same as the difference between the score of 110 and 120.

Let us consider another example. Suppose you have figures of annual income of three people measured in US$10,000, US$15,000 and US$20,000. The second person makes US$5,000 more than the first person and US$5,000 less than the third person and the size of these intervals is the same. If there were two other people who make US$85,000 and US$90,000, the size of that interval between these two people is also the same (US$5,000). However, you cannot apply mathematical operations of multiplication and division to interval data or variables. For instance, you cannot accurately say that 100 degrees is twice as hot as 50 degrees. Additionally, interval variables often do not have a meaningful 0 point. For example, a temperature of 0 degrees (on Celsius and Fahrenheit scales) does not mean a complete absence of heat.

An interval variable can be used to compute commonly used statistical measures such as the average (mean), SD and the Pearson correlation coefficient. Many other advanced statistical tests and techniques also require interval data.

Nominal, ordinal and interval data are important because the type of data the researcher collects influences his or her choice of statistical test. Simple distributions and summary data are commonly used to calculate several measures of central tendency and dispersion in the distribution. Generally, non-parametric tests are used on simple nominal and ordinal data having no particular pattern or distribution. Examples include chi-square test for comparing variables and the Spearman's rank-order correlation for testing correlation between variables.

A distinct feature of quantitative data is that it can be illustrated graphically very neatly and succinctly in many forms of charts and tables, such as pie charts and bar charts.

Ratio Data

With ratio variables, you can apply all sorts of arithmetic operations. Consider, for example, the ratio variable weight in kg. We can very precisely and accurately say that 20 pounds is twice as heavy as 10 pounds. Additionally, ratio variables have a meaningful 0 point (e.g., exactly 0 pounds means the object has no weight). In the example of the intelligence test scores mentioned above, there is not a true 0 in that a score of 0 on the test does not mean that there is a complete absence of intelligence. However, a 0 score in ratio data signifies a complete absence of intelligence among people observed in the research study. Other examples of ratio variables include the public education expenditure and revenue, and so on.

In ratio data, you can compare numbers as multiples of one another. For instance, one person can have as much weight as another person. Ratio data can be multiplied and divided because not only is the difference between score 1 and score 2 the same as between score 3 and score 4 but also that score 4 is twice as much as score 2.

A ratio variable can be used as a dependent variable for most parametric statistical tests such as *t*-tests, *F*-tests, correlation and regression.

Interval and ratio data measure quantities and hence are 'quantitative'. Because they can be measured on a scale, they are also called 'scale data'.

In short, quantitative data:

- Are more reliable and objective
- Help the researcher to generalise a finding
- Establish cause and effect relationship in highly controlled circumstances
- Help the researcher to test theories or hypotheses

- Assume that the sample is representative of the population
- Control subjectivity of the researcher in methodology
- Less detailed than qualitative data

Qualitative Data

Qualitative data are measures of 'types'. They may be represented by a name, symbol or a number code. Qualitative data are extremely varied in nature. It includes virtually any information that can be captured that is not numerical in nature. Qualitative data are unstructured data which ask questions of 'why not' and 'how' things happen. It approximates or characterises but does not measure the attributes, characteristics, properties and so on of a thing or phenomenon.

Qualitative data can be arranged into categories that are not numerical. These categories can be physical traits, gender, colours or anything that does not have a number associated to it. Qualitative data are sometimes referred to as categorical data.

According to Gay, Mills and Airasian (2006) and Sekaran (2003):

> Qualitative data are narrative or illustrative in format and, consequently, inherently subjective in nature. These data describe the phenomenon of interest and do not measure them. For qualitative data the meaning of context from which the data is collected is significantly important. Included in that context are the researcher as well as his or her background, perspectives, capabilities, and personal biases.

Qualitative data do not permit the researchers to take any kind of standard statistical test. The obvious reason is that data are not numeric. The basic process followed by them includes organisation and categorisation, identification and synthesising specific patterns, creating a narrative and finally describing the phenomenon of interest. So when is qualitative data used? Qualitative data are often used for case studies and for questionnaires which have open-ended questions.

Importance of Quantitative and Qualitative Data

Since quantitative and qualitative data use different analytical methods and provide different outcomes, researchers often use both methods hand in hand to get a full picture of a population. For example, if the researcher has collected data on annual income (quantitative), he could also collect occupation data (qualitative) at the same time to get more details on the average annual income by type of occupation.

Quantitative and qualitative data can be collected from the same data unit depending on whether the variable of interest is numerical or categorical. This is shown in Table 7.1 as an example.

It is important to note from the earlier table that data collected about a numeric variable will always be quantitative and data collected about a categorical variable will always be qualitative. Therefore, you can identify the type of data, prior to collection, based on whether the variable is numeric or categorical.

Qualitative data are more in-depth and detailed data. In principle, it does not attempt to find out how 'many' people do something, or think something, it tries to find out 'why' they do it—what motivates them, what their feelings are towards things and so on.

Quantitative data, on the other hand, works more by asking simple questions to a lot of people—they can find out for example, 'how' many people do a certain thing and so on, but does not go any deeper

TABLE 7.1 Examples of Quantitative and Qualitative Data

Data Unit	*Numeric Variable = Quantitative Data*		*Categorical Variable = Qualitative Data*	
A Person	• How many children do you have? • How much do you earn? • How many hours do you work?	• 3 Children • US$50,000 p.a. • 36 hours per week	• In which country were your children born? • What is your occupation? • Do you work full-time or part-time?	• France • Teaching • Full-time
A House	• How many square metres is the house?	• 300 square metres	• In which city or town is the house located?	• Lyon
A Business	• How many workers are currently employed?	• 270 employees	• What is the nature of the business?	• Retail
A Farm	• How many milk cows are located on the farm?	• 30 cows	• What is the main activity of the farm?	• Dairy

Source: Authors.

than this to find reasons, anomalies and so on. Qualitative data can help the researcher produce 'new ideas' and hypotheses to test that the researcher had not thought of, whereas quantitative data can test 'existing theories'. But answers to your research problem require an intensive use of both types of data meaningfully and intelligently.

In conclusion, both quantitative and qualitative data are needed to answer the many questions there are in education, so qualitative data can be considered highly important and should be more respected within the domain of educational research.

Uses of Quantitative and Qualitative Data

There are different ways you can use quantitative and qualitative data. But you should keep in mind that both methods can affect the statistics that you wish to produce.

Frequency Counts

Frequency counts—the number of times an observation occurs for a data item (variable)—can be shown for both quantitative and qualitative data.

Let us look at Figures 7.1(A) and 7.1(B) in which both quantitative and qualitative data are respectively arranged to show the frequency distribution of the data. These figures show how absolute and relative frequencies can be calculated on quantitative and qualitative data, such as percentages, proportions, rates and ratios. For example, figure 7.1(A) shows four people (20%) worked less than 30 hours per week, while figure 7.1(B) shows that six people (30%) are teachers.

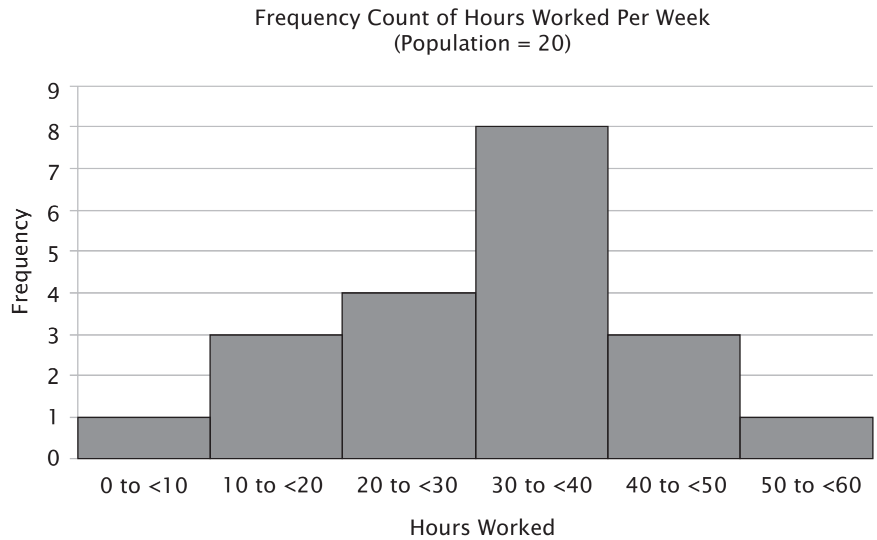


FIGURE 7.1(A) Quantitative Data

Source: Authors.

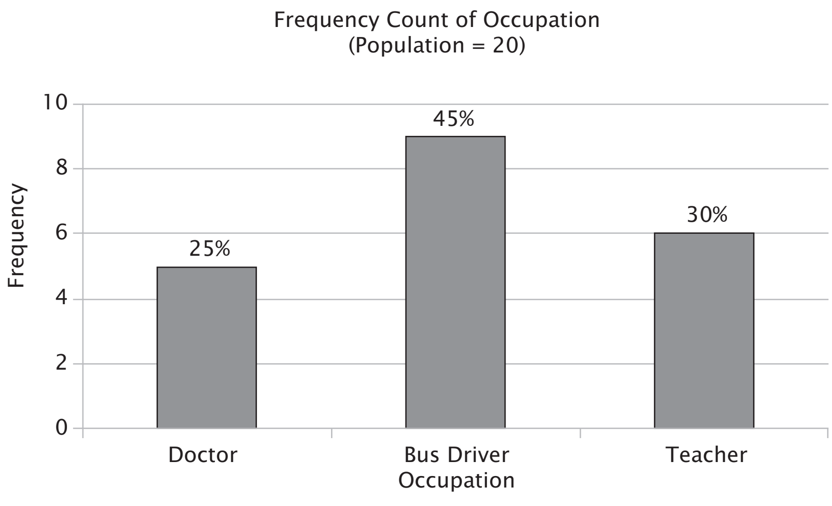


FIGURE 7.1(B) Qualitative Data

Source: Authors.

Descriptive (Summary) Statistics

Descriptive or summary statistics can be produced for quantitative data and to a lesser extent for qualitative data.

Since quantitative data are numeric, researchers can manipulate them and put them in order, add them and can easily count the frequency of an observation. Therefore, you can easily calculate all DS (mean, median, mode, weighted average and so on) by using quantitative data.

With descriptive data, which are primarily quantitative in nature, the researcher can do all kinds of mathematical or statistical computations. The data allow calculation of all sorts of measures of central tendency, measures of dispersion and range, measures of shape, correlation and multiple regression and other advanced statistics.

Inferential Statistics

Inferential statistics help you draw inferences from the quantitative data taken from the sample. Based on the sample statistic, inferential statistics are extensively used to generalise the research findings and observations for the total population.

Quantitative data are highly useful in making informed decisions and acquisition of broader understanding of a population from which the sample has been taken and analysed. The data are commonly used in research to study and consider patterns and trends as to how that population may change or progress into the future.

For example, if you want to compute a simple income projection for an employee for the year 2015, this may be inferred from the rate of change for data collected in 2000, 2005 and 2010.

Consider Figure 7.2. It shows annual income of an employee for the period 2000–2010. It also shows a 5 per cent increase in his income every five years. If we want to calculate his income in the year 2015 assuming 5 per cent increase for every five year, then his projected income would be US$46,305, which is the 2010 wage of US$44,100 increased by an additional 5 per cent.

Qualitative data are not compatible with inferential statistics as all techniques are based on numeric values.

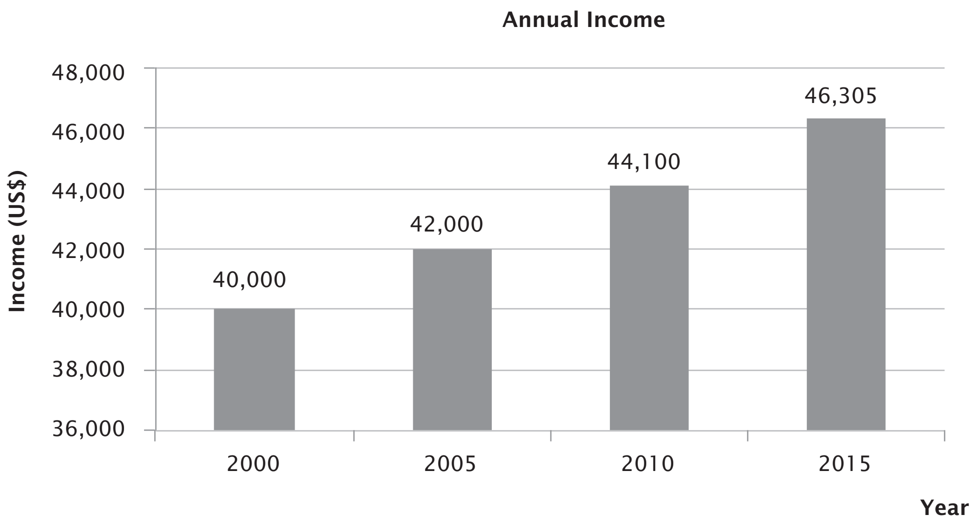


FIGURE 7.2 Annual Income of an Employee (2000–2005)

Source: Authors.

Data Sources

Data, of course, be categorised in a number of ways. Figure 7.3 shows data categorisation on the basis of sources of data. Researchers classify data sources into three categories:

- Primary
- Secondary
- Tertiary

The prime objective of any investigation is to explore a true, authentic, reliable and unbiased answer to the research question. Researchers use several tools to realise this objective. They collect and analyse data and in the end to find out possible solution(s) to the research problem. To do this, researchers have at their disposal a variety of methods to collect and analyse data. Figure 7.3 lists the most prominent sources of data.

Primary Data Sources

Primary data are the data that are gathered never before. These are actually unvarnished data that have never been manipulated before. For this purpose, the researcher can use the questionnaires specifying the

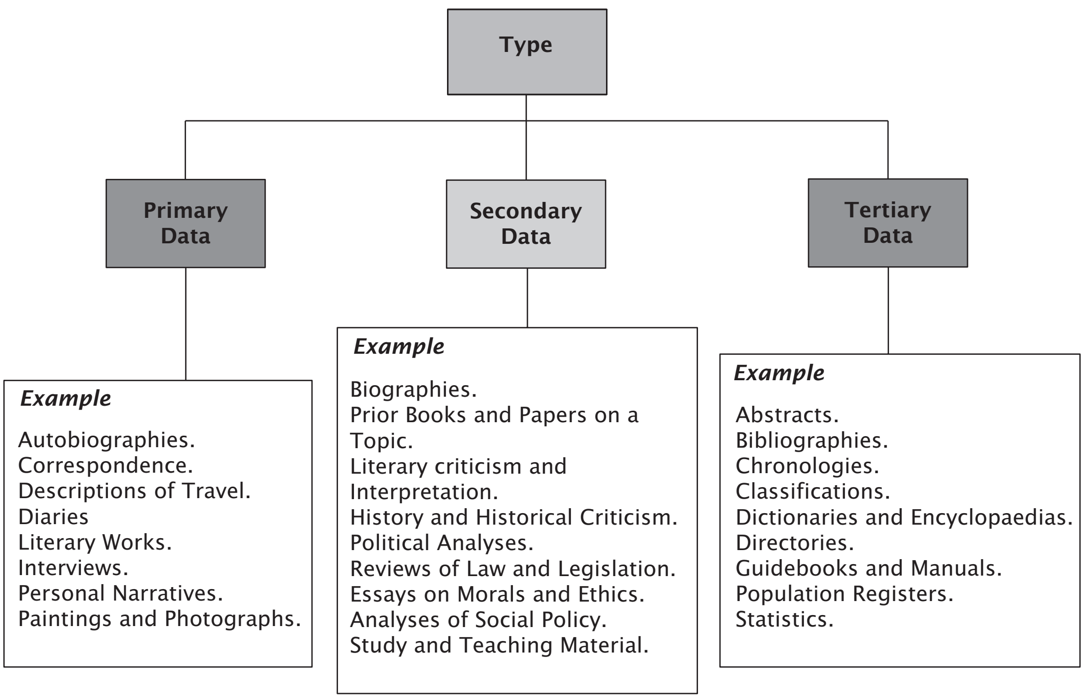


FIGURE 7.3 Classification of Data and Data Sources

Source: Authors.

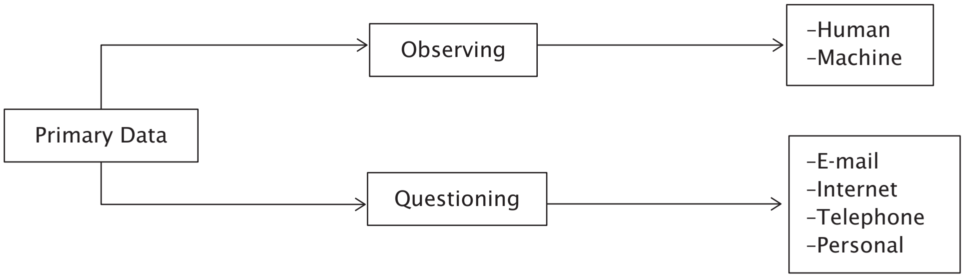


FIGURE 7.4 Primary Data Collection Methods

Source: Authors.

special factors that he needs to collect. Therefore, before collecting the primary data, it is crucial that you investigate if there is any other source available with the information which you are looking for. Primary data are collected with 'the objective of identifying some specific factors needed by the researcher'.

Primary data allow researchers to get as close as possible to original ideas, events and empirical research as possible. Such data sources may include creative works, first-hand or contemporary accounts of events and the publication of the results of empirical observations or research.

Researchers in the field of education can obtain their data by getting it directly from the subjects they are interested in. Primary data are those that are original and researchers collect them for the first time.

Figure 7.4 shows a broad classification of primary data collection method.

Each method shown in Figure 7.4 has its relative merits and demerits. Each method is used according to the type of study. Some of the most prominently used methods of primary data collection are observation, interview, questionnaire and experiments. There are many other methods but they are less recognised compared to the methods presented above.

If you are interested in obtaining the primary data, the most popular method is the questionnaires (Gardner and Haeffele 2012). The reason for this is that you as researcher can build up the questionnaires according to your requirements.

Depending on the research method, primary data may be both quantitative and qualitative. It may include subjects in large quantities or one smaller in number. One of the problems using primary data is that the researcher has nothing to back it up against if the data are not normally distributed (skewed) or incorrect or if there is some degree of experimental error.

While using primary data, you have full control over when and how to collect the information, but at the same time you are responsible and accountable for the data you have collected and may be questioned about its source(s), authenticity and reliability and at some stage on the process.

Primary data are highly useful for undertaking small experiment or running a small survey. You should always use them when the aim of your research is to make claims or criticisms, as evidence for theories, or to gain timely perspectives on a topic.

If the researcher is using this method to obtain data for the interested party, he or she should always consider how much would be the total cost to complete his or her study. Considering the cost and time involved in collecting primary data, the researcher should first check if any secondary data that suit the purpose of his or her enquiry or are flexible to use, after doing some modification, are available. If such data sources are not available, then only he or she should proceed with the methods of collecting primary data.

Advantages and Limitations of Primary Data

It is the information fathered first-hand by the researcher, who collects the data and completes the study process without seeking recourse or reference to any second-hand sources. Large organisations (generally government and private sector organisations) with good funding prospects can perform primary research such as surveys, face-to-face interviews, social media surveys/polls/feedback, analysis of customer feedback, getting response via e-mail and so on.

Advantages

- As the data has been collected personally, primary data relate directly to the researcher's study.
- Research is oriented for specific goals and purpose, cutting out possibility of wasting resources.
- Researchers can change the course of study whenever needed and choose platforms for observation well suited for projects.
- Original research quality is given, and the data do not carry bias or opinions of the third parties.
- Researcher is in full control of how and from where he or she collects the information.

Limitations

- Primary data are expensive to collect and take a long time to process.
- As the researcher is fully in charge, he or she has nothing to back up if he or she makes a mistake. In case of an error, it would be difficult to tell as there is nothing for comparison.
- Primary data collection procedure is more time-consuming.
- The outcome from research audience may not be always feasible.

Secondary Data Sources

In conducting research, educational researchers also draw upon data from a variety of sources on different subjects: economy, finance, demography, health, education, culture, environment, agriculture and so on. These data are gathered and made available by governments, international organisations, national and international NGOs, social science scholars and students from various disciplines.

Many research studies in the field of education do not require original data for analysis. This is simply because there are so many agencies and researchers gathering, publishing or otherwise distributing data all the time. The researcher should explore, analyse and illuminate these secondary data in new ways for different purposes.

Secondary data are the data that have been already collected by and readily available from other sources. A secondary source of data contains commentary on or discussion about a primary source. A key feature of secondary sources is that they offer an interpretation of information gathered from primary sources.

Use of secondary data is highly meaningful when the researcher reviews the literature to see and understand the views of others on his or her topic of research. Secondary data sources can be a good place to gather background information on a topic. Secondary data can also be used to identify what sub-topics have already been explored on a given topic.

Advantages and Limitations of Secondary Data

Advantages

- Secondary data help other future researchers in referring the data for studies.
- Trustable and ethical practices existing to support or organise other researches.
- Cost-effect, ready-made observations, less time spent on gathering information. Secondary data sources are cheaper and more quickly obtainable than the primary data sources and also may be available when primary data cannot be obtained at all.
- Statistically reliable, less requirement of expertise from internal team.

Limitations

Secondary data, however, are not free from disadvantages. Secondary data hardly fit in the framework of the marketing research. Reasons for its non-fitting are as follows:

- Information may be unsuitable for current research project
- The data may lack details that fulfil goal of the client at present
- Not customised, may require intensive study to judge validity of data
- A serious disadvantage of secondary data is the unit of secondary data collection. Let us consider a case where you want information on disposable income, but the data are available on gross income. The information may not be in accordance to your needs
- The researcher is not sure of the accuracy of secondary data
- Data may be outdated
- Class intervals (CIs) or boundaries may be different when units are same. For example:

Class boundary in 2010	Class boundary in 2015
US$2,500–5,000	US$5,000–6,000
US$5,001–7,500	US$6,001–7,000
US$7,500–10,000	US$7,001–10,000

In this case, the data collected in 2010 are of no use to the researcher in 2015.

Because of the above-mentioned disadvantages of secondary data, you must always evaluate the secondary data prior to its use in your research. Evaluation means fulfilment of the following four requirements:

- Availability: It has to be seen that the kind of data the researcher is looking for is available or not. If it is not available then he has to use the primary data.
- Relevance: It should be meeting the requirements of your research problem. There are two criteria that need to be taken into account:
    - Units of measurement should be the same.
    - Concepts and terms used must be same and currency of data should not be outdated.
- Accuracy: For finding out the accuracy of data, the following points must be considered:
    - Specification and methodology used.
    - Margin of error should be examined.
    - The dependability of the source(s) must be seen.
- Sufficiency: The availability of adequate data must be ensured for the research.

Tertiary Data Sources

By synthesising information gathered from other sources, tertiary data sources provide overviews of topics. Tertiary data sources often provide data in a convenient form or provide information with context by which to interpret it.

Tertiary data refer to data that are generated about us, rather than by us. In fact, tertiary data sources are distillations and collection of primary and secondary data sources. As the information presented through tertiary data has been filtered through many reviews, it tends to consist of highly reliable and accurate information.

No particular author is a claimant of tertiary data sources. Researchers use these sources to have general overview of the research topic, its basic terminology and background information and for references. But they rarely contain original material.

When you are beginning research on a wholly unknown topic, your first step would be to consult general references. The main sources of tertiary information are shown in Table 7.1. These tertiary sources provide general explanations condensed from common knowledge on the topic intended for a broad public audience.

In short, the primary data sources collect information afresh, and for the first time, and thus happen to be original in character. Data are collected first-hand by the researcher himself or herself without relying on any kind of pre-researched information for the specific purpose or analysis under consideration. Here, a research team conceives of and develops a research project, collects data designed to address specific questions and performs its own analyses of the data the team collected. The people involved in the 'data analysis' should therefore be familiar with the research design and data collection process.

Secondary data are collected from other means instead of the researcher's own means. The analysis of qualitative data can be a long and extended process requiring much imagination and reflection on the part of the researcher.

Significance of Statistical Unit in Research

Selection of unit of analysis is one of the most important aspects of a research study. 'The unit of analysis is the major entity that the researcher analyses in his or her study'. Units of analysis are 'what' or 'who' being studied. Choosing the unit of analysis is done at the conceptualisation stage after the formulation of the research objectives. The researcher should always choose the units of analysis prior to choosing his or her research instrument.

In social science research, the most typical units of analysis are individual people, groups, organisations and social artefacts. Other types of units of analysis include institutions, cultures, societies and so on. There are some studies that make descriptions or explanations pertaining to more than one unit of analysis. In studies using multiple units of analysis, it is important to anticipate what conclusions the researcher wishes to draw with regard to units of analysis.

Why is it not called the unit of sampling? This is because it is the analysis a researcher does in his or her study that determines what the unit is. For instance, if you are comparing children in two classrooms on achievement test scores, the unit is the individual child because you have a score for each child. But if you are comparing the two classes on classroom climate, your unit of analysis is the group.

Thus, any of the following could be a unit of analysis in a research study:

- Individuals
- Groups
- Organisations
- Social artefacts

Individuals as Unit of Analysis

In social sciences, individuals are the most typical units of analysis. The reason is the norm followed in social science research. The norm suggests that scientific findings are most valuable when applied to all kinds of people. If the units of analysis are chosen from a very small group of people, its results would be generalisable only for much narrower population of the people within a nation, region, city and so on.

Some comparative studies are specifically designed to examine phenomenon across national boundaries. Thus, in such studies, the groups whose members may be units of analysis are restricted (aka circumscribed). Restricted groups whose members may be units of analysis at the individual level include: students, residents, workers, voters, parents, faculty and so on. Each of these units of analysis implies some restricted population, that is, a sub-population.

In descriptive studies, the researcher makes observations describing the characteristics of a large number of individual people, each as their sex, ages, regions of birth, attitudes and so on. In most studies, the results of examining individual units of analysis are aggregated to describe the general sub-population restricted by the units of analysis. Descriptive studies having individuals as their units of analysis aim to describe the population that comprises those individuals.

Explanatory studies, on the other hand, aim to discover the social dynamics operating within that population. Here, individual may be examined as the factors that cause them to act in a particular manner. The results of these studies are then aggregated to explain the behaviours of the general population.

Individuals, as the units of analysis, may be characterised in terms of their membership in social groupings. A research project might examine whether high school graduates in rich families are more likely to attend college than those in poor families. Each individual unit of analysis implies a circumscribed population that must be delineated at the time of operationalisation of the research design.

Groups as Units of Analysis

In research dealing with social problems, social groups themselves may be the units of analysis. It is important to remember that the groups as the units of analysis for social research are not the same as studying the individuals within a group. Consider the following example. A researcher wants to learn about a criminal gang. In this case, individual gangster would be his or her unit of analysis. But if he or she wants to learn the differences between big and small gangs, or uptown and downtown gangs, then a group would be his or her unit of analysis. Other units of analysis at the group level include: the family, cliques (circles), couples, census blocks, cities, regions and so on.

When your units of analysis are social groups, their characteristics may be derived from the characteristics of their individual members. A family might be described in terms of age, race or education of its head. Groups and individuals may also be characterised in other ways, for instance, according to their environments or their membership in larger groupings.

Organisations as Unit of Analysis

Formal organisations, such as corporations, implying the population of all corporations, may be the units of analysis in social research. As units of analysis, it is easier, in many ways, to operationalise organisations than informal social groups because the organisation itself or the law often provides clear boundaries.

At the level of social groups and organisations, the researcher may examine characteristics of individuals or characteristics of the groups and organisations themselves and still use the population of the group or organisation as the unit of analysis. Experts suggest that organisations are unit of analysis in case the aim of the researcher is to examine characteristics of individuals in a sample of organisation and to compare organisations by aggregating the information of individuals. If the researcher examines characteristics of the organisations themselves, such as the income of the organisation, then the organisations are the units of analysis and individuals are not even looked at.

Social Artefacts as Units of Analysis

Social artefacts are the products of social beings or their behaviour such as music, homes, cars, ideas, poems, books, paintings, pottery, jokes, scientific discoveries and so on. Each social artefact comprises a population of all such objects.

Another important class of artefacts include social interactions such as weddings, divorces, parties, church services and so on.

For social interactions, the researcher should be very precise whether he is studying a social artefact or the individuals who create or use the social artefact, that is, the purpose of the research is to examine brides or weddings. Other examples of social artefacts include friendships, court cases, traffic accidents, fistfights, ship launchings, airline hijackings, race riots, congressional hearings and so on.

In the light of the earlier discussion, the selection of statistical units for data collection and analysis is the key of any research study. If the units of analysis are not clearly defined and understood, the data collection exercise would not be systematic and the findings of the study not valid and meaningful. When selecting the units of analysis, it is important for the researcher to ensure that the statistical units meet the following requirements:

- Statistical units are simple, specific and unambiguous.
- They are stable.
- They are comparable.
- They are appropriate to the objective of enquiry.
- They are precise and ascertainable.

Table 7.2 shows some examples of units of analysis and variables. Table 7.2 will enable you understand which variables you might consider within each unit of analysis for your research study.

In short, unit is the basis of all these measures in which we, collect, analyses, edit and interpret the data used for the survey. The selection of a given unit of analysis will be determined by the nature and scope of our research enquiry.

TABLE 7.2 Examples of Units of Analysis and Variables

Unit of Analysis	*Variables*
Objects	Characteristics of objects which vary
Individuals	Income, age, sex, attitude towards abortion, how voted in 2016
Household	Income (not the same as individual income), size (number of people), marital status of head
Organisation	Number of organisational levels, size and sex composition
Census tracts	Level of median income, per cent of people belonging to specific religion (Muslims), per cent of owner-occupied dwellings
Metropolitan areas	Population, number of jobs in manufacturing, region (of the country in which is), per cent graduates in total
Nations	Population, GNP, per cent of population with access to drinking water, system of government

Source: Authors.

Limitations of Statistics

Although statistics is a very useful science, it suffers from certain limitations. According to Newsholme, 'It must be regarded as an instrument of research of great value but having several limitations which are not possible to overcome and as such they need a careful attention' (cited in Billings 2012).

Qualitative Aspect Ignored

As the qualitative phenomenon is not expressed in quantitative terms, research studies comprising such information cannot be studied and analysed by statistical methods. Such phenomena (qualitative aspects) cannot be a part of the study of statistics. For example, qualities such as beauty, honesty, intelligence, health, riches and so on cannot be numerically expressed. So these qualitative characteristics cannot be examined statistically.

Statistics Does Not Deal with Individual Items

Statistics deals only with the aggregates rather than individual item or isolated figures. An individual item like height of a student in a class is 155 centimetres cannot be considered statistics. Researchers use statistical methods which deal with aggregates and not with a single figure. For example, the average height of a class 160 centimetres, this single figure refers to the aggregate of all individuals in the class. Statistics cannot be of much help for making a study of the changes that may have taken place in individual cases.

Statistics Laws are Not Exact

The 'law of inertia of large numbers' and the 'law of statistical regularity' are not as good or perfect as the laws of science. Both are based on probability. For example, on the basis of probability or interpolation, we can only estimate the number of output of trained teachers from teacher training institutes in the country but cannot make a claim that this number be exactly 100 per cent. Here only approximations are made.

Results are True Only on Average

The researcher can interpolate the results for which time series or regression or probability can be used. These are not absolutely true. For instance, if average of two sections of students in mathematics is same, it does not mean that all the 50 students in Section A of Grade I have secured same marks as in Section B of the same grade. There may be much variation between the two. So the results we get reflect average results.

Statistics Can Be Misused

Research studies use statistical techniques to analyse and interpret the information collected for an enquiry. They neither prove nor disprove anything. They are just a means to an end. As a matter of fact, statements supported by statistics are more appealing and are commonly believed. For this, statistics is often misused. Statistical methods rightly used are beneficial but if misused the analysis will be disastrous. Statistical methods used by inexperienced hands will lead to inaccurate results. Here the fault does not lie with the method of statistics but with the researcher who makes wrong use of it.

Too Many Methods to Study Problems

We can use several statistical methods to find a single result. For instance, the researchers can study variations in pupils' learning achievement by using quartile deviation, mean deviation or SDs. But the results will certainly vary in each case.

Statistical Results are Not Always Beyond Doubt

Although the researchers use a variety of laws and formulae in statistics, one cannot say with fair amount of certainty that the results achieved are final and conclusive. As they are unable to give complete solution to a problem, the result must be taken and used with much wisdom.

In a nutshell, the limitations of use of statistics in research are as follows:

- Statistics does not deal with isolated measurement.
- Statistics deals with only quantitative characteristics.
- Statistics laws are true on average. Statistics are aggregates of facts. Thus, a single observation is not statistics. It deals with groups and aggregates only.

- Statistical methods are best applicable on quantitative data.
- Statistical cannot be applied to heterogeneous data.
- The results of the study might be misleading if data are not collected, analysed and interpreted scientifically.
- Only a qualified and trained statistician can handle statistical data efficiently.
- Some errors are possible in statistical decisions. Particularly the inferential statistics involves certain errors.

Knowing the fact that methods of data collection differ significantly from one study to another and also from one discipline to another, the importance of ensuring accurate and honest data collection remains the top priority. No matter for which field of study or preference for defining data (quantitative, qualitative) you follow for your research, the fact remains that accurate data collection is vital to maintaining the integrity of your research. Both the selection of an appropriate data collection instruments and well defined, explained and delineated instructions for their correct use minimise the likelihood of errors taking place. If data are not collected properly, they may lead to the following consequences:

- Inaccurate answer of research questions
- Inability to repeat and validate the research study
- Waste of resources due to distorted findings
- Misleading other researchers to pursue meaningless avenues of investigation
- Compromising decisions for public policy

Data Needs in an Educational Setting

Educational administrators rely on data in the process of scientific decision-making and to manage their operations, whereas educational researchers use data to explore, investigate or describe pressing education problem(s) and suggest possible solutions for resolve. It follows from this that the relevance of research in an education setting depends greatly on the selection of appropriate instruments of data collection, its analytical methods and interpretation techniques.

Data are endemic in educational settings. Education systems are set up to achieve certain objectives. Educational researchers collect data on both human and material resources as inputs to be processed in order to produce some outputs and outcomes. Such outputs and outcomes are usually subjected to some forms of evaluation to check if they meet expected standards. To enable an education system to realise the planned goals, there is a great need for the planners and managers of the education system to plan strategically. Scientifically collected data help educational managers and researchers in identifying the system's goals, ascertaining resources available, prioritising the goals, identifying alternative strategies for achieving the goals, choosing the best strategies and preparing programmes and activities that will lead to the achievement of stated goals. Figure 7.5 highlights the needs and uses of data in an educational setting.

It is clear from this figure that in an educational setting data should cater for the needs of:

- Agencies responsible for policy, planning, coordination and administration of educational and training activities

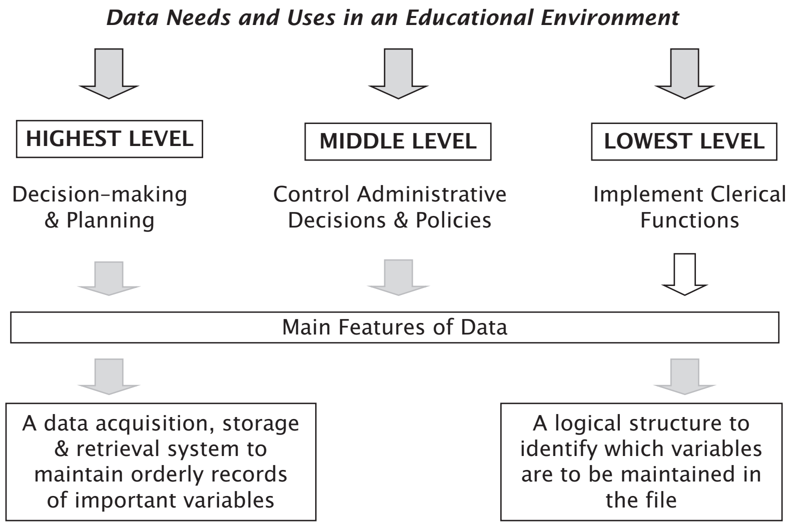


FIGURE 7.5 Features of Data in an Educational Setting

Source: Authors.

- Agencies supporting and organising education activities both within and outside the government sector
- Educators, researchers, community leaders, general public and so on
- Potential learners

Data in an educational environment are needed to increase our understanding of progress and disparities in the development of various levels and types of education programmes in terms of availability, resource inputs, participation, efficiency, quality output, learning achievement and impact.

You might have noticed that ministries of education and school districts usually gather much more data than they can effectively use. One of the key challenges facing the educational planners and researchers is: How to analyse the data and use it wisely? Many believe the logical starting point is to make better use of existing or archival data. These data include statistics on attendance, grades, referrals, retentions and standardised test results. When compiled and reported on a regular basis, researchers can use these data to provide a baseline of school operations and to make comparisons among similar schools.

What kinds of data are important for continuous school improvement, and how can you best organise the data for easy access and analysis? These are fundamental questions in educational research. More specifically, these include, among others, the following fundamental questions:

- What could be the learning impact in terms of personal growth and well-being, employment and other socio-economic cultural development?
- Does the supply of educational provisions match the demand? What are the gaps, problems and possible solutions?
- What measures and actions could be taken in order to solve the problems, coordinate the education activities, promote support to them, increase participation and improve their relevance and impact on overall development?

Briefly speaking, the education system is a production system that uses both human and material resources for meeting the qualified and trained manpower needs of the country. For achieving this goal, research should help planning, implementation and evaluation of success of education development plans. Research should also aid public accountability, curriculum review and staff recruitment and development. These functions of research cannot be effectively and properly met in the absence of reliable data. Thus, for the educational research to be meaningful accurate data and its scientific analysis is an imperative. The success and relevance of research will depend primarily on accurate data collection and its analysis since poor data collection and analysis could lead to recording of fiction instead of facts and this could distort sound decision-making.

It should, however, be borne in mind that for research studies data alone is simply not enough to make the judgement and draw conclusions. It should be balanced by researchers' understanding of the wider issues facing our education systems such as the quality of provisions including teaching and learning, the curriculum, pupil performance, care, support and guidance. Such data are vital and equally important in conducting research in the field of education.

Data Elements for Educational Research

Modern education systems do face from time to time a number of challenges in terms of policy and planning. In the absence of reliable statistics, administrators can barely provide objective information particularly when making organisational and curricular decisions. Without these hard data in place, often they fail to justify their arguments and policy decisions on scientific reasoning. Thus, for relevant and meaningful inquiry in education, researchers must take into account data collection on the following four key elements.

Student Learning

The term student learning refers to a wide variety of educational programmes, learning experiences, instructional approaches and academic-support strategies that are intended to address the distinct learning needs, interests, aspirations or cultural backgrounds of individual students and groups of students (Postlethwaite 2005).

Researchers use students' learning data for a variety of tests such as (a) norm-referenced tests, (b) criterion-referenced tests, (c) standards assessments, (d) teacher-assigned grades and (e) authentic assessments. All of these tests measure the impact of our education system on our students. Educational researchers quite often analyse data on student learning to reveal the hard facts such as what students know, what they should know and what can be done to meet their learning needs. When analysed and interpreted appropriately, researchers can help educators make informed decisions that positively affect students' learning outcomes.

According to Blaich and Wise (2011), data sources that are useful to consider in assessing student learning are as follows:

Evidence of Learning Outcomes

- Direct measures of learning: These allow students to demonstrate to faculty to assess how well students attached to a given programme are meeting expected level of proficiency for skills or knowledge. These measures include papers, standardised school board tests, observations of students in a clinical setting and quiz questions related to a key area of knowledge needed.

- Indirect measures of learning: They are meant to gather students' perception of and satisfaction with their learning. Focus groups, and student and alumni surveys are common examples of indirect measures of learning.
- Multiple methods of learning: Research evidence suggests that students quite often not able to self-assess precisely their learning outcomes. As a result, the use of indirect measures alone may be inaccurate. Mixed methods that use both direct and indirect measures are highly valuable.
- Background information about the students in the course or curriculum: Information on issues such as: what sort of academic preparation do students reveal and/or demonstrate in a course or in an academic programme? What are the key demographic characteristics of students in a class or programme? What are the aspirations of graduate students? (Postlethwaite 2005)
- Documentation of the learning outcomes: What are the learning experiences and outcomes of students?

In educationally developed countries, the Monitoring and Evaluation Division/Unit of the MOE collaborates with departments or school/colleges to collect assessment data for educational grant evaluation or curriculum decision. Interviews, focus groups and surveys of students, faculty and alumni are widely used methods to collect feedback data.

Perceptions

The term perception refers to 'a belief or opinion, often held by many people and based on how things seem' (*Oxford Dictionary*).

Educational researchers collect perception data through questionnaires, interviews and observations. These data are highly useful for understanding the diverse views and opinions of students, parents, teachers and the community members about (a) making informed decisions about students and the learning environment; (b) measuring stakeholders' perception of the learning community—because perception does shape reality; (c) measuring the school's self-perception against the community's image of the school and (d) identifying programme effectiveness. For example, student perception can help you understand factors which encourage and motivate student to learn. Similarly, staff opinions may suggest the kind of change that is possible and necessary within the school. It is the data that tell us how people feel about being part of an educational institution, that is, how they see it.

As regards Ministries of Education, they also want information on behaviours and perceptions which directly associate and correlate with student achievement. Research studies planned in this manner can certainly generate data that the school improvement team can use to identify intervening variables that need improvement to positively impact student achievement. These research studies can also help ministries in contributing to the reporting requirements outlined in national policy documents.

School Processes

School processes define what learning organisations, and those who work in them, are doing to help students learn: what they teach, how they teach and how they assess students. School processes include programmes, curriculum, instruction and assessment strategies, interventions and all other classroom practices that teachers use to help students learn.

Keeping track of school processes through careful documentation helps you build a continuum of learning for all students.

These data are necessary as it provides information on the educational environment, programmes, vision and so on. Education leaders are accountable and have unique responsibilities for developing and implementing a vision of learning to guide organisational decisions and actions on the one hand and to provide appropriate and effective learning opportunities on the other hand.

The types of questions asked when researching the school process and the corresponding data needs are different from those asked for student learning and people's perception. The researcher needs data to answer questions how well interventions are being implemented. Educational researchers need data on school processes to examine and answer the following key questions:

- What intervention activities are taking place?
- Who is conducting the intervention activities?
- Who is being reached through the intervention activities?
- How much resources have been allocated for school development programme?
- What are possible weaknesses and threats that deserve immediate attention for improvement?

In short, school processed data are diagnostic tool. Its use in research facilitates rich and deep collaborative discussions among all the education stakeholders about key problems and challenges facing their education system. The data can serve as a guide to determine a school's strengths and challenges as well as directions for improvement based on analysis of data and responses to series of data-related questions in content areas. These data should include the identification of achievement gaps as well as reflection on possible causes for these gaps.

Demographics

School demographics are the statistical information that describes a school's population and organisation. They usually include items such as the ethnicity and SES of students, teacher qualifications, class size and graduation rates.

Demographic data provide descriptive information about the school community—students, school staff, the school and the surrounding community. Examples on students' demographics may include enrolment, attendance, grade levels, race/ethnicity, gender, students with disabilities, English learners, SES, graduation rate, suspensions/expulsions and so on.

For teachers, demographic data usually provide information, among others, on number of years teaching and administrative experience of school teachers and headmasters, respectively, absenteeism, teacher training (pre-service and in-service), teachers' professional development and impact of these on students' learning achievement.

About school-related demographic data, the key data should provide information about the practices and procedures school uses to plan, deliver and monitor curriculum, instruction and assessment, strands/standards/indicators which stand out as strengths, how might these challenges impact student achievement, actions which could be incorporated into the school.

Data on school community are key in any educational research. While intending to analyse the impact of involvement of school community, UNESCO-IIEP suggests that the researcher should collect relevant information on following key dimensions:

- What is the educational attainment of the school community?
- How might it impact educational performance and opportunities of the community school?
- What is the enrolment by type of educational institution/level in the school community?
- How might larger or smaller education and enrolment impact the livelihood patterns of the school community?
- What are demographic features (age, gender, race/ethnicity) of taxpayers in the school community, who typically finance much or most of the community school?
- How do income and property characteristics impact the community school?
- What are the characteristics of household composition and family structure? How might these attributes affect growth of the school community?
- What are the special learning needs of the foreign-born population and language spoken at home?

The aforementioned data clearly show that in an educational research, school/school community data are broader than those data only about students, staff and facilities. They include all the other dimensions about the school community.

In summary, the importance of demographic data in all sorts of educational research studies cannot be underestimated. Depending on the understanding and the availability of pertinent data for a given region, researchers can study and analyse trends (e.g., trends in enrolment, teachers' demand and supply, expenditure on education and so on) make future projections and forecasts, and suggest proper strategies to use in order to provide learning spaces to prospective learners. Demographic data are also crucial in the education process because it helps the researcher to have a better and clear understanding of who makes up the country and how the lifestyles of its inhabitants have undergone changes and how they have evolved over the years.

It is important to remember that

> [O]n the surface, demographic data can appear as little more than sterile numbers and statistics. However, if the researcher delves into those numbers, he will find out that these numbers illuminate aspects of life that are not found anywhere else. By analysing demographic data the researcher can understand where and how specific population groups live, what is their background, their economic profile, their family lives, their educational attainment, and more. (UNESCO Institute for Statistics Databank 2017)

Summary

The analysis of quantitative and qualitative data can be complex and time-consuming process. This chapter presented a discussion about the basic elements of data and their role as key element of educational research. The chapter described the key elements of data and statistics to enable you undertake basic analysis of both qualitative and quantitative data that will encourage you and new researchers to research inquiry and give some reminders for those already engaged in educational research. It gave a helpful grounding to begin thinking about salient features of quantitative and qualitative data that you may develop for your research in the field of education.

The basic concepts described in this chapter become meaningful for your research inquiry only when you have some scientific tools to describe and provide accurate and precise information to illuminate the condition of education and contribute to its improvement. The information generated will be neither possible to grasp through casual observation nor generally available from other efforts to collect, report and analyse data that you have collected about schooling. For formulating specific schooling goals and for translating those goals into action, you need a complete understanding of a good education indicator system. Chapter 8 provides key educational indicators and a description of how and when they might be used in any research inquiry.

Self-test Exercises

Exercise 7.1:

7.1.1. Describe the five major characteristics of statistics.

7.1.2. Define statistics. Explain its types and importance to education.

7.1.3. 'Statistics is all-pervading'. Elucidate this statement.

7.1.4. Write a note on the scope and limitations of statistics. Explain with suitable examples.

7.1.5. Distinguish between descriptive statistics and inferential statistics.

7.1.6. Imagine that you have been asked to evaluate your teacher. Determine which aspects of teaching you would consider important and develop a set of indicators that might reflect these.

References

Billings, J.S. 2012. 'Newsholme's Vital Statistics'. *Journal of the American Statistical Association* 1 (6): 269–271.

Blaich, C.F. and K.S. Wise. 2011. *From Gathering to Using Assessment Results: Lessons from the Wabash National Study* (NILOA Occasional Paper No. 8). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Bowley, A.L. 1901. *Elements of Statistics*, 6th ed. New York, NY: Scribner; London: King.

Croxton, F.E. and D.J. Cowden. 1946. *Applied General Statistics*. New York, NY: Prentice Hall.

Gardner, D.C. and L.M. Haeffele. 2012. *When to Use What Research Design*. New York, NY: Guilford Press.

Gay, L.R., G.E. Mills and P. Airasian. 2006. *Educational Research: Competencies for Analysis and Applications*, 8th ed. Upper Saddle River, NJ: Pearson.

Postlethwaite, N. 2005. *Educational Research: Some Basic Concepts and Terminology Module 1*. Hamburg: UNESCO International Institute for Educational Planning, Institute of Comparative Education, University of Hamburg.

Sekaran, U. 2003. *Research Methods for Business*, 4th ed. Hoboken, NJ: John Wiley & Sons.

UNESCO Institute for Statistics Databank. 2017. Available at: http://www.geolytics.com/resources/demographic-data.html (accessed on 29 January 2018).

8
CHAPTER

Education Indicators

Introduction

There is a growing body of opinion throughout the international community that a crucial task of national governments, researchers and international support agencies is the collection of data to construct key educational indicators. Since an 'education indicators system' is dependent upon a multitude of complex factors, it is necessary for you to understand the uses and applications of these indicators to scientifically interpret the outcomes of your research and to explore and highlight gaps in data and resources and, in turn, to inform educational planners the need to collect the relevant data. Such indictors are powerful tools for monitoring, evaluation and also for advocacy purposes and for seeking cooperation of the international community to coordinate.

The work on education indicators should, therefore, be complemented by support to develop rich databases about the state of education. A common impression is that the priority indicators are the only statistical outputs a country needs—even for its own governance. Another problem, and perhaps one of the most difficult, is the dilemma between open accountability and national sovereignty in relation to what data are collected, the methods used to collect and analyse them and who is to have access to the results.

This chapter attempts to answer this fallacy. It defines the concept of an 'education indicator' and explains the key role of indicators as an integral part of an educational research. It then describes the method and process of instituting a body of education indicators and how an educational researcher can use them in analysing data and in interpreting the research findings by supplying useful, valid, reliable, timely and cost-effective information. Finally, the chapter highlights how educational indicators enable researchers to assess progress towards the achievement of intended outputs, outcomes, goals and objectives in an education system.

More specifically, the education indicator framework covered in this chapter identifies three aspects of education: (a) institutions, service providers and resources; (b) individual-level behaviours, engagement and outcomes; and (c) contextual factors that influence learning.

The chapter provides guidelines to your research focusing in each of these three above-mentioned categories to think about what information will be most needed to measure the state of education in the country as it transforms and adapts to fast-changing technologies and global economic forces. It is anticipated that these guidelines would enable you address the following questions and to suggest answers to each of the following questions:

- What are the key indicators about (a) institutions, service providers and resources; (b) individual-level behaviours, engagement and outcomes; and (c) contextual issues?
- What is the proof or evidence that justifies the use of these indicators? What is the argument for using them?
- In what direction would you and decision-makers want the indicator to change over time? That is, whether you would want to see the indicator increase, decrease or stay the same and how such changes would be interpreted?
- What are the likely consequences associated with these indicators?
- What are the equity concerns to be considered for these indicators?
- What are the relevant data sources for developing these indicators?

Important Terms and Concepts

Before we define and describe the salient features of several types of education indicators, it is important that you understand the following most commonly used terms and concepts in the development of indicators.

- Activities: activities refer to the actions taken by the national project staff and their partners (donor partners if the project is externally funded) that are designed to meet a project's objectives. Examples of activities include hiring school staff (teaching and non-teaching staff), purchasing equipment, constructing school buildings or other facilities and providing other forms of technical assistance.
- Data source: data source refers to the information that may be used as the basis for measurement. It includes public surveys, administrative records, interviews, focus groups and observations. It also includes information collected specifically for the purpose of measurement and pre-existing sources, such as budgets, reports and legislative documents.
- Impacts: impacts refer to key strategic goals. Examples of impacts are increased access to education or improvements in teachers' professionalisation.
- Inputs: inputs are all those things that we wish to embed in a system for some type of use. Inputs could be the people, raw materials, energy, information, technical expertise, relationship with personnel or finance that provides a basis for security for the programmes. Inputs denote all sorts of aspects provided to the system to achieve a certain outcome.
  A simple example of inputs is a computer. A variety of sources is used to input: keyboard, scanner, microphone, mouse or even another computer. What we input has a purpose—but

until it is processed and generated in some form of output, it does not do us much good. In an educational setting, examples of inputs include teachers, students, staff, books, chalk, blackboard and so on.

- Log frame (logical framework): a project planning and oversight tool consisting of indicators and milestones for key inputs, activities, outputs, outcomes and impacts.
- Objectivity: it is an important feature of research. It is based on facts and includes steps designed to minimise the impact of the biases or personal preferences of the researcher, funder or others involved in the research on the findings.
- Outcomes: the benefits that a project or intervention is designed to deliver.
- Outputs: the tangible and intangible products that result from project activities. Output or processed information, in case of our computer example, comes in many different forms: monitor or printer for visual work, a speaker for audio.
- Process: a process transforms inputs and resources into outputs and outcomes. For our above example of computer, it takes place in the internal parts of the computer. It refers to the 'act of taking inputted data and converting it to something usable'.

Meaning of an Indicator

An indicator is simply a way of saying 'how much' or 'how many' or 'to what extent' or 'what size'. Indicators are tools to measure a phenomenon. Measuring is not new to none of us. We all measure all the time. We start using measures at an early age—who got the biggest pie of cake, who can run the fastest, who obtained the highest marks in a test—the examples are endless.

Measuring is part of our daily lives and there is nothing wrong with measuring and comparing. How many of us have set ourselves a goal for something that we really wanted to achieve? As a child, we all saved money to buy toys; later on perhaps we wanted a mobile phone. The amount of money you have in your bank account is an 'indicator'. However, the cost of toys or mobile phone is the 'goal'.

How many of you have ever counted the number of course credits you needed to graduate? How many of you watch the highway signs (milestones) showing the number of miles left to cover wherever you are going on a trip? We all set goals and use indicators to measure our progress towards those goals.

Different authors and organisations have defined the word 'indicator'. Despite the different expressions used in these definitions, their central concepts are more or less similar. For instance, UNDP in its *Handbook for Monitoring and Evaluation* defines indicators 'as signals that reveal progress (or lack thereof) towards objectives; means of measuring what actually happens against what has been planned in terms of quantity, quality and timeliness'. On the other hand, the Organisation for Economic Co-operation and Development (OECD) in its publication *Education at a Glance* defines an indicator as 'a quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect the changes connected to an intervention, or to help assess the performance of a development actor'.

From the above two well-known definitions, we can safely say that an indicator is a variable that:

- Reflects the efficient and effective performance programme
- Provides a standard against which to measure changes brought by programme activities
- Helps identify possible changes that need to be made in the programme structure, strategies or implementation scheme

Similarly, an 'education indicator measures the status of, or change in, an educational system with regard to its goals'. It provides evidence that a certain condition exists or certain results of the education system have or have not been achieved. Education indicators enable researchers and policy- and decision-makers to assess progress towards the achievement of intended outputs, outcomes, goals and objectives. As such, indicators are an integral part of a results-based accountability of an education system.

UNESCO Institute for Statistics (UNESCO UIS, 2009) considers

> [E]ducation indicators are statistics that enable management to monitor effectiveness and efficiency in the delivery of education services. The indicator provides a clear insight on judgements on key aspects of the functioning of the education system. Many education indicators reflect something about the quality of schooling. An education indicator must consist of at least one other reference point—an agreed standard, past measurement or comparison across schools against which it can be compared.

Thus, an education indicator is a tool that assesses the state-of-the-art of an education system. It gives an indication to report on that state to the whole of the education community, in other words, to the whole of nation. An indicator is a way to measure, indicate, point out or point to with more or less exactness. It is something that is a sign, symptom or index. It is something used to show visually the condition of an education system.

Education Indicator System

An education indicator system is an important central concept in the discussion of education indicators. Before you begin the data preparation and validation stages in your research, make sure that you have a clear sense of the purpose of the education indicator system—how you will use it in your research. Further, you should also clearly define student outcomes that the indicators are intended to predict.

Whether indicators are single or composite statistics, a single indicator can rarely provide you useful information about such a complex phenomenon as schooling. Your choice for selecting indicators for your research should be such that it generates more and more accurate and relevant information about conditions. Thus, you should remember that 'an indicator system' is more than just a collection of indicator statistics. Ideally, a system of indicators must measure distinct components of the system and also provide information about how the individual components work together to produce the overall impact or effect, that is, the whole of the information provided by a system of indicators is greater than the sum of its parts.

At the national level, indicators must represent, at least roughly, the important components of an educational system. Further to monitoring outcomes, indicators should highlight the characteristics of students and communities served by schools, the financial and human resources (especially teachers) available to the schools and other educational inputs. Moreover, they should reflect the relevance and adequacy of the curriculum and instruction received by students, the nature of the school as an organisation in pursuit of educational excellence and equity, and other educational processes. Finally, indicators must be associated to one another so that their relationships, and changes in these relationships, can be ascertained to suggest possible explanations for observed changes in outcomes.

We expect a good education indicator system to provide accurate and precise information to illuminate the condition of education and contribute to its improvement. The information generated will be neither possible to grasp through casual observation nor generally available from other efforts to collect, report and analyse data about schooling. We anticipate that indicators to assist policymakers in formulating schooling goals and translating those goals into actions.

Characteristics of an Effective Indicator

The purpose of an education indicator is to show you how well the education system is working. If there is a problem in the system, an indicator can help you determine what direction to take to address the issue. Education indicators are as varied as the types of systems they monitor. However, there are certain characteristics that effective education indicators should have in common. Good and effective indicators are:

- Relevant: They highlight and exhibit something about the system that you need to know.
- Easy to understand: Even people who are not experts can easily understand and interpret an indicator.
- Reliable: You can trust the information that is provided by the indicator.
- Based on accessible data: You can easily compute them using the available information or can gather them while there is still time to act.

Let us take a simple example. A gas gauge in your car is an indicator. The gauge shows you the amount of gasoline (petrol or diesel) left in your car tank at a given point in time. If the gauge shows the tank is almost empty, then you know it is time to fill up the gasoline tank. Another example of an indicator is a mid-term report card of a student. It shows you whether the student is doing well enough to be promoted to the next grade or if extra help is needed. Both of these indicators provide you relevant, easy-to-understand and reliable information to help you prevent or solve problems, hopefully before they become too severe.

For measuring conditions that are so complex and for which there is no direct measurement, you can use indicators as proxies or substitute. For instance, it is quite difficult to measure the 'quality of education' because there are several factors that determine the quality of education, and experts may have different opinions on which these determinants count most. A very simple substitute indicator could be 'number of pupils passing the exam'.

Let us describe in some detail these features of an effective and good indicator.

Relevant

An indicator must be relevant. It must fit the purpose for measuring. As indicators, the gas gauge and the report card in examples given above, both measure facts that are relevant. If, instead of showing (measuring) the amount of gas in the tank, the gas gauge showed the hydrocarbon (octane) rating of the gasoline, it would not help you decide when to refill the tank. Likewise, a pupil report card that measures the number of pencils used by the student would be a poor indicator of academic performance.

Understandable

An indicator must be understandable. You need to know what it is telling you. For instance, there are many different types of gas gauges. There are gauges in the car which have a lever that moves between 'full' and 'empty' marks. Some other cars have gauges that use lights to achieve the same effect. Still some gauges point to the number of gallons of gasoline left in the tank. Indeed all the gauges are different, but each gauge is understandable to the driver. Likewise, with the report card, different schools use different ways of reporting academic progress. Some schools use letter grades A through E. Other schools use numbers from 100 to 0. Still other schools use written comments. Similar to the gas gauge, these different measures express the student's progress or lack of progress in a way that is understandable to the person reading the report card.

A report card that gave grades in ancient Latin or Greek script would be a mystery to most people. In order for you to know when action is needed, you must be able to understand what an indicator is telling you.

Reliable

An indicator must be reliable. You must trust what the indicator shows. A good gas gauge and an accurate report card provide information one can rely on. If the gas gauge shows that the tank of your car is half-full, you would stop the car for gasoline before it is needed. On the other hand, if the gas gauge that shows the tank is half-full when in fact it is empty would cause you to run out of petrol in an inconvenient place. Similarly, if a student's grade were reported wrong, he or she could be sent for remedial work and a student who needs help would not get it. An indicator is only useful if it shows the thing that one can believe and rely on.

You must remember that reliability is not the same as precision. When your gas gauge shows an empty gasoline tank, you know there is still a gallon or so of gasoline left as a reserve. The gas gauge under-reports the amount of gasoline. Thus, 'an indicator does not necessarily need to be precise'; it just needs to show a reliable picture of the system it is measuring.

Accessible Data

Indicators must provide timely information. They must give you information when you still have time to act. For example, imagine a gas gauge that only shows you the amount of gasoline in the tank when you started the engine of the car. After you have been driving for several hours, that reading would no longer be useful. You must know how much gasoline is there in the tank at each moment. Similarly, a report card distributed a week before the examination arrives too late to give a student remedial help. Thus, for an indicator to be useful in preventing or solving a problem, it must provide you the information while there is still time to correct the problem.

Consequently, good and effective indicators need to:

- Cover all relevant areas (i.e., should not have 'blind spots' where changes remain undetected)
- Be sensitive enough to detect changes when they occur
- Be informative, that is, provide evidence about the reason for a change
- Be measurable, statistical reliable, comprehensive and relevant
- Summarise information without distortions
- Describe central features of the programme implementation
- Be related to other indicators for global analysis of the system
- Be understandable to all major interest users
- Be applicable or useful at all levels (national, local and so on)

What Indicators Can and Cannot Do?

The expectations for education indicators are quite modest: to describe and state problems more clearly, to signal new problems more quickly, to obtain clues about promising educational programmes and the like.

Indicators Can Do

More specifically, you can use indicators to:

- Provide information at different levels for use by various actors within an indicator system
- Identify individual students for intervention or leverage points for schools to focus their efforts on
- Assess a student's progress or to evaluate the effectiveness of policies and practices for improving college readiness
- Identify areas for improvement in schools and across the district (system-level indicators) or to inform the process of planning supports and interventions

However, your choice for the selection of indicators and their validation should always depend on how they are intended to drive improvements in student outcomes and school processes.

While the ultimate purpose of an indicator system is to improve students' educational attainment, you can also use as mechanisms for districts to communicate their priorities and evaluate progress towards intermediate goals around school culture and educational practices.

Indicators Cannot Do

The literature on education indicators appears to agree on what indicators cannot do. They cannot help you to:

- Set goals and priorities: We know that it is the elected representatives of the public who establish educational goals and priorities. Indicators simply generate information and on the basis of this information can inform those goals and objectives. Indicators cannot help you set goals and priorities.
- Evaluate programmes: Since indicators do not provide the level of rigour or necessary details, they cannot be used as substitutes for well-designed, in-depth evaluations of education programmes.
- Develop a balance sheet: Education indicators lack the common referent available to economic indicators. As Rivlin (1973) pointed out:

> No amount of disaggregation of inputs...will provide a basis for answering the how-are-we-doing question in the education sector. As long as cost is used as a proxy for value there is no way to compare inputs with outputs or to see whether a given amount of education is being produced with fewer resources.

Rivlin also noted that 'because students help produce education, it is difficult to disentangle the quality of the output from the quality of student input'.

How Can You Select an Indicator?

While selecting an indicator in your research, you should be sure that the indicator makes it possible to:

- Measure how far or how close you are from an objective
- Identify problematic or unacceptable situations
- Meet policy concerns and to answer the questions leading to its choice
- Compare its value to a reference value, to a standard or itself, as computed for different observation period

Thus, choosing the most appropriate indicators can be difficult. Development of a successful accountability system demands involvement of several people when identifying indicators. These people should include those who will collect the data, those who will use the data and those who have the technical expertise to understand the strengths and limitations of specific measures. The following questions may guide your selection of indicators (Box 8.1).

Does this indicator enable you to know about the expected result or condition?

Indicators should, to the extent possible, provide the most direct evidence of the condition or result they are measuring. For example, if the desired result is a reduction in adult illiteracy, then to measure

BOX 8.1: General Criteria for Indicators

Indicators should be:

Generally relevant

- Related to a specific question or issue of concern
- Education-related and linked to environment/development factors
- Sensitive to changes in the conditions in question
- Give early warning of pending changes

Scientifically sound

- Unbiased and representative of the conditions in question
- Scientifically credible, reliable and valid
- Based on the best available data of acceptable quality
- Robust and unaffected by minor changes in the method or scale used in their construction
- Consistent and comparable over time and space

Applicable to users

- Relevant to policy and management needs
- Based on data that are available or can be collected or monitored with a reasonable financial/time resource input
- Easily understood and applied by potential users
- Acceptable to stakeholders

Source: Modified and adapted from Briggs (1999).

achievement you should best choose an outcome indicator, such as the adult illiteracy rate. The absolute number of adult illiterates receiving literacy instruction services would not be an optimal measure for this result; however, it might well be a good output measure for monitoring the service delivery necessary to reduce adult illiteracy rates.

You should note that proxy measures may sometimes be necessary due to data collection or time constraints. For example, there are few direct measures of school readiness. Instead, a number of measures are used to approximate this: children's participation in high-quality developmentally appropriate preschool, parents' exposure to parenthood education services and family literacy levels. When using proxy measures, you must acknowledge that you will not always provide the best evidence of conditions or results.

Is the indicator defined in the same way over time? Are data for the indicator collected in the same way over time?

To draw conclusions over a period of time, you must be certain that you are looking at data which measure the same phenomenon (often called reliability). The definition of an indicator must therefore remain consistent each time it is measured. Similarly, where percentages are used, the denominator must be clearly identified and consistently applied. For example, when measuring teen pregnancy rates over time, the population of girls from which pregnant teenagers are counted must be consistent (i.e., 10% of girls ages 12 to 18). Additionally, care must be taken to use the same measurement instrument or data collection protocol to ensure consistent data collection.

Data Sources and Data Collection Priorities for Indicators

Will data be available for an indicator? This is the key issue in the process of development of an education indicator system. National Statistical Offices must collect data on indicators frequently enough to be useful to decision-makers. The general trend is that data on outcomes are often available only on an annual basis; those measuring outputs, processes and inputs are typically available more frequently.

Are data currently being collected? If not, can cost-effective instruments for data collection be developed?

There has been increasing demand for accountability. At the same time, resources for monitoring and evaluation are decreasing. The MOE and several other agencies often collect data, especially data for input and output indicators and some standard outcome indicators. Where data are not currently collected, the cost of additional collection efforts must be weighed against the potential utility of the additional data.

Is this indicator important to most people? Will this indicator provide sufficient information about a condition or result to convince both supporters and sceptics?

Indicators which are publicly reported must have high credibility (UNESCO UIS 2009). They must provide information that will be both easily understood and accepted by the key stakeholders. However, indicators that are highly technical or which require a lot of explanation (such as indices) may be necessary for those more intimately involved in programmes.

Is the indicator quantitative? Numeric indicators often provide the most useful and understandable information to decision-makers. In some cases, however, qualitative information may be essential to understand the measured phenomenon.

Data are also required on indicators that you will use for an accountability system. A result-based accountability system often requires data on a number of different indicators, reflecting the information needs of different decision-makers. Educational planners and decision-makers frequently require information on long-term outcomes (and, in some cases, on inputs) while programme providers require details on inputs, processes and outputs as well as on outcomes (OECD 2015). For each indicator, baseline data should be collected to identify the starting point from which progress is examined. A comparison of actual indicator results to anticipated levels (often called performance standards or targets) permits decision-makers to evaluate the progress of programmes and policies. Assigning responsibility for indicator data collection to individuals or entities in an organisation helps to assure that data will be regularly collected (Parsons, Gokey and Thorton 2013).

It is important to note that indicators are red flags; good indicators merely provide a sense of whether expected results are being achieved. They do not answer questions about why results are or are not achieved, unintended results and the linkages existing between interventions and outcomes or actions that should be taken to improve results. You should always exercise utmost caution when interpreting indicators. You should use them to point to results that need further exploration, rather than as definitive assessments of programme success or failure.

Types of Education Indicators

In an educational setting, researchers and planners make use of several types of indicators. An education indicator provides information about the health and status of the education system. A statistics becomes an education indicator when it is useful in a policy context. For example, for your research, it would not be particularly useful to know that there are 1.5 million teachers and 35 million students in your country. These numbers describe the size of the system rather than its health. It would be more useful to compute a pupil–teacher ratio (PTR)—in this case, roughly 23.3. This statistics is an indicator.

These indicators serve as measures of progress of the programme. UNESCO UIS (European Commission 2009; UNESCO Institute for Statistics, 2009) suggests the following education indicators that researchers and educational planners might use in interpreting their research findings and in policy formulation, respectively.

Qualitative Indicators

Qualitative indicators provide a detailed description of complex phenomena based on interviews, documents or other sources of narrative information. They provide detailed information on the activities, outputs and outcomes of a project.

Qualitative indicators are used to measure the quality of the outcome of the system and also the quality of schooling inputs. They measure performance relative to some given standards and norms.

A list of qualitative indicators of inputs, process and outputs widely used in educational research and planning is given in Table 8.1.

Since qualitative measures are relatively flexible, these qualitative measures are particularly suited to issues that are complex, nuanced or where there is little existing information to provide a basis for quantitative measures. Researchers often combine qualitative indicators with quantitative measures to provide a detailed assessment of issues that are not easily quantifiable or to provide nuance and contextual detail to numerical findings.

TABLE 8.1 Qualitative Indicators

Inputs	• Number of trained teachers • Shortage of teachers • Surplus of teachers • PTR • Pupil–classroom ratio • Textbook–pupil ratio • Percentage of qualified teachers • Percentage of classrooms in good condition
Process	• Hours of study in class • Absenteeism (pupils, teachers) • Use of teaching aids
Outputs	• Percentage passing final examination • Employment status of graduates

Source: Authors.

Focus groups, in-depth interviews, observation and interviews constitute the most commonly used qualitative methodologies. Although in some research studies, researchers have tried to quantify the results of qualitative techniques, the most appropriate and useful analyses capture the main ideas of respondents through narrative text rather than through percentages and other statistics. Thus, qualitative indicators require a process to make them operational based on a total pre-order of scale values (i.e., ordinal scale). The qualitative indicators use scale values which have a total order, for example, by using labels such as 'poor', 'medium' or 'good' to express the order. Minimal qualitative scales contain just two values, for example, 'exists' and 'does not exist', or 'yes' and 'no'.

Qualitative indicators in programme evaluation complement quantitative techniques and quantifiable indicators and are particularly useful in four areas:

- Conducting needs assessments or formative research (to learn more about the local situation before designing the programme);
- Understanding the local context and terminology for a given subject (prior to finalising quantitative data collection instruments);
- Evaluating process (documenting the dynamics of how a programme works, as well as its strengths and weaknesses) and
- Developing a clearer and better understanding of the results obtained from a quantitative instrument (e.g., the attitudes, beliefs and values that underlie a given finding). By contrast, quantification is essential for measuring results and impact.

Quantitative Indicators

As the name suggests, quantitative measures use numerical summaries such as percentages, rates or absolute numbers. These measures are commonly used as indicators of activity and output, such as counts of the number teachers receiving in-service teachers training or the amount of food provided by a midday meal project. Quantitative indicators permit and are amenable to comparison over time or between settings. You could use quantitative indicators to track changes over time. However, indicators that are based exclusively on counting numbers can be misleading. However, you should make sure that valuable information is not lost in the process of turning complex concepts into a numerical measure.

Quantitative indicators are always designed as a proportion between two quantifiable factors—this makes the specific indicator independent of the size of the population being looked at in the indicator. The outcome in itself does not immediately provide information on whether the achieved and observed value is to be considered good or not. Rather, the comparison with other values facilitates the assessment of the quantitative indicator. You can use quantitative indicators for comparison either with values of the same country measured at different times (e.g., in the context of a trend analysis) to determine whether a certain situation develops in the intended direction or with values of other countries to facilitate benchmarking and mutual learning.

Public surveys and information from administrative records or budgeting systems are the most commonly used quantitative data sources. However, anything that can be counted can be turned into a quantitative measure. For example, observations of the proportion of teachers that are women, the average length of time taken to announce the examination results for students awaiting admissions at next level of education and counts of the number of articles in the local media mentioning school principals' corruption cases could all provide quantitative data.

Quantitative Indicators measure the amount or value of input or resources available. Table 8.2 lists the examples of these indicators.

In the paragraphs that follow, we discuss the most relevant qualitative and quantitative education indicators you may use in your research in areas such as monitoring and evaluating the performance and for measuring the cost-effectiveness and equity of educational services. It is important to remind you that indicators are precisely 'indicative' and cannot be a substitute for in-depth analysis and evaluative work.

A review of the extensive literature on social indicators and poverty measurement might show you the importance of an analytical distinction between types of indicators. You may notice a stress to classify them on the basis of the means, the process or the end in achieving the objective of a particular set of

TABLE 8.2 Quantitative Indicators

Inputs	• Student enrolments ▪ Gross enrolment ratio (GER) ▪ Net enrolment ratio (NER) ▪ Average daily attendance • Expenditure per pupil • Facilities and equipment ▪ Facilities ▪ Classrooms • Textbooks ▪ Issued to students ▪ Available in school library
Output	• Students progression • Examination passes

Source: Authors.

development policies, programmes or projects. Your attempt in systematic and good monitoring and evaluation should make use of an appropriate balance between different types of indicators that can establish the link between means and ends. The several classifications of indictors are roughly similar, though some important differences exist. Here, we propose to distinguish four types of education indicators:

- Input indicators;
- Process indicators;
- Output indicators and
- Outcome indicators

The purpose of each indicator is shown as follows:

Indicators			
Input (What comes into the system)	**Process** (What is done with the inputs)	**Output** (How many)	**Outcome** (What is the effect)

Input Indicators

Input indicators enable you measure the effectiveness of resources used in the education production activity. In other words, they measure the means or the resources employed to facilitate the satisfaction of learning needs and, hence, reaching development objectives. The determinants are characteristics of students, schools, teachers, facilities, instructional materials, equipment, cost and level of expenditures (public and private) on education. In each case, the term 'characteristics' refers to the availability of a resource, its nature and quality, and its manner and rate of use. Table 8.3 shows the most commonly used input indicators in an education system.

Since absolute numbers may not be very indicative for policy decisions, input indicators are often specified as some match of supply and demand variables, such as PTRs and average cost per pupil. The

TABLE 8.3 Input Indicators

Inputs	• Characteristics of learners • Teachers • Curriculum • Textbooks (teaching–learning materials) • Facilities • Equipment • Learner's capacity for learning • Other resources

Source: Authors.

latter type of indicators are input indicators, since they give an indication of the amount of services (inputs) in relation to some identified need or demand. You can use input indicators to address questions such as:

- Who participates in education?
- Who studies abroad and where?
- How successful are students in moving from education to world of work?

Input indicators enable you identify demand factors of potential users and may comprise variables that determine the use and accessibility of the supplied services. Examples of this type of indicators in education include the geographical distance to school facilities, family and cultural background of students, foregone earnings of individuals and households and direct private costs of education (fees, utensils, uniforms and so on). Some of these demand factors are essential in textbook analyses of the economics of education, but rarely are given due importance in educational information systems, let alone in the practical application of monitoring and evaluating educational programmes.

Process Indicators

Researchers and planners use process indicators to describe the important processes that contribute to the achievement of outcomes. The qualities of training, assessment and needs assessment are examples of process indicators. Basically, the process indicators are indirect indicators of merit and as such do not guarantee the achievement of outcomes. The process indicators highlight things that educational planners and decision-makers do that are expected to lead to desirable outcomes and which can be observed and described. Process indicators are particularly useful in reviews focused on improvement.

Process indicators are particularly useful in reviews focused on improvement. Where the outcome indicators suggest that the performance is below the expected level, you can use process indicators diagnostically to explore and identify the reasons. You can also use process indicators for developing recommendations for improvement.

Whereas the outcome indicators suggest performance above the level expected, process indicators can be used to validate or explain the processes that contributed to the outcomes. These indicators measure and determine the interaction that takes place among inputs. Normally, they require the collection of data through the observation of behaviour. Table 8.4 highlights key process education indicators.

TABLE 8.4 Process Indicators

Process	• Teachers • Learners • Administrators • Materials • Technology

Source: Authors.

Output Indicators

Output indicators measure the immediate effects of the educational activity, for example, attainment effect, achievement effects, attitudinal/behaviour effects and equity effects. These indicators are results of the education system and are readily observable upon completion of a level of education. They are reflection of value added.

The immediate objective of our educational policies is to raise the coverage of the educational system (as measured through enrolment ratios), improve its internal efficiency (retention rates) and/or raise the skills and knowledge of graduates (which can be measured through achievement tests). Output indicators attempts to measure to what extent such immediate objectives are achieved. In an educational setting, researchers and planners use the following three output indicators.

- Students' attainment
- Students' achievement
- Test scores

Outcome Indicators

There is often some confusion in differentiating output and outcome indicators. On many occasions, we use these terms interchangeably. Thus, understanding the difference between educational outputs and outcomes is important to developing meaningful assessment processes that position an institution to respond to these changes.

Better education may serve broader development goals, such as higher labour productivity, better health and enhanced capabilities of individuals to participate in modern society. Such 'higher' goals could be referred to as outcomes beyond the immediate influence of educational policies and programmes.

For developing meaningful expectations for student outcomes, you should make sure that you understand the institutional culture to which the students belong. Expectations for institutional outcomes should be aligned with the other elements of the institution. For example, if 30 per cent of entering students require remedial work in mathematics or writing (inputs), it should inform the nature of the services that are offered (processes). These services should be designed to enhance the likelihood of student success and may be such things as tutoring, remedial courses and interactions with teachers of high school feeder schools.

> If the services are well aligned with the inputs, they should have an impact on the quality of institutional outputs in the form of grades, graduation rates and employment statistics. In addition to looking at outputs as indicators of effectiveness, institutions are expected to demonstrate that student learning has taken place and that learning outcome goals have been met. When outcomes assessment data are collected, understanding how the results inform the effectiveness of other institutional processes will enhance the ability of institutions to improve their educational processes. (UNESCO Institute for Statistics 2009)

Outcomes relate to the likely or achieved short-term and medium-term effects of an intervention's outputs: better allocation of educational resources, increased affordability of education, quality of education, equitable access to education and optimal employment. These outcomes have both specific impact and intermediate impacts. Specific impacts cover positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended such as skills and learning enhancement (literacy and numeracy rates), examination outcomes and social views. On the other hand, immediate impacts are similar to specific impact but are longer in nature and are the last cause-and-effect chain level that can be monitored effectively and could be: greater income opportunities, improved participation in society and improved family planning and health awareness (Vos 1996).

Thus, outcome indicators are results and effects on individuals an society as a whole that are evident over time as a consequence or following interaction of educational outputs with the societal and socio-economic context. These education outcomes are effects more distant in time after completing education and are usually more dispersed in occurrence than education outputs. The main indicators of outcome are shown in Table 8.5.

In each of the above four core categories of education indicators, educational research look at in more details indicators of equity, efficiency, quality and performance and socio-economic/environment-related indicators. They are explained as follows:

TABLE 8.5 Outcome Indicators

Affordability Level Outcomes

1. Transparent decision-making
   Public current expenditure per pupil as per cent of GNP per capital
2. Cost for household—cost of education (i.e., fees, materials/equipment, lunches and so on) as a per cent of household expenditure; cost per pupil as a per cent of income per capita

Quality Outcomes

3. Absenteeism and dropout
   Teacher–pupil absenteeism rates; dropout rates; CE
4. Teachers ratios/rates
   PTRs; rate of double shifting
5. Qualification of teachers
   Average qualification of teachers at each educational level
6. Teaching material ratios
   Pupil–textbook ratio
7. Transition and repetition
   School life expectancy; transition rates, repetition rates, survival rates by grade, years-inputs per graduate

Equitable Access Outcomes

8. Equitable educational participation
   Ratios of girls to boys in primary, secondary and tertiary education; ratio of other disadvantaged groups compared to X per cent of population
9. Enrolment ratios
   Gross and net enrolment rates; apparent intake rate; age-specific enrolment ratio (ASER)
10. Completion rates
    Gross and net enrolment rates, on schedule completion at appropriate age; progression to next level of education
11. Distance to school
    X per cent of pupils within x kilometre of a school
12. School places
    Educational places as an X per cent of eligible pupils

Optimal Employment Outcomes

13. Employment: school construction and maintenance
    Number of people directly and indirectly employed in school construction and maintenance
14. Employment: teachers and government workers
    Number of people directly and indirectly employed as teachers and government workers linked to education sector

Specific Impacts

15. Literacy and numeracy rates
    Literacy and numeracy rates by age, sex and social groups.
16. Examination outcomes
    Examination outcomes for various educational levels and subject areas by age, sex and social groups; educational attainment of the population aged 25 years and above
17. Social views
    Support for violent conflict; prejudiced views towards other social groups

Intermediate Impact

18. Employability
    Per cent of school leavers/graduates gaining paid employment within x time (split by type of employment)

Source: UNESCO Institute for Statistics.

Equity Indicators

Equity is concerned with access (distribution) participation (opportunities) and achievements (consequences) in education.

The groups that are addressed in this category are distinguishable and overlapping and include females in general. In many countries, girls are invariably one of the disadvantaged categories. Other groups include the poor, refugees, minority language populations, ethnic minorities, nomads, refugees and street children. Also, there are disadvantaged groups in most countries—for example, many rural dwellers and inhabitants of poor urban districts populations living in areas with difficult access or of low demographic density.

Equity indicators in education are used to measure the degree to which provisions (physical, material, financial) for education are provided for the population regardless of economic status, place of residence and intellectual capability. These indicators are measures of achievement, fairness and opportunity in education. They also measure quality of access not only to physical facilities such as schools but also to quality education.

You can use equity indicators for interpreting and monitoring the extent and experience of educational opportunity for example between different student groups. The concept of equity refers to equality of educational opportunity among pupils in getting into and progressing through the education system and when leaving it.

The most commonly use equity indicators are:

- Gender Parity Index (GPI)
- Representative Index

Efficiency Indicators

Efficiency indicators are used to monitor the attainment of one of the programme's results at the least possible cost. They include both indicators of internal efficiency and external efficiency. Both type of indicators are given in Table 8.6.

TABLE 8.6 Efficiency Indicators

Indicators of Internal Efficiency	*Indicators of External Efficiency*
• Flow rates (promotion, repetition and dropout rates by grade) • Pupil years wasted (due to repetition, dropout) • Rates of transfers in/out by grade • Average study time (of dropouts, graduates) • Years of study per graduate • Survival rates by grade • Capacity utilisation rates (teachers and facilities) • Expenditure per learner	• How well are graduates prepared for integration into the labour force? • To what degree does their work maximise economic growth? • How much do their qualifications help improve social development? • To what extent has the education system provided school leavers with life skills to contribute to the mainstream? • What are the patterns of graduates' earnings over time? • What is the employment experience of graduates? ▪ Attainment effects ▪ Learning achievement effects ▪ Immediate attitudinal/behavioural effects ▪ Equity effects

Source: Authors.

Quality and Performance Indicators

Indicators of quality are usually performance indicators (Table 8.7). These indicators are defined as the relationship between one component of the educational system—pupils—and another such as teachers in terms of their interaction within the system. Two variables are combined to form a ratio indicator and thus it is measured or 'derived' from two variables.

For the country as a whole, appropriate indicators of educational quality are likely related to the overall performance of the system. These include such factors as system-wide dropout rates, repetition rates, levels of teacher absenteeism or disparity between male and female literacy rates in rural areas of the country.

Socio-economic/Environment-related Indicators

These indicators are external to the education system and include:

- Incidence of poverty
- Gender disparities
- Regional disparities
- Level of community involvement

We should recognise that regardless of how good the specification of indicators can be, there are sometimes limitations in the indicators themselves. These limitations emanate from the use of an

TABLE 8.7 Quality and Performance Indicators

Country level	• Student characteristics • Teacher and administrator characteristics • Curriculum, educational materials and instructional practices • Facilities and equipment (type, quantity, utility) • Attainment and achievement data • Education and training outcomes costs
Local level	• Contextual indicators (precondition for improving quality) ▪ Improved library facilities • Process indicators (what is taught and how it is taught) ▪ Quality of subject matter knowledge on the part of teachers • Learning performance indicator (quality of evaluation strategies) ▪ Assessment by teachers and student self-assessment • Improvement in educational services ▪ Average academic qualifications of teachers in the school ▪ Number of sitting and writing places in the classroom related to number of students in the class ▪ Number of textbooks ▪ Number of exercise books notebooks and so on ▪ Number of books shelves ▪ Amount of space in the classroom ▪ Number of toilets for girls ▪ PTR

Source: Authors.

indicator, including the matter of interpretation. It should be noted that some of the indicators mentioned or discussed above overlap with the listing of characteristics, but here the focus is on the factor of constraint.

Computation Techniques for Education Indicators

We have discussed above the several types of education indicators commonly used for measuring the educational system's performance. We describe and explain below the standard definitions proposed by UNESCO-UIS, 2009 for some selected education indicators which can measure the efficiency, effectiveness, equity and quality of the education system. UNESCO-UIS[1] suggests the following widely used key indicators of an education system.[2]

Indicator 1: Gross Intake Ratio (GIR) or Apparent Intake Rate in Grade I of Primary Education

- Definition: Total number of new entrants in the first grade of primary education, regardless of age, expressed as a percentage of the population at the official primary school entrance age.
- Purpose: To indicate the general level of access to primary education. It also indicates the capacity of the education system to provide access to Grade 1 of primary level for the official school entrance age population.
- Calculation method: Divide the number of new entrants in Grade 1, irrespective of age, by the population of official school entrance age and multiply the result by 100.
- Interpretation: A high GIR indicates a high degree of access to primary education. As this calculation includes all new entrants to first grade (regardless of age), the ratio can exceed 100 per cent, due to overaged and underaged children entering primary school for the first time.

Indicator 2: Net Intake Rate (NIR) in the Grade I of Primary Education

- Definition: New entrants in the first grade of primary education who are of the official primary school entrance age, expressed as a percentage of the population of the same age.
- Purpose: The indicator measures precisely the access to primary education by the eligible population of primary school entrance age.
- Calculation method: Divide the number of children of official primary school entrance age who enter the first grade of primary education for the first time by the population of the same age and multiply the result by 100.
- Interpretation: A high NIR indicates a high degree of access to primary education for the official primary school entrance age children. NIR of 100 per cent is a necessary condition for the policy goal of universal primary education.

Indicator 3: Transition Rate (TR)

- Definition: The number of pupils (or students) admitted to the first grade of a higher level of education in a given year, expressed as a percentage of the number of pupils (or students) enrolled in the final grade of the lower level of education in the previous year.

---

[1] Adapted from the UNESCO-UIS Databank 2017.

[2] For computation of indicators, refer to UNESCO Institute for Statistics.

- Purpose: To convey information on the degree of access or transition from one cycle or level of education to a higher one. Viewed from the lower cycle or level of education, it is considered as an output indicator, viewed from the higher educational cycle or level, it constitutes an indicator of access. It can also help in assessing the relative selectivity of an education system, which can be due to pedagogical or financial requirements.
- Calculation method: Divide the number of new entrants in the first grade of the specified higher cycle or level of education by the number of pupils who were enrolled in the final grade of the preceding cycle or level of education in the previous school year and multiply by 100.
- Interpretation: High transition rates indicate a high level of access or transition from one level of education to the next. They also reflect the intake capacity of the next level of education. Inversely, low transition rates can signal problems in the bridging between two cycles or levels of education, due to either deficiencies in the examination system or inadequate admission capacity in the higher cycle or level of education or both.

Indicator 4: Gross Enrolment Ratio

- Definition: Total enrolment in a specific level of education, regardless of age, expressed as a percentage of the eligible official school age population corresponding to the same level of education in a given school year.
- Purpose: To show the general level of participation in a given level of education. It indicates the capacity of the education system to enrol students of a particular age group. It can also be a complementary indicator to net enrolment ratio (NER) by indicating the extent of overaged and underaged enrolment.
- Calculation method: Divide the number of pupils (or students) enrolled in a given level of education regardless of age by the population of the age group which officially corresponds to the given level of education and multiply the result by 100.
- Interpretation: A high GER generally indicates a high degree of participation, whether the pupils belong to the official age group or not. A GER value approaching or exceeding 100 per cent indicates that a country is, in principle, able to accommodate all of its school age population, but it does not indicate the proportion already enrolled. The achievement of a GER of 100 per cent is therefore a necessary but not sufficient condition for enrolling all eligible children in school. When the GER exceeds 90 per cent for a particular level of education, the aggregate number of places for pupils is approaching the number required for universal access of the official age group. However, this is a meaningful interpretation only if one can expect the underaged and overaged enrolments to decline in the future to free places for pupils from the expected age group.

Indicator 5: Net Enrolment Ratio

- Definition: Enrolment of the official age group for a given level of education expressed as a percentage of the corresponding population.
- Purpose: To show the extent of coverage in a given level of education of children and youths belonging to the official age group corresponding to the given level of education.
- Calculation method: Divide the number of pupils (or students) enrolled who are of the official age group for a given level of education by the population for the same age group and multiply the result by 100.

- Interpretation: A high NER denotes a high degree of coverage for the official school age population. The theoretical maximum value is 100 per cent. Increasing trends can be considered as reflecting improving coverage at the specified level of education. When the NER is compared with the GER, the difference between the two highlights the incidence of underaged and overaged enrolment. If the NER is below 100 per cent, then the complement, that is, the difference with 100 per cent, provides a measure of the proportion of children not enrolled at the specified level of education. However, since some of these children/youth could be enrolled at other levels of education, this difference should in no way be considered as indicating the percentage of students not enrolled. To measure universal primary education, for example, adjusted primary NER is calculated on the basis of the percentage of children in the official primary school age range who are enrolled in either primary or secondary education. A more precise complementary indicator is the ASER which shows the participation in education of the population of each particular age, regardless of the level of education.

Indicator 6: Repetition Rate by Grade (RR)

- Definition: Proportion of pupils from a cohort enrolled in a given grade at a given school year who study in the same grade in the following school year.
- Purpose: To measure the rate at which pupils from a cohort repeat a grade, and its effect on the internal efficiency of educational systems. In addition, it is one of the key indicators for analysing and projecting pupil flows from grade to grade within the educational cycle.
- Calculation method: Divide the number of repeaters in a given grade in school year $t + 1$ by the number of pupils from the same cohort enrolled in the same grade in the previous school year t.
- Interpretation: Repetition rate ideally should approach 0 per cent. High repetition rate reveals problems in the internal efficiency of the educational system and possibly reflect a poor level of instruction. When compared across grades, the patterns can indicate specific grades for which there is higher repetition, hence requiring more in-depth study of causes and possible remedies.

Indicator 7: Survival Rate by Grade (SR)

- Definition: Percentage of a cohort of pupils (or students) enrolled in the first grade of a given level or cycle of education in a given school year who are expected to reach successive grades at that level of education.
- Purpose: To measure the retention capacity and internal efficiency of an education system. It illustrates the situation regarding retention of pupils (or students) from grade to grade in schools and conversely the magnitude of dropout by grade.
- Calculation method: Divide the total number of pupils belonging to a school cohort who reached each successive grade of the specified level of education by the number of pupils in the school cohort, that is, those originally enrolled in the first grade of primary education and multiply the result by 100. The survival rate is calculated on the basis of the reconstructed cohort method, which uses data on enrolment and repeaters for two consecutive years.
- Interpretation: Survival rates approaching 100 per cent indicate a high level of retention and low incidence of dropout. The distinction between survival rate with and without repetition is necessary to compare the extent of wastage due to dropout and repetition. Survival rate to the last grade of primary education is of particular interest for monitoring universal primary education, a central objective for EFA and the MDGs.

Indicator 8: Coefficient of Efficiency

- Definition: The ideal (optimal) number of pupil years required (i.e., in the absence of repetition and dropout) to produce a number of graduates from a given school cohort for a cycle or level of education expressed as a percentage of the actual number of pupil years spent to produce the same number of graduates. Input–output ratio, which is the reciprocal of the CE, is often used as an alternative. N.B.: one school year spent in a grade by a pupil is counted as one pupil year.
- Purpose: This is a synthetic indicator of the internal efficiency of an educational system. It summarises the consequences of repetition and dropout on the efficiency of the educational process in producing graduates.
- Calculation method: Divide the ideal number of pupil years required to produce a number of graduates from a given school cohort for the specified level of education, by the actual number of pupil years spent to produce the same number of graduates and multiply the result by 100. The CE is calculated on the basis of the reconstructed cohort method, which uses data on enrolment and repeaters for two consecutive years.
- Interpretation: Results approaching 100 per cent indicate a high overall level of internal efficiency in producing graduates and no wastage due to repetition and dropout. Coefficients below 100 per cent reflect the impact of repetition and dropout on the efficiency of the educational process in producing graduates. As the reciprocal, the optimum input–output ratio is one and inefficiency arises from any point which is greater than one.

Indicator 9: Years–Input per Graduate

- Definition: The estimated average number of pupil years spent by pupils (or students) from a given cohort who graduates from a given cycle or level of education, taking into account the pupil years wasted due to dropout and repetition (note that one school year spent in a grade by a pupil is equal to one pupil year).
- Purpose: To assess the extent of educational internal efficiency in terms of the estimated average number of years to be invested in producing a graduate.
- Calculation method: Divide the total number of pupil years spent by a pupil cohort (graduates plus dropouts) in the specified level of education by the sum of the successive batch of graduates belonging to the same cohort. This indicator is calculated on the basis of the reconstructed cohort method, which uses data on enrolment and repeaters for two consecutive years.
- Calculation method: Divide the total number of pupil years spent by a pupil cohort (graduates plus dropouts) in the specified level of education by the sum of successive batch of graduates belonging to the same cohort.
- Interpretation: The closer the value of this indicator is to the theoretical number of grades (or duration) of the specified education cycle, the higher the internal efficiency and the lesser the negative effects of repetition and dropout. A high number of pupil years per graduate as compared to the normal duration denotes waste of resources and hence inefficiency.

Indicator 10: Percentage of Repeaters

- Definition: Total number of pupils who are enrolled in the same grade as in a previous year, expressed as a percentage of the total enrolment to the specified grade.

- Purpose: To measure the extent and patterns of repetition by grade, as part of the internal efficiency of education system.
- Calculation method: Divide the number of pupils/students repeating a given grade in a given school year by the number of pupils or students enrolled in the same grade in the same school year and multiply by 100. It can be calculated for the whole level of education by dividing the sum of repeaters in all grades of the given level by the total enrolment of that level of education and multiply the result by100.
- Interpretation: High values reflect serious problems of grade repetition or the internal efficiency of the education system.

Indicator 11: Public Expenditure on Education as Percentage of Gross National Income (GNI)

- Definition: Total public expenditure on education (current and capital) expressed as a percentage of the GNI in a given financial year. GNI is also referred to as gross national product (GNP).
- Purpose: This indicator shows the proportion of a country's wealth generated during a given financial year that has been spent by government authorities on education. The indicator can be also calculated based on gross domestic product (GDP).
- Calculation method: Divide total public expenditure on education in a given financial year by the GNI of the country for the corresponding year and multiply by 100.
- Interpretation: In principle, a high percentage of GNI devoted to public expenditure on education denotes a high level of attention given to investment in education by the government and vice versa.

Indicator 12: Public Current Expenditure per Pupil (Student) as Percentage of GNI Per Capita

- Definition: Public current expenditure per pupil (or student) at each level of education, expressed as a percentage of GNI per capita in a given financial year.
- Purpose: To measure the share of per capita income spent on each pupil or student. It helps in assessing a country's level of investment in human capital development. When calculated by level of education, it also indicates the relative costs and emphasis placed by the country on a particular level of education. The indicator can be also calculated based on GDP.
- Calculation method: Divide per pupil public current expenditure on each level of education in a given year by the GNI per capita for the same year and multiply by 100.
- Interpretation: A high percentage figure for this indicator denotes a high share of per capita income being spent on each pupil/student in a specified level of education. It represents a measure of the financial cost per pupil/student in relation to average per capita income. A high level of spending per pupil should be interpreted with caution because this could simply reflect low enrolment. This indicator should therefore be used in conjunction with enrolment ratios. Low expenditure per pupil and low enrolment in primary education when compared to high expenditure and/or low enrolment in tertiary education suggests a need to reconsider resource allocations within the education sector, especially if universal primary education is a priority.

Indicator 13: PTR

- Definition: Average number of pupils (students) per teacher at a specific level of education in a given school year.

- Purpose: To measure the level of human resources input in terms of the number of teachers in relation to the size of the pupil population. The results can be compared with established national norms on the number of pupils per teacher for each level or type of education.
- Calculation method: Divide the total number of pupils enrolled at the specified level of education by the number of teachers at the same level.
- Interpretation: A high PTR suggests that each teacher has to be responsible for a large number of pupils. In other words, the higher the PTR, the lower the relative access of pupils to teachers. It is generally assumed that a low PTR signifies smaller classes, which enables the teacher to pay more attention to individual students, which may, in the long run, result in a better performance of the pupils.

Indicator 14: Adult Literacy or Illiteracy Rates

- Definition: Adult literacy rate is defined as the percentage of population aged 15 years and over who can both read and write with understanding a short simple statement on his or her everyday life. Adult illiteracy is defined as the percentage of the population aged 15 years and over who cannot both read and write with understanding a short simple statement on his or her everyday life.
- Purpose: Adult literacy rate shows the accumulated achievement of primary education and literacy programmes in imparting basic literacy skills to the population, thereby enabling them to apply such skills in daily life and to continue learning and communicating using the written word. Literacy represents a potential for further intellectual growth and contribution to economic-socio-cultural development of society. Illiteracy rates indicate the extent of need for policies and efforts in organising adult literacy programmes and quality primary education.
- Calculation method: Divide the number of literates by the corresponding age group population and multiply the result by 100. Alternatively, apply the same method using the number of illiterates to derive the illiteracy rate or by subtracting literacy rate from 100 per cent.
- Interpretation: High literacy rate (or low illiteracy rate) indicates a wide coverage of the primary education system and/or literacy programmes in that a large proportion of the population have acquired the ability of using the written word in daily life and to continue learning. It is common practice to present and analyse literacy rates together with the absolute number of adult illiterates as improvements in literacy rates may sometimes be accompanied by increases in the illiterate population due to the changing demographic structure.

Indicator 15: Number of Adult Illiterates

- Definition: The population aged 15 years and above who cannot both read and write with understanding a short simple statement on their everyday life.
- Purpose: The purpose of this indicator is to identify the size and if possible also the whereabouts and characteristics of the illiterate population aged 15 years and above who should be targeted for policies and efforts in expanding adult literacy programmes.
- Calculation method: Either use data on the number of adult illiterates collected during population census or survey or subtract the number of adult literates from the total population aged 15 years and above.
- Interpretation: The higher the illiterate population of the country, the more the need for expanding primary education and adult literacy programmes. When disaggregated by

geographical locations, it can pinpoint the areas needing most literacy efforts, and policies may be set to target such efforts at priority population groups of a particular gender and age group(s).

Indicator 16: Gender Parity Index

- Definition: Ratio of female to male values of a given indicator.
- Purpose: The GPI measures progress towards gender parity in education participation and/or learning opportunities available for women in relation to those available to men. It also reflects the level of women's empowerment in society.
- Calculation method: Divide the female value of a given indicator by that of the male.
- Interpretation: A GPI equal to 1 indicates parity between females and males. In general, a value less than 1 indicates disparity in favour of boys/men and a value greater than 1 indicates disparity in favour of girls/women. However, the interpretation should be the other way round for indicators that should ideally approach 0 per cent (e.g., repetition, dropout, illiteracy rates and so on). In these cases, a GPI of less than 1 indicates a disparity in favour of girls/women and a value greater than 1 indicates a disparity in favour of boys/men.
- Further to these basic indicators widely used in measuring the performance and progress of an education system, statisticians have also developed several other indicators. We present only those which are highly useful in educational research.

Index for Inclusion

Booth and Ainscow (2011) have developed a number of indicators to support the inclusive development of schools. The index offers schools a supportive process of self-review and development, which draws on the views of staff, pupils/students and parents, as well as other members of the surrounding communities. It involves a detailed examination of how barriers to learning and participation can be reduced for any pupil/student. The indicators cover three dimensions:

- Creating inclusive cultures (building community, establishing inclusive values)
- Producing inclusive policies (developing the school for all, organising support for diversity)
- Evolving inclusive practice (orchestrating learning, mobilising resources)

Quality Indicators in Special Needs Education

Hollenweger and Haskell (2002) have developed a number of quality indicators covering the aspects of educational inputs and resources, processes and results:

- Educational inputs and resources: policies, community characteristics, resources, personnel, students' characteristics, family characteristics
- Educational processes: state/school district practice, school building
- Level practice, classroom instructional practice, student-oriented domains
- Educational results for systems and individuals: academic and functional literacy, physical health, responsibility and independence, citizenship, personal and social well-being, satisfaction and so on.

Summary

This chapter tried to emphasise that an improvement of educational statistics should start off with a clear definition of what needs to be measured. A framework was sketched for selecting and systematising data for the construction of indicators for policy analysis. An analytical system of indicators is proposed which will enable to monitor the causal chain running from inputs and demand factors to outputs and outcomes.

The analysis of quantitative and qualitative data can be complex and time-consuming process. The chapter presented a discussion about the basic elements of data and its role as a key element of developing monitoring and evaluation indictors commonly used in the process of educational research. The chapter provided key educational indicators and a description of how and when they might be used in any research inquiry.

Finally, the chapter described the key elements to enable you undertake basic analysis of both qualitative and quantitative data that will encourage new researchers to research inquiry and give some reminders for those already engaged in educational research. It gave a helpful grounding to begin thinking about salient features of quantitative and qualitative data and education indicators that you may develop for your research in the field of education. Data collection and data analysis techniques are discussed in Chapters 9 and 10, respectively.

Self-test Exercise

Exercise 8.1:
8.1.1. Why do we use indicators in educational research?
8.1.2. What are the main characteristics of a good education indicator?
8.1.3. What do you understand by the term 'education indicator system'?

Exercise 8.2:
8.2.1. Which five indicators would you use to measure the educational access and coverage?
8.2.2. How would you calculate and interpret them? Give practical cases.

Exercise 8.3:
8.3.1. Which are the indicators you would propose to measure the 'quality of primary education'?
8.3.2. What are the limitations of these indicators?

Exercise 8.4:
8.4.1. In 2014, there were 25 million new entrants in Grade I of primary level education with a total official entrance age population (6 years old) of 20 million. Calculate the apparent intake rate.
8.4.2. How many children in Grade 1 are overage and underage? Calculate.

Exercise 8.5:
From the data given in the following table, calculate:
8.5.1. Students flow rates (repetition, dropout and promotion rates) by grades.
8.5.2. Transition rates
8.5.3. Survival rates
8.5.4. Wastage at primary level of education

Enrolment and Repeaters by Grade (2014–2015 and 2015–2016)

	Grades						
Year/Grade	*I*	*2*	*3*	*4*	*5*	*1–5*	*6*
Enrolment							
2014–2015	23,592	22,019	17,832	15,394	13,514	92,351	11,862
2015–2016	27,062	20,999	18,866	16,152	14,296	97,375	12,913
Repeaters							
2014–2015	1,005	696	719	688	668	3,686	1,109
2015–2016	1,230	598	576	628	571	3,603	634

References

Booth, T. and M. Ainscow. 2011. *Index for Inclusion: Developing Learning and Participation in Schools*. Bristol: CSIE.

Briggs, D. 1999. *Environmental Health Indicators: Framework and Methodologies*. Geneva: World Health Organization (WHO/SDE/OEH/99.10).

European Commission. 2009. *Outcome and Impact level Indicators Education Sector* (Working Paper: September 2009). EC External Services Evaluation Unit. Guidelines and UNESCO Institute for Statistics, Education Indicators, Technical Guidelines. Available at: http://ec.europa.eu/europeaid/how/evaluation/methodology/impact_indicators/wp_edu_en.pdf (accessed on 31 January 2018).

Hollenweger, J. and S. Haskell, eds. 2002. *Quality Indicators in Special Needs Education* (SNE). Lucerne: Edition SZH/SPC.

OECD. 2015. *Education at a Glance 2015: OECD Indicators*. Available at https://www.education.ie/en/Publications/Statistics/International-Statistical-Reports/Education-at-a-Glance-OECD-Indicators-2015-Briefing-Note.pdf (accessed on 17 November 2017).

Parsons, J., C. Gokey and M. Thornton. 2013. *Indicators of Inputs, Activities, Outputs, Outcomes and Impacts in Security and Justice Programming*. DFID. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/304626/Indicators.pdf (accessed on 31 January 2018).

Rivlin, A.M. 1973. 'Measuring Performance in Education'. In *Studies in Income and Wealth*, edited by M. Moss. New York, NY: Columbia University Press, 411–437.

UNESCO Institute for Statistics. 2009, November. *Education Indicators, Technical Guidelines*. UNESCO. Available at: http://uis.unesco.org/sites/default/files/documents/education-indicators-technical-guidelines-en_0.pdf (accessed on 31 January 2018).

Vos, Rob. 1996. *Educational Indicators: What's to be Measured* (Working Paper Series I-1). Washington, DC: Inter-American Development Bank.

9

CHAPTER

Data Collection: Observation, Interviews and Tests

Introduction

In Chapters 7 and 8, we explained the salient features of several types of data commonly used in educational research and educational indicators researchers use for monitoring and evaluation of programmes and projects, respectively. By now, it should be fairly clear that educational research involves the collection of data and that there are a variety of ways to do so. For example, if a teacher wanted to measure aggressive behaviour in children in his or her class, he or she would collect those data by observing children, by using equipment to measure the force with which they hit an object, by examining juvenile crime records, by surveying and interviewing parents and other teachers or by administering an aggression scale to children. This example does illustrate several distinctly different methods you can use to collect data. Research in education relies heavily on all such methods.

This chapter describes three data collection techniques—(a) observation, (b) interviews and (c) tests—commonly used in social sciences particularly in educational research. The chapter highlights some of the critical problems in communicating with the users of research outcomes (audiences); it then provides helpful information on how researchers can collect reliable data. The chapter also presents an overview of research instrument of data collection and variables included in the process of data collection. The other two important data collection methods, namely, surveys and questionnaire are discussed in chapter 10.

Data Collection Techniques

Data collection is a method in which information related to a study is gathered by suitable and reliable mediums. The procedures refer to how and when the data should be collected. The techniques used to collect data can influence research validity. It is, therefore, critical that data collection exercise should be a well thought out process, and its instruments should have both validity and reliability. In general, instruments of data collection are valid only when they measure what they are designed to measure.

In educational research and other social sciences, the following data collection techniques are widely used:

- Observations
- Interviews
- Tests
- Questionnaires
- Surveys

Observation Method

Perhaps the most interesting and challenging of all methods of data collection is the method of observation. With this method, the researcher records the behaviour of the participants by using his or her sensory systems (including eyes and ears). In this method, judgements about the occurrence of the behaviour, its frequency, its duration or its latency (the state of existing but not being developed or manifest) are made.

> Marshall and Rossman (2011) define observation as 'an observation is the recording of behavioural patterns of people, objects, and events in a systematic manner to obtain information about the phenomenon of interest'.

Observation method attempts to study individuals without interfering with their behaviour, if possible. The focus is upon what one can learn from individuals in their own habitat acting 'normally'. For Lofland (1971), 'the observation method is the most penetrating of strategies, the most close and telling mode of gathering information. The researchers can collect both qualitative and quantitative data by these methods'.

Observation is a qualitative method. The main objective of the method is to help researchers learn the perspectives held by study populations. In an observation, the researcher will make himself or herself part of the community that he or she is observing.

Types of Observation

You can observe participants in different ways:

- Overt observations: where everyone knows that they are being observed. They refer to the researcher being open about their intentions in the field and ensuring all members of the social group are aware of what is happening. Overt observation method allows the researcher to be honest with the participants, thus avoiding problematic ethical issues such as deception or lack of informed consent.

  However, one of the disadvantages of overt observation is that the participants understand the aims of the observer and so there is likely to be possible observer effects, that is, the participants of the study might change their behaviour by acting in a way that they believe is expected by the experimenter. An example of an overt observation would be the study conducted by Williams (1986)[1] on the media effects on anti-social behaviour in children. In this study, Williams assessed children 6–11 years old from an isolated society on their levels of aggression after the introduction of TV. The children were aware that they were being studied and so they may have shown demand characteristics or social desirability effects.

- Covert observations: In this method, no one knows they are being observed and the observer is concealed (K.M. DeWalt and B.R. DeWalt 2011). The researcher does not inform members of the group the reasons for their presence and keeps their true intentions secret. This automatically raises ethical concerns.

  An important advantage of the covert observation is that people are more likely to behave naturally if they do not know they are being observed. It allows the researcher access to social groups that normally would not provide consent to being involved in studies.

  An example of a covert observation would be the famous Bandura's 1961 Bobo doll study. First, children watched a model act violently. Then these children were unknowingly observed as they played with toys so that the researcher could measure the children's level of aggression after watching a model acting aggressively. The children did not know that they were being observed and also the reasons why they were involved in this study and therefore they would more likely to show natural displays of behaviour.

  For ethical reasons, you will typically choose overt observations because the respondents may not like that the information you gather from them should be concealed.

  Observations can also be either direct or indirect.

- Direct observations: Here the researcher watches interactions, processes or behaviours as they occur, for example, observing a teaching lesson from a written curriculum to determine whether the teacher delivering it with loyalty.

  You can use direct observation if the purpose of your research is to learn about the normal behaviour of people in their natural environments. A researcher observes and records while the action is happening. In direct observation, observers remain inconspicuous and do not allow their presence to affect the behaviour of the subjects they are studying. Educational researchers studying students' behaviour can observe the action occurring in schools, in homes or playground, depending on what aspect of behaviour they are studying.

- Indirect observations: In indirect observations, the researcher watches the results of interactions, processes or behaviours. He or she relies exclusively on the reported observations (including self-observations of others). For example, if you want to measure food wastage left by the student in the plate in the school dining hall to determine whether the students like a new menu. Indirect observation depends on the reports of others.

  Observations can be highly structured or non-structured.

- Structured observations: They consist of a checklist of, for example, the incidence, presence or frequency of predetermined evidence to be observed in the situation which will either support or refute a preconceived theory. While conducting structured observation, you should always determine beforehand the focus of the observation.

  Structured observation is the process of collecting data with two different characteristics. First, it is a part of the observational techniques where the investigators gather the required

[1] See http://www.thesuccessfulparent.com/categories/children-and-media/item/the-impact-of-tv-violence-on-children-and-adolescents#.Wg6l4VWWbIU (accessed on 17 November 2017).

information without any respondents, interviewees and so on, that is, the technique is a structured technique where the researcher collects data following some pre-defined rules and procedures. Structured observation can be applied to a wide array in educational settings.

Structured observations are highly useful when you look for direct information and try to understand the process, behaviour, event or situation. In case you feel that other data gathering methods are inappropriate, you should use structured observations because they directly measure the behaviour that provides direct information. The method is very easy to follow as well as time saving. You can also use this method in experiments as well as natural settings.

- Non-structured observations: They allow for issues to emerge from the observation. In a semi-structured observation, the researcher focuses on how to collect the data with observations and uses a predetermined process for this purpose.

  Non-structured observations are used often at the beginning of research study. The aim is to record information as much as possible. The information gathered not only includes the specified behaviour but also the context and surroundings of the behaviour. Data collected with this method are usually 'narrative data' (i.e., qualitative) rather than statistics (quantitative).

When to Use Observational Method in Educational Research?

Observation is one of the important methods of data collection. In educational research, you can use observation method for:

- Understanding an ongoing process or situation
- Gathering information on individual behaviours or interactions between people
- Knowing about a physical setting
- Linking quality improvement resources to positive outcomes for teachers and students
- Streamlining and promoting effective classroom interactions and teaching practices

In an education setting, you can make observations on many aspects of change in learning and teaching, for example, lectures, seminars, practical classes in lab, documents and other teaching material presented to students, learning resources, learning environments, interactions between participants. You are trying to understand an ongoing process or situation that you are monitoring and evaluating as it occurs.

Although there are several types of observational procedures or techniques that have been used to examine effective teaching (e.g., charts, rating scales, checklists and narrative descriptions), the most widely used procedure or research method has been systematic classroom observation based on student coding system (Phellas, Bloch and Seale 2011). The method specifies both the events/behaviours that are to be observed and how they are to be recorded. Generally, the data that are collected from this procedure focuses on the frequency with which specific behaviours or types of behaviour occur in the classroom and measure their duration.

With the students' coding system, you can record nearly everything that students and teachers do during a given time interval. These interaction are very objective and typically do not require the observer to make any high inferences or judgements about the behaviours they observe in the classroom. In other words, these low inference observational systems provide specific and easy identifiable behaviours that observers can easily code.

The other reasons for using observational methods in educational research are:

- To provide contextual information required framing the evaluation and making sense of data collected using other methods
- To develop insight into the learning and teaching context, the environment, events, activities, interactions, language used and so on
- To collect information about how a change in learning and teaching has been implemented independently of participant perceptions
- To collect different types of data to learn about sensitive issues that participants may be unwilling to talk about
- To identify and guide relationship with informants (participants)
- To get the feel for how things are organised and prioritised, how people interrelate and what are the cultural parameters (manners, leadership, politics, social interaction, taboos and so on)

What to Observe During an Observation?

When you plan to use observations for gathering data for your research study, there are certain important areas that you need to observe carefully. These areas are:

- Setting: The first thing you need to do is to observe and understand the setting (e.g., a child in the classroom) and its context from where you will gather the information. Observe what the setting is like, size, facilities, number of participants to be observed and so on. In other words, observe anything that might have an influence on the behaviour you are going to observe.
- Track events and event sequences: Observe carefully what participants of your study (teachers and students in the case of a classroom) do or what events take place during observation. Wait for and record a specific preselected behaviour. Reflect on the conditions under which particular behaviours occur or the frequency of behaviours.
- Counting, census taking and ethnographic mapping (e.g., decisions, kinships): Try to understand the environment in which the behaviour under study will take place. Are there relevant artefacts? Map out the spaces where relevant action could happen. This could involve drawing maps of retail space, picking up pamphlets, noticing billboards or logging television and radio messaging related to the topic at hand. The important thing here is to take in and understand the public messages that your participants are receiving around the topic at hand.
- Search for indicators of socio-economic difference: Search indicators that will help you establish the association between students' socio-economic status (SES) and academic achievement, making SES one of the powerful predictor of student performance.

Table 9.1 summarises the elements that you should watch while using the observation method for gathering information.

It is important to point out again that no single data source or methodology will sufficiently answer all critical educational questions. For enabling the researcher to capture a more comprehensive picture

TABLE 9.1 Elements to Be Observed During an Observation

Category	*Includes*	*Researchers should observe and note:*
Appearance	Clothing, age, gender, physical appearance	Anything that might indicate membership in groups or in sub-populations of interest to the study, such as profession, social status, socio-economic class, religion or ethnicity
Verbal behaviour and interactions	Who speaks to whom and for how long; who initiates interaction; languages or dialects spoken; tone of voice	Gender, age, ethnicity and profession of speakers; dynamics of interaction
Physical behaviour and gestures	What people do, who does what, who interacts with whom, who is not interacting	How people use their bodies and voices to communicate different emotions; what individuals' behaviours indicate about their feelings towards one another, their social rank or their profession
Personal Space	How close people stand to one another	What individuals' preferences concerning personal space suggest about their relationships
Human traffic	People who enter, leave and spend time at the observation site	Where people enter and exit; how long they stay; who they are (ethnicity, age, gender); whether they are alone or accompanied; number of people
People who stand out	Identification of people who receive a lot of attention from others	The characteristics of these individuals; what differentiates them from others; whether people consult them or they approach other people; whether they seem to be strangers or well known by others present

Source: Authors.

of what has been happening in classrooms, multiple measures or indicators of instruction are needed. Some of the new directions for classroom observation research could be:

- A combination of both qualitative and quantitative methods in observation systems
- Design of observation instruments in accordance to 'standards' of pedagogy
- Use of student-centred observation instruments that allow for comparisons between groups of students within the class
- Use of instruments that assess authentic, interactive instructional practices that have been found to relate to student gains on higher level cognitive outcomes

How Can You Plan Your Observational Method?

There are certain key considerations that you have to keep in mind prior the collection of data using the observation method. These are described next.

- Determine the focus of your observations: Reflect on the evaluation question(s) you want to answer through observation and select a few areas of focus for your data collection. For example, you may want to know how well and effectively an HIV curriculum is being implemented in the classroom. Your focus areas might be interactions between students and teachers, and teachers' knowledge, skills and behaviours.

- Design a system for data collection: Once you have focused on your evaluation, think and identify the specific items for which you want to collect data followed by the manner you will use to collect the information you need. You can choose any of the three following ways for collecting observation data. You can also combine these three methods to meet your data collection needs.
    - Recording sheets and checklists: These are the most standardised way of collecting observation data and include both preset questions and responses. These forms are typically used for collecting data that can be easily described in advance (e.g., topics that might be covered in an HIV prevention lesson).
    - Observation guides: List the interactions, processes or behaviours to be observed with space to record open-ended narrative data.
    - Field notes: They are the least standardised way of collecting observation data and do not include preset questions or responses. Field notes are open-ended narrative data that can be written or dictated onto a tape recorder.
- Select the sites: You have to identify and select an adequate number of sites to ensure that they are representative of the larger population and provide an understanding of the situation you are observing.
- Select the observers: You may choose yourself as the only observer or you may like to include others in conducting observations. The potential observers could be: stakeholders, other professional staff members, interns and graduate students.
- Train the observers: It is critical that the observers are well trained in your data collection methods and processes so as to ensure high-quality, consistent and reliable data. The level of training will vary on the complexity of the data collection and the individual capabilities of the observers.
- Schedule your observations appropriately: Programmes and processes typically follow a sequence of events. It is critical that you schedule your observations in a way that you can observe the components of the activity that will answer your evaluation questions. This requires advance planning.

Advantages and Limitations of Observational Method

Advantages

The novelty of this method is its simplicity. It is difficult neither for the enumerator nor for the informants because both are present at the spot of data collection. Since the investigator collects the information personally, this method provides most accurate information. But as the investigator alone is involved in the process, his or her personal bias may influence the accuracy of the data (K.M. DeWalt and B.R. DeWalt 2011). The method, therefore, demands an honest, unbiased and experienced investigator. In such cases, the data collected may be fairly accurate. However, the method is quite costly and time-consuming. Thus, the method should be used primarily for those research inquiries whose scopes are relatively small.

Observations permit the measurement of actual behaviours rather than reports of intended or preferred behaviours. The method is free from reporting bias, and potential bias caused by the interviewer and the interviewing process is either eliminated or reduced. There are specific types of data that can be collected only by observation. If the observed phenomenon happens frequently or is of short duration, observational methods may be cheaper and faster than survey methods. ~~As such~~:

- Observations can provide good insights into how different participants are behaving and interacting.
- Observations may enable the researcher see things that are taken for granted by participants in the learning–teaching context.
- They help the researcher to collect data where and when an event or activity is occurring.
- The method does not rely on people's willingness or ability to provide information.
- It allows you to directly see what people do rather than relying on what people say they did.

Limitations

The reasons for the observed behaviour may not be determined since little is known about the underlying motives, beliefs, attitudes and preferences. There is likelihood that bias in the researcher's perception can bias the data. Observational data are often time-consuming and expensive. Quite often researchers encounter problems in observing certain forms of behaviour. In some cases, the use of observational methods may be unethical, as in observing people without their knowledge or consent. It is best to view observation as a complement to survey methods, rather than as being in competition with them. Thus,

- Observations can be time-consuming.
- Observations of activity may affect the behaviour of those involved in it and hence what you observe.
- The thinking that underlies participants' observed actions could not be observed.
- Being able to make sense of the context of evaluation in a limited amount of time with limited resources may require some knowledge of the academic discipline and its culture.
- They do not increase our understanding of why people behave as they do.

To sum up, the process of conducting this type of field work involves several aspects such as gaining entry into the community, selecting key informants, participating in as many different activities as are permitted by the community members, clarifying one's findings through member checks, formal interviews and informal conversations and keeping organised, structured field notes to facilitate the development of a narrative that explains various cultural aspects to the reader. An observation is used as a mainstay in fieldwork in education research, and, as such, has proven to be a beneficial tool for producing studies that provide accurate representation of a culture.

In addition to helping define clear goals and creating a shared purpose, classroom observations can provide an objective method for measuring progress towards these goals and provide a framework for giving constructive and focused feedback that helps teachers incorporate higher levels of desired behaviours into their interactions with students in the classroom.

Interviews

Interviews have been used extensively for data collection across all the disciplines of the social sciences and in educational research in particular. Interviews are the most widely used qualitative design for gathering information. They are used to find social meanings and to answer the critical question of 'why'.

In other words, they help you find out about things that cannot be seen or heard such as the interviewee's inner state, the reasoning behind their actions and feelings. Interviews are also used to get detailed and extensive data, called 'life histories', which is common in feminist research. The method is highly useful when the subject matter is sensitive or complicated and involves only one interviewee.

> 'An *interview* is a process of asking questions and getting answers from participants in a study. Interviews are surveys that are administered verbally, either individually or in groups' (Kvale 1996).

Research interview aims at exploring the views, experiences, beliefs and/or motivations of individuals on specific matters. Interviews are considered to provide a 'deeper' understanding of social phenomena than would be obtained from purely quantitative methods such as questionnaires (Silverman 2000). Interviews are, therefore, most effective and appropriate where little is information is available about the study phenomenon or where detailed insights are required from individual participants. They are also particularly useful for exploring sensitive topics, where participants may not want to talk about such issues in a group environment.

An interview is a qualitative method of gathering evidence, data or information. Responses offered by interviewees are not usually expressed in numerical terms, as might be the case with questionnaires. Interviews are more reactive measures than are paper-and-pencil questionnaires. For this reason, interviewers should have a complete knowledge of training in conducting the interview. This is especially true when more than one interviewer is gathering data.

Interviews are particularly useful for getting the information behind a participant's experiences. The interviewer can chase in-depth information around a topic. You can also use interviews as follow-up to certain respondents to questionnaires, for example, to further investigate their responses. Usually interviewer asks open-ended questions during interviews.

You can interview people in a variety of forms including:

One-to-one: where interviewees are seen

Focus groups: where the interviewer (called a 'facilitator' or 'moderator') guides the discussion among a small group of respondents. One of the strengths of group discussions is the insight they offer into the dynamic effects of interaction on expressed opinion.

Telephone/online: where standardised interviews are usually the preferred medium. Semi-structured or non-structured interviews are sometimes conducted in this manner. The interviewer, however, cannot observe body language using this method.

The Interview Format

An interview is not a conversation. You have to follow a rigorous format for conducting an interview properly. You will have to watch your body language and make sure you maintain eye contact. For doing an interview, you have to keep in mind the following steps:

Step 1: A clear idea and understanding about: (a) who you will interview, (b) what kind of information you want to obtain and (c) the type of interview that will help you to do that.

Step 2: Interviews are different from questionnaires as they involve social interaction. Unlike questionnaires, you need training in how to interview which costs money.

Step 3: Before an interview takes place, respondents should be informed about the study details and given assurance about ethical principles such as anonymity and confidentiality.
Step 4: Conduct interviews in areas free from distractions and at times and locations that are most suitable for participants. Establish rapport with participants prior to the interview.
Step 5: Ensure that the interview is as productive as possible, that comprehensive and representative data are being collected during the interview. Listen attentively to what is being said without unnecessary interruptions.
Step 6: Ask different types of questions which help you in generating 'different types of data'. For example, closed questions provide respondents with a fixed set of responses, whereas open questions allow respondents to express what they think in their own words.
Step 7: Use an interview schedule. An interview schedule is a set of prepared questions designed to ask exactly as worded. Interviews schedules have a standardised format, that is, the same questions should be asked to each interviewee in the same order (see Appendix 8.1 as an example of an interview schedule).
Step 8: Record interviews and write the data generated by the interview as a transcript—a written account of interview questions and answers) which can be analysed at a later date.
Step 9: Exercise special care when interviewing vulnerable groups such as the children. Avoid lengthy interviews.
Step 10: Use appropriate vocabulary and language for the group of people being studied. For example, change the language of questions to match the social background of respondents' age, educational level, social class, ethnicity and so on.
Step 11: At the end of the interview, thank participants for their time and ask them if there is anything they would like to add, anything that have not been dealt with by the interviewer. This can often lead to the discovery of new, unanticipated information. Respondents should also be debriefed about the study after the interview has finished.

It should, however, be noted that interviews may not be the best method to use for researching sensitive topics, for example, truancy in schools (staying away from school without good reason), discrimination and so on as people may feel more comfortable completing a questionnaire in private.

Another equally important aspect for conducting an interview is communication skills. While doing interviews, you should always keep in mind the following tips.

- Your questions should be open-ended as much as possible for the information to be spontaneous rather than a rehearsed position.
- Your questioning style should encourage interviewees to be as frank as possible. This should help overcome any inhibitions a shy or timid interviewee may have.
- Your prompting skills should encourage the respondent to produce an answer. The mildest technique is to repeat the question. If this fails, rephrase the question slightly.
- Your probing skills should help encourage the interviewee to give a full response as much as possible. It means you should start follow-up questioning to get a fuller response. This can be verbal or non-verbal. An expectant glance can work as much as a direct request like 'please tell me more about that' or 'any other reasons?' Try to keep your probe as neutral as possible. You should not force your interviewee to state something that is not true to them (Roulston 2010).

Design of Interviews

When designing interviews, the first thing you have to do is to choose whether to use a structured or non-structured interview. Next, you must decide who will be the interviewer, and this will depend on what type of person you are going to interview. You should consider the following variables:

- Gender and age: This can have a big effect on respondents answer, particularly on personal issues.
- Personal characteristics: Some people are easier to get on with than others. Also, your accent and appearance (e.g., dress) can have an effect on the rapport between you and the interviewee.
- Ethnicity: Researchers have difficulty in interviewing people from a different ethnic group.

Kinds of Questions in Interviews

Interviews have become the main data collection procedure closely associated with qualitative, human scientific research. Kvale (1996) has written extensively on this subject matter and his books and articles on interviewing are probably the most cited in the entire field of qualitative research.

Kvale has identified the following nine types of questions to be asked in qualitative interviews. Remember these tips while composing your interview guide.

- Introducing questions: 'Why did you...?' or 'Can you tell me about...?' Through these questions, you introduce the topic.
- Follow-up questions: Through these, you can elaborate on their initial answer. Questions may include: 'What did you mean...?' or 'Can you give more detail...?'
- Probing questions: You can employ direct questioning to follow up what has been said and to get more detail. 'Do you have any examples?' or 'Could you say more about...?'
- Specifying questions: Such as 'What happened when you said that?' or 'What did he say next?'
- Direct questions: Questions with a yes or no answer are direct questions. You might want to leave these questions until the end so you do not lead the interviewee to answer in a certain way.
- Indirect questions: You can ask these to get the interviewee's true opinion.
- Structuring questions: These questions move the interview on to the next subject. For example, 'Moving on to...'
- Silence: Through pauses, you can suggest to the interviewee that you want them to answer the question.
- Interpreting questions: 'Do you mean that…?' or 'Is it correct that…?'

Sampling for Interviews

When you design your research project, you need to take into account how many people you need to interview to make the research valid or for 'population validity'. If the scope of your investigating is a narrow but deep subject, you may not need to interview that many people. You may be interested in the opinions and experiences of experts or people with direct experience—a purposive rather than a random sample (Seidman 2013). In this case, the sample size should be as appropriate as possible for your research. While in quantitative research, you should normally use questionnaires.

Preparing an Interview Guide

When preparing an interview guide (a clear set of instructions for interviewers), you need to keep in mind the following points.

- Identify your topic that is appropriate to be studied by interviewing.
- What is it you find problematic about the topic? In other words, what is your puzzle?
- Jot down questions that express your puzzlement with the topic.
- Try to expand the range of enquiry through a literature review and discussion with colleagues, friends and so on.
- Who are your subjects?
- What is your setting?
- Design your probes.

Example of an Interview Guide

Subjects	*Topic Areas*	*Probe Questions*
Young Asian women	Family history Parents Siblings Wider kin Education and work Leisure The future	Leisure What do you do in your spare time? Do you go out often? With Whom?

In addition to these points, also keep in mind the following points as suggested by Kvale (1996):

- Make sure you introduce yourself and explain the aim of the interview. Also, adhere to academic ethics by making sure the interviewee is fully aware of the purpose of the research.
- Devise your questions so they help to answer your research question, and make sure all the questions are relevant.
- Try to have a sequence to your questions or topics by grouping them in themes that follow a logical sequence.
- Make sure that you can easily move back and forth between questions or topic areas, as your interviewee may naturally move on to another subject.
- Make sure your questions are clear and easy to understand—only use technical or academic language if you are sure your interviewee will understand what you mean.
- Do not ask leading questions leading question (a question asked in a way that is intended to produce a desired answer) in a sentence. Make sure people are free to give their own, honest answers.

Transcription of Interviews

Once you have completed the interview, its transcription is very important. It reflects the following: How the interview went? Was there anything you could have done better? Is there a need to add any questions or topic areas? Is there anything you should have explained to the interviewees?

An accurate transcription (a written or printed representation of your interview) makes sure that all the quotations are authentic and correct. If an interviewee speaks in a clear manner, it saves a great deal of time because an accurate transcription requires fewer playbacks during the transcription process. Thus,

- Record the interview with a tape recorder and take notes.
- Write it immediately or as soon as possible. Do not sit on it. You will forget what had been said and how.
- Interview scripts are written like a drama script. For example:

  Interviewer: Hello, how are you today?
  Teacher: (coughs) I'm fine, though I've had better days.
  (Teacher takes out a tissue and blows her nose)

Analysis of Interviews

Upon the completion of transcription of your interview(s), you may have a lot of data. How could you analyse all of this? There will certainly be some data which are of no use perhaps because the interviewee did not keep to the subject or gave background information which is not needed.

Of the relevant information, you just pick out key points and quotes to illustrate your key points. Another thing that you could do is that you can code the information, that is, you could turn a qualitative interview into quantitative data. The simplest way to do this is by identifying passages of text and applying labels to them to show that they are an example of a theme. Suppose, if you asked 20 people how they travelled to work and one of the answers given was 'by car', this would be your thematic code 1. 'By bike' could be code 2 and 'walking' as code 3 and so on. By this coding system, you can analyse the data in a spreadsheet, which in turn will give you the chance to generate charts and graphs to better illustrate your answers. You could also use a qualitative research tool such as NViVO, a programme that helps you to classify your data using codes. Alternatively, if your sample is a small sample, you just design a table on a piece of paper listing how many people said 'car' and how many said 'bike'.

When to Use Interview

Interviews can be used at any stage of the evaluation process. They are especially useful in answering questions such as those suggested by Patton (1990):

- What does the programme look and feel like to the participants and to other stakeholders?
- What are the experiences of programme participants?
- What do stakeholders know about the project?
- What thoughts do stakeholders, who are knowledgeable about the programme, have concerning programme operations, processes and outcomes?

- What are participants' and stakeholders' expectations?
- What features of the project are most salient to the participants?
- What changes do participants perceive in them as a result of their involvement in the project?

Specific circumstances for which interviews are particularly appropriate include

- Complex subject matter
- Detailed information sought
- Busy, high-status respondents
- Highly sensitive subject matter

Advantages and Limitations of Interviews

Advantages

- Usually yield richest data, details, new insights
- Permit face-to-face contact with respondents
- Provide opportunity to explore topics in depth
- Afford ability to experience the affective as well as cognitive aspects of responses
- Allow interviewer to explain or help clarify questions, increasing the likelihood of useful responses
- Allow interviewer to be flexible in administering interview to particular individuals or circumstances

Limitations

- Expensive and time-consuming
- Need well-qualified, highly trained interviewers
- Interviewee may distort information through recall error, selective perceptions, desire to please interviewer
- Flexibility can result in inconsistencies across interviews

Do's and Don'ts of Interviewing

The rules of interviewing follow the rules which govern most human interactions and the rules which govern most investigative and problem-solving processes. In effect, they can be called the rules of the 'game'.

Do's

- Ask questions which start with who, what, where, when, why and how, where possible
- Ask both open and closed questions
- Verify understanding through probing and confirming questions
- Avoid confrontation
- Act in a friendly but professional manner
- Listen actively
- Take notes, but do not be obtrusive about it

- Let the interviewee do most of the talking
- Establish rapport early and maintain it
- Maintain control over the subject matter
- Establish a time frame for the interview and stick to it
- Conclude positively
- Allow for follow-up or clarification interviews later on
- Be polite and courteous

Don'ts

- 'Not' assume anything
- 'Not' form prejudgements
- 'Not' interrupt
- 'Not' go off on tangents
- Volume of information too large; may be difficult to transcribe and reduce data

Types of Research Interviews

There are many types of interviews, as suggested in the literature. However, in this section we attempt to look at those types of interviews, which are frequently used in educational research and report on how we can use these research methods to collect data.

Interviews may be highly formalised and structured, using standardised questions for each research participant (often called a respondent) or they may be informal and non-structured conversations. In between, there are intermediate positions. Table 9.2 shows the different types of research interviews.

Structured Interviews

A 'structured interview' or a standardised interview is a quantitative research method commonly used in survey research. The purpose of this method is to confirm that each interview is offered with exactly the same questions in the same order. This guarantees that answers can be reliably collected and that comparisons can be made with confidence between sample subgroups or between different survey periods.

TABLE 9.2 Types of Interviews

Interviews					
	Non-standardised				
Standardised	*One-to-One*			*One-to-Many*	
Interviewer-Administered Questionnaire	Face-to-Face Interviews	Telephone Interviews	Internet and Intranet-mediated (Electronic Interviews)	Group Interviews	Internet and Intranet-mediated (Electronic) Group Interviews
				Focus Groups	Focus Groups

Source: Authors.

In this type of interview, we use a formal interview schedule with a list of specific questions. The wording of the questions and the order in which they are asked is in the same form from one interview to another. The interviewer holds the interview schedule and ticks the boxes on behalf of the interviewee. The interviewee does not really get the opportunity to speak openly on an issue. The method relies on a predetermined coding system.

The interviewer using this approach prepares a set of open-ended questions which are carefully worded and arranged for the purpose of minimising variation in the questions posed to the interviewees. In view of this, this method is often preferred for collecting interviewing data when two or more researchers are involved in the data collecting process. Although this method provides less flexibility for questions, probing is still possible, depending on the nature of the interview and the skills of the interviewers (Patten 2007).

Strengths of Structured Interviews

- Structured interviews are easy to replicate as a fixed set of closed questions are used, which are easy to quantify—this means it is easy to test for reliability.
- Structured interviews are fairly quick to conduct and thus allow many interviews to take place within a short span of time. This means a large sample can be obtained resulting in the findings being representative and having the ability to be generalised to a large population.

Limitations of Structured Interviews

- Structure interviews are not flexible. This means new questions cannot be asked impromptu (i.e., during the interview) as an interview schedule must be followed.
- The answers from structured interviews lack details as only closed questions are asked which generates quantitative data. This means a research would not know why a person behaves in a certain way.

When to Use Structured Interviews

You should only use this form of interview if you already know what is happening in relation to your research topic and where there is no danger of loss of meaning as a result of imposing a standard way of asking questions. You should, therefore, best use structured interviews when the literature in a topical area is highly developed or following the use of observational and other less structured interviewing approaches that will provide you an adequate understanding of a topic to construct meaningful and relevant close-ended questions.

Semi-structured Interviews

With this type of interview, the interviewer asks major questions in the same way each time but he or she is free to alter the sequence and to probe for further information. He or she can adapt the research instrument to the level of comprehension and clarity of the interviewee. The interviewee can speak more openly. Box 9.1 presents an example of semi-structured interview.

Characteristics of Semi-structured Interviews

- The interviewer and respondents engage in a formal interview.
- The interviewer develops and uses an ‘interview guide’.
- The interviewer follows the guide. In the process of interviewing, he or she may follow topical trajectories in the conversation that may stray him or her from the guide.

BOX 9.1: Class, Gender, (Hetero) Sexuality and Schooling: Working-class Girls' Engagement with Education and Post-16 Age Aspirations

Context

Working-class girls may not be doing as badly as working-class boys, but a significant number are leaving school at the age of 16 with few or no qualifications. In order to explain this, the study draws attention to two processes—behaving in a hyper-heterosexual manner and working girls' expectations after leaving school at age 16 years.

Methods

The researchers used a multi-method, mainly qualitative, approach. First, data were collected from 89 pupils aged 14 to 16 using semi-structured interviews from six schools chosen because they served working-class areas suffering from severe economic and social deprivation.

The sample of 89 pupils was made up of pupils who had been identified by their schools as being at risk of dropping out of schooling at age 16. The sample included boys and girls from a variety of ethnic backgrounds.

Discussion groups were set up with an additional 36 pupils. Third, eight female pupils were asked to complete photographic diaries, focusing on their everyday activities and interests. Finally, semi-structured interviews were conducted with 19 members of staff and a small sample of five parents.

Findings

The study revealed that most of the female pupils were keen to be seen as 'desirable' and 'glamorous'. They spent a great deal of time and effort working on their hair, make-up and dress styles, in order to construct what the researchers called a 'sexualised hyper-feminine identity'. This 'work' was regarded by the girls as far more important than the academic work demanded by the school. The primary importance placed on appearance was highlighted in the sample's photo diaries, which included pictures of their favourite glamour products.

Evaluation

The strength of this study is its multi-strategy approach to gathering a range of qualitative data over a significant period of time. The longitudinal nature of the research allowed trends over time to be identified and the development of pupils to be regularly monitored in terms of their interaction with teachers and their peer group.

The qualitative data were obtained from both the teachers and the pupils so that the researchers managed to obtain the trust of both parties. For the pupils, guarantees of anonymity and confidentiality contributed to this. However, although extensive qualitative data resulted from the group discussions, we need to be aware that the validity of the data can be affected by peer pressure and fears of ridicule and exclusion.

Source: Adapted from Archar, Halsall and Hollingworth (2007).

Recording Semi-structured Interviews

Typically, the interviewer has a paper-based interview guide that he or she follows. Since semi-structured interviews often contain open-ended questions, discussions may diverge from the interview guide; it is generally best to tape-record interviews and later transcript these tapes for analysis.

While it is possible to try to jot down notes to capture respondents' answers, it is difficult to focus on conducting an interview and at the same time jotting notes. The notes can be poor. If tape recording an interview is not possible, a note taker could help the researcher.

Advantages of Semi-structured Interviews

Semi-structured interview is a preferred method for many researchers because questions can be prepared ahead of time. The method allows the interviewer to be well prepared and appear knowledgeable, skilled and competent during the interview. Semi-structured interviews also permit informants to feel free and relaxed to express their views in their own terms. Semi-structure interviews provide reliable, comparable qualitative data.

Limitations of Semi-structured Interviews

The key limitations of semi-structured interviews are as follows:

- Only a relatively small number of these interviews can take place because each interview can last for a long time.
- It is difficult to directly compare the results of in-depth interviews because each interview is unique.
- Because the sample size is small, the results are unlikely to be representative of a particular population.
- They are time-consuming—both in terms of data collection and data analysis.

When to Use Semi-structured Interviews?

Semi-structured interviewing is best used when you would not get more than one chance to interview someone and when you will be sending several interviewers out into the field to collect data. You can use semi-structured interviews when the aim of your study is to reveal people's knowledge, views, understandings, interpretations and experiences, especially their situational aspects.

Non-standardised Interviews

In research, this term is used interchangeably with the terms, informal conversational interview, in-depth interview, non-standardised interview and ethnographic interview. These are interviews in which neither the question nor the answer categories are predetermined. Instead, their basis is social interaction between the researcher and the informant.

This type of interview has no interview schedule (a set of questions with structured answers containing 'the opening', 'the body' and 'the closing'). Instead, it uses an interview guide. This means that interviewers only have a list of topics to cover with the interviewees. The interviewee is free to speak openly. Questions can be phrased as the interviewer wishes and he or she can ask them in any order. Discussion usually begins as a general conversation. If necessary, the interviewers can provide prompts and can join in the interview by discussing what they think of the topic themselves. The interview should be organised in an environment that will foster informality in case the interviewer is carrying out an in-depth work. The whole idea is to create a friendly, relaxed atmosphere. The non-structured interview is the most common form of qualitative interview and is used mostly for exploration in little researched areas. However, an important prerequisite of this technique is that you should be a skilled interviewer to conduct this type of interview properly. It is a lot harder than it sounds.

Recording Non-standardised Interviews

Non-structured interviews often contain open-ended questions. Therefore, there is every possibility for developing discussions in unanticipated directions. The best thing the research could do is to tape record interviews and later transcript these tapes for analysis. In this way, the interviewer can focus on interacting with the participant and follow the discussion.

The interviewer cannot focus on conducting an interview while jotting down the notes to capture respondents' answers. The quality of notes will be poor and also it will detract rapport between interviewer and interviewee. Development of rapport and dialogue are essential in non-structured interviews.

Strengths of Non-standardised Interviews

Non-structured interviews help us appreciate culture, experiences or setting which have not yet been fully understood or appreciated. They provide researchers an opportunity to focus the respondents' talk on a particular topic of interest. They can test their preliminary understanding on the topic of research and look for an opportunity for new ways of seeing and understanding to develop.

Non-structured interviews can be can be instrumental in the development of more structured interview guides or surveys.

- The interviewer and respondents involve in a formal interview as they follow a scheduled time for the interview.
- The interviewer has a clear plan, focus and goal of the interview in his or her mind.
- There is not a structured interview guide.
- Questions are open-ended with little control over informants' responses.
- As questions can be adapted and changed following respondents' answers, non-structured interviews are more flexible.
- Non-structured interviews generate qualitative data.

Limitations of Non-standardised Interviews

- The non-structured interviews can be long and thus time-consuming.
- The analysis of qualitative data can be an arduous exercise.

When to Use Non-standardised Interviews

You should use non-structured interviewing only when you have acquired sufficient understanding of a setting and your topic of interest.

Types of Non-structured Interviews

One-to-one Interviews

One-to-one interview is a common qualitative research tool. One-to-one interviews are exactly what they sound like—in-depth interviews conducted one-to-one between an individual respondent and a professional qualitative researcher. One-to-one interviews exist in the form of case studies, job interviews, book interviews or radio and television interviews. Interviewer can also use phone, e-mail or in person. The latter is most frequent method used in one-to-one interviews.

One-to-one interview is a survey method that is utilised when a specific target population is involved. The purpose of conducting one-to-one interview is to explore the responses of the people to collect more and deeper information. One-to-one interviews are used to probe the answers of the respondents and, at the same time, to observe the behaviour of the respondent.

Basically, there are two-types of one-to-one interview surveys according to how the interviewer approaches the respondents: 'intercept' and 'door-to-door interviews'. In an intercept approach, the interviewer usually conducts a short but concise survey by means of getting the sample from public places such as malls, theatres, food courts or tourist spots. On the other hand, in a door-to-door interview survey, the interviewer (or moderator) visits the respondent at his or her residence and conducts the interview either on the spot or at a scheduled date.

It is important to mention here that one-to-one interviews look similar to focus groups; however, according to Mcintyre (2012), there are a few key differences worth noting.

- One-to-one interviews tend to be more structured than focus groups and therefore require more active intervention from the moderator.
- One-to-one interviews tend to last between 20 and 40 minutes each on average—shorter than focus groups, which generally run between 90 minutes and 2 hours in duration.
- These interviews get to the core of the matter quickly by building a personal rapport between the moderator and the respondent and allowing each respondent to share their personal experiences, views and opinions in greater depth and detail than can generally be achieved in a focus group setting.
- The interview guide in the one-to-one interview is equivalent to the focus group discussion guide, and, similar to a discussion guide it should be developed by the moderator in collaboration with the client and with the objectives of the research study in mind.
- Using one-to-one interview does not mean that good process and methodical rigour are not required. Just as with focus groups, conducting a series of in-depth interviews calls for a thorough project planning, appropriate recruiting, the right incentives for participants, a detailed discussion guide and disciplined analysis and reporting.

Because only one respondent is involved at a time, the group discussion that occurs during focus groups is not realised with the one-to-one interview methodology. It is also important to note that although there are differences between one-to-one interviews and focus groups, the qualities of a good moderator are the same.

A good rule of thumb is to plan to conduct between 12 and 15 one-to-one interviews per target segment of interest. Experience shows that after this number of interviews, common themes and trends should begin to emerge.

Further, as an interviewer, your goal is to gain useful information from your interviewee through your questioning. For example, if the interview is based on the interviewee's personal life, then ask personal questions. Or if conducting a job interview, ask questions relating to the candidate's job qualifications and professional behaviour.

When to Use One-to-one Interviews

One-to-one interviews are likely the right choice for your research study if the subject matter is sensitive or very detailed, or if the research is required to generate insights based on the types of personal viewpoints, stories and experiences that are unlikely to be shared in a group discussion (confidential information). You can use one-to-one interview method for:

- Sensitive topics: When the subject matter is too sensitive or too personal or highly embarrassing in nature.
- Potential for bias: When participants' opinions could easily be coloured by the facial expressions or body language of others in the group.
- Busy audiences: When respondents are extremely busy such as time-pressed executives.
- Confidential information: When the objective is to ensure the strictest confidentiality and respondents' anonymity.
- Detailed understanding: When the objective to obtain a detailed understanding of complicated behaviours (motivations, beliefs, attitudes and feelings) is required.

This method is highly useful when your research objective is to acquire factual information, consumer evaluations, attitudes, preferences and other information coming out during the conversation with the respondent. Thus, one-to-one interview method ensures the quality of the obtained data and increases the response rate. You can also use this method in national or certain region population inquiries, consumer and customer or reader surveys.

One-to-one surveys have several key strengths. They are well structured, flexible and adaptable. They are based on personal interaction and can be controlled within the survey environment. Physical stimuli can be used and respondent can be observed. On the other hand, they are not free from some disadvantages. These disadvantages include interviewer bias, high cost per respondent, geographical limitations and time pressure on respondents.

Telephone Interviews

Researchers have been using increasingly telephone as a means of collecting data in several areas, but there are only few examples of studies where telephone interviews have been used to collect qualitative data. In particular, the field of educational is no exception to this.

With telephone interviews, a researcher gathers information rapidly. Telephone interviews are extensively used in major public opinion polls. They allow for some personal contact between the interviewer and the respondent and also permit the interviewer to ask follow-up questions.

A salient feature and main attraction of telephone interviewing is that the researcher can collect data from geographically scattered samples more cheaply and quickly, and avoid the well-known limitations of postal surveys. Another advantage is that interviewing from a central telephone unit lends itself to careful supervision and control. The researcher can approach and contact anyone on the planet who has a telephone.

But they also have some major disadvantages. With the advent of mobile phones, many people now use mobiles phones that are not listed in official telephone directories. Some do not have telephones. People often do not like the intrusion of a call to their homes. Finally, telephone interviews have to be relatively short or people will feel imposed upon.

Other problems are to obtain adequately representative samples of the general population and adequate response rates. In other words, the response rate is not as high as the face-to-face interview but considerably higher than the mailed questionnaire. The sample may be biased to the extent that people without phones are part of the population about whom the researcher wants to draw inferences. One can also raise doubts about the quality of the data, compared with face-to-face interviewing.

When to Use Telephone Interviews

You can use telephone interviews for conducting attitude surveys, customer satisfaction surveys or exploration of the potential for new products or services and respondents belonging to 'hard-to-reach' groups who are accessible via the telephone. For example, if the aim of your research study is to include members of the general public on topics of wide range such as opinion poll of parents on an increase in primary school tuition fees, head teachers on a specific school reform, recipients of grants for disabled children and so on. Use of telephone interviews can be highly useful and less costly method to collect information. Likewise, you can use this method when the respondents of your research study to be interviewed are widely geographically distributed.

Internet (Online) Interviews

Computers and the Internet have completely changed the styles and ways people communicate, access and interact with the information and take up research. Web technologies are expanding the process of research by enabling researchers to reach a larger and more geographically diverse sample. In the past, researchers used to collect qualitative data through interviews, observations and document analysis. Now all of these they can do online. Qualitative researchers can now collect information from participants from all around the world without having to travel.

The online interview comes in two basic forms: a 'screening interview' or a 'formal interview'. In any case, the interview is usually a behavioural interview, based on job criteria and standard questions. As such, the setting for your interview should be comfortable and well presented on camera. It is important to remember that for those of you who are working internationally, this is a common form of media for job interviews. Online interviews are usually video interviews. Thus, it is advisable to make a point of familiarising yourself with the online media and become fluent in its use.

The online interview is very similar to a face-to-face job interview. Its preparation and organisation both should be done in the same way as any face-to-face job interview:

- Essential job criteria: Revise each necessary element in the advertisement.
- Core skills: Treat each skill seriously and in depth.
- Communications: Focus on key communication skills, develop some standard interview questions and rehearse your answers.
- Interview content: Speak briefly and concisely and do not repeat and overstate information.
- Time management at interviews: Warn the interviewer in advance if you consider that your answer would be complex. This will save time.

When to Use Online Interviews

Online interviews 'are used for primary Internet-mediated research (IMR), that is, they are used to gather original data via the Internet with the intention of subjecting them to analysis to provide new evidence in relation to a specific research question' (Hewson 2010).

You can use online interviews when the objective of your research study is to explore:

- Patterns of technology use, modes of participation in online communities or human–computer interaction
- Behaviours or phenomena unrelated to the Internet that occur offline

Focus Group Interviews

Focus group interview is rapidly becoming one of the major research tools to understand people's thoughts and feelings. Through them, the researcher gets insights into how people think. They provide a deeper understanding of the phenomena being studied. Researchers use them for generating information on collective views and the meanings that lie behind those views. They are also useful in generating a rich understanding of participants' experiences and beliefs (Stewart and Shamdasani 1990).

There are various forms of focus group interviews. The most widely used are those where there is one moderator who acts as the observer also. Another type of interview can be where two moderators intentionally raise two different viewpoints to gather information about the contrasting views of the group members. Researchers can study two focus groups at the same time, where one group observes and reacts to the discussion of the other group. In another scenario, a member of the group can perform temporarily the role of a moderator. The researcher here acts only as the observer and simply takes notes of the behaviour of the group members. The selection of a given type of focus group and its method depend on the field of research and the problem being addressed. There are often cases when focus groups are used as an individual method of research or in combination with other methods including surveys, observations and interviews (Kitzinger 1994).

An important feature of group discussion is that corrections by the group concerning views that are not correct, not socially shared or extreme are available as means for validating statements and views. As focus groups involve interactions between the participants, it is important to ask the right questions.

How to Organise Focus Group Interviews?

A focus group interview is an interview with a small group of people on a specific topic. Groups typically consist of 6–8 people who participate in the interview for 90–120 minutes. Group discussions are used for various reasons. It is usually conducted with a trained moderator to talk about a product, service or organisation. The role of the interviewer is to make sure the group interact with each other and do not drift off topic. The meeting is conducted in a pleasant place, and refreshments are served to create a relaxed environment.

The selection of participants in the focus group is critical. A large sample size will give you a clearer picture of your participants than a small one. While there is no set method for deciding how large or small your sample should be, it will depend on what type of primary research you are conducting and the time and budget you can allocate to your research.

In the process of selection of participants for your research, you must be assured that all participants are able to answer your questions. You must see that they are familiar with the topic you are discussing, as targeting the wrong population can result in inconclusive or misleading data. In principle, the selection of your focus group participants should be based on criteria relevant to your customer profile.

Prior conducting your research with your participants, it is important to ensure that they agree to participate and outline any potential issues, such as confidentiality. You should explain to them how you will use their information and give them the opportunity to opt out at any time if they want.

In principle, focus groups are based on a set of questions or discussion points. Sometimes they may include visual concepts and even product samples for trial. Focus groups are most effective for qualitative research, as they let you explore people's opinions and attitudes (e.g., how they feel about a given education policy reform and what improvements they would like to see in the education system). Thus, a researcher must be highly skilled to conduct a group interview. For example, the interviewer should possess certain skills such as the ability to establish rapport and knowing when to probe.

As regards the moderator of the focus group, he or she needs objectivity, knowledge of the subject and the education system as well as some understanding of group and individual participant's behaviour. The moderator starts with a broad question before moving to more specific issues. He or she encourages an open and easy discussion to have a sense of true feelings and thoughts.

The nature of your research question and the count of different population subgroups required for the study will determine the number of group interviews you should conduct. Also, it is more appropriate to work with strangers instead of groups of friends or people who know each other. This is simply because the level of things taken for granted, which remain implicit tends to be higher in groups, which share an acquaintance with each other. Further, you should start with groups as heterogeneous as much as possible and then run a second set of groups that are more homogeneous.

When preparing for focus groups and interviews, you should prepare an outline/list of questions and key points so that you can have a productive conversation with your participants. In order to get the most out of a focus group or interview:

- Ask open-ended questions—those questions that cannot be answered with 'yes', 'no' or another single word. For example, instead of asking 'are you happy with our products?', you better ask 'which of our products are you happy with and why?'
- Repeat participants' responses to make sure you understand what they are saying.
- Ask follow-up questions in case you need more information about a participant's answer.

For the conduct of focus group interviews, you should plan where and when you will conduct them and think about the responsible person who will conduct them. If you are not familiar with the process, you may decide to hire the services of a professional expert for help. It is important that you record the answers from the interview and conversations in the focus group. You can use either a sound recorder or take written notes. The focus group should last 1–3 hours and as a nice gesture, you should pay participants to attend.

Strengths of Focus Group Interviews

- Group interviews generate qualitative data from open questions.
- The respondent talks freely and in some depth using his or her own words. This helps the researcher develop a real sense of a person's understanding of a situation.
- They increase validity because some participants may feel more comfortable being with others as they are used to talking in groups in real life, that is, the environment it is very natural.

Limitations of Focus Group Interviews

- The researcher cannot guarantee confidentiality and respect privacy.
- There is less reliability of group interviews as they use open questions.
- Group interviews may sometimes lack validity as participants may lie to impress the other group members. They may conform to peer pressure and give false answers.

When to Use Focus Group Interviews

In fact, focus group interviews have been widely used in the private sector, particularly in market research. In educational research, you can use them when:

- Limited resources prevent more than a small number of interviews being undertaken.
- It is possible to identify a number of individuals who share a common factor and it is desirable to collect the views of several people within that population subgroup.
- Group interaction among participants has the potential for greater insights to be developed.

Tests

As data-gathering devices, 'tests' are among the most frequently used tools of educational research, for they provide the data for most experimental and descriptive studies in education. The instruments have been designed to describe and measure sample of aspects of human behaviour. These instruments assess a variety of human abilities, skills, potential accomplishment and behaviour tendencies. They possess different degrees of validity reliability and applicability (Marwat 2010).

Types of Tests

The following types of tests are designed for different purposes.

Achievement Tests

Achievement tests attempt to measure what an individual has learned from his or her present level of performance. The achievement tests that most of us are familiar with are the standard examinations taken by every student in school. With these tests, students are regularly expected to demonstrate their learning and proficiency in a variety of subjects. These test determine individual or group status in academic learning. Achievement test scores are used in placing, advancing or retaining students at particular grade levels (Wise GEEK). In other words, achievement tests are needed in order to pass a class or continue on to the next grade level.

> An achievement test is an exam designed to assess how much knowledge a person has in a certain area or set of areas.

Achievement tests are used to ascertain what sort of information the test taker already knows at a specific point in time about a certain topic. They are not used to determine what you are capable of; they are designed to evaluate what you know and your level of skill at the given moment. This makes achievement test somewhat different from aptitude or abilities tests, both of which are designed to gauge how much potential a person has for later learning. Administrators looking only at achievement can have a good idea of where the test taker is at the present moment when it comes to knowledge of a specific topic or subject area. For this reason, most tests are designed to be straightforward and they tend to present material in a clear and unambiguous way.

Purposes of Achievement Tests

Achievement tests are frequently used in educational and training settings. In schools, for example, achievement tests are frequently used for admissions and the administrative roles of its staff and to determine the level of education for which students might be prepared. Students take such a test to assess if they are ready to enter into a particular grade level or if they are ready to pass particular subject or grade level and move on to the next. They are also used to determine if students have met specific learning goals. As there are certain educational expectations at each grade level, testing is used to determine if schools, teachers and students are meeting those standards.

Schools use these tests to both place students into appropriate grade levels and skill groupings and to assess teacher efficiency. Tests are not used to draw a conclusion about the general intelligence of the child being tested, but rather to ensure that each student is placed in a classroom where the student will have the best opportunity to learn. This enables the student to move on to more advanced material in later years.

In some cases, schools may administer achievement tests to teachers as a way of determining the best candidates for promotion or advancement. This sort of exam is usually given by a senior educational administrator to help identify the top contenders (teachers) for certain administrative positions.

Private schools and elite institutions may also use this sort of test as a way to measure the basic knowledge of applicants. Results so obtained from these tests enable school administrators to measure exactly where a prospective student is in terms of actual learning. Essays, interviews and past school records can give a rough idea of how a student will perform, but the test can provide greater insight into what he or she already knows. This can help determine whether the student would be a good fit for the academic environment.

Tests can also be used as a benchmark for either 'gifted' or 'remedial' programmes, both of which offer extra support to students who need it. For example, a student whose performance is relatively poor in basic mathematics on an achievement test is likely to be placed in a remedial learning class. This special assistance provides the student with the opportunity to master the basics before attempting to learn more advanced mathematical concepts like algebra or geometry. At a later date, the student may be examined again; should the results indicate that the student is sufficiently prepared to move on to something more complicated, he or she can be reassigned to a more challenging course of study.

Limitations of Achievement Tests

Though they are widely used in schools, achievement tests are not without their limitations. Many opponents argue that using the tests encourages teachers and educators to skew their curricula to encourage higher scores rather than focusing on individual student learning. This can be particularly troublesome in schools that offer financial incentives or other bonuses for high scores or penalty for low ones.

Some other opponents also argue that standardised achievement tests are skewed to favour certain subsets of students over others. The way questions are framed, the vocabulary employed and even the basic baselines in use make it easier for students of certain socio-economic, racial or ethnic backgrounds to do well, these opponents argue. In response, most school boards try to regularly examine their questions to ensure that they will adequately measure all students.

Aptitude Tests

Aptitude tests seek to assess the level of achievement that an individual is capable of doing or predict what he or she is able to learn or do given the right education and instruction and capable of attaining in some particular academic or vocational field (Career Explorer).

The main purpose of aptitude testing is to evaluate and watch how an individual reacts to situations and accomplishes tasks in his or her daily life. These tests contain several questions on number of individual's traits such as hobbies, talents, interests and overall personality. School administrators widely use aptitude tests to offer guidance to students taking the test regarding his professional or academic path.

There are numerous reasons for using aptitude test. These tests can be administered and monitored by a variety of sources. Employers use majority of these tests. However, certain government bodies utilise such tests as well, depending on the position an individual is seeking.

Purposes of Aptitude Tests

In an educational setting, aptitude tests are designed for the scholastic assessment test (SAT) and the college testing exam (CTE). These aptitude tests are created to monitor the overall academic competency of an individual. They are also used at various times during a child's life in order to determine if the child is in the appropriate grade in school or if adjustments need to be made. Educators use them to measure a student's performance and comprehension in a variety of subjects.

These tests serve very useful purpose in business where employers can determine the suitability of a candidate for a position or promotion. A career aptitude test is typically given at some point during the interview process; however, an employer may also choose to utilise a career assessment test when determining which staff member is competent and best suited for an upcoming promotion or to decide whether a person should be given more responsibility in his current position. For instance, the career aptitude tests are designed to measure an individual's skills in a particular field. It is quite difficult to cheat in an aptitude test. Therefore, these tests are considered reliable by most interviewers.

Design of Aptitude Tests

Aptitude tests are typically designed by educators or outside sources such as the Department of Education. Basically, educators do not need formal guideline for carrying out these tests. However, they follow the same format and structure regardless of the source of the test. Educators often use computers to help them achieve the best results when creating aptitude and assessment tests.

While no test is an infallible gauge of one's ability, standardised aptitude tests provide students, school admission boards and potential employers skills and ability to pinpoint areas where an individual is weak and strong, as well as where they rank among their peers. It is for this reason these tests are a beneficial tool in the business and academic worlds.

It is important to mention here that achievement tests differ in important ways from aptitude tests. An aptitude test is designed to assess our potential for success in a certain area. For example, a student might take an aptitude test to decide which types of career they might be best suited for. However, an achievement test is designed to determine what a student already knows about a specific subject.

Personality Tests

Personality tests are designed to study the non-intellectual aspect of human behaviour. They are used to accurately and consistently measure personality. Personality tests are primarily self-report instruments. The individual checks responses to certain questions or statements. These instruments yield scores, which are assumed or have been shown to measure certain personality traits or tendencies (Hewson 2010).

Personality tests are usually used:

- For assessing theories
- To look at changes in personality
- To evaluate the effectiveness of treatment
- For screening candidates for jobs

Personality tests seek to assess an individual's motivations in particular fields. Personality assessments, when validated properly, have been shown to adequately predict employment success. However, limiting assessments to an individual's personality ignores the impact mental aptitudes have on employment success (Career Explorer).

Tests are also classified on the basis of norm and criterion as shown further.

Norm-referenced Test

A norm-reference test is a test which compares the individual's performance with those other persons taking the same test. They serve the following purposes:

- Measure individual differences
- Classify and grade learners in various categories
- Determine the meaning of any particular score by comparing it to other scores achieved by students taking the test
- Select a candidate
- Make comparative decision regarding individuals
- Maximise the variability of test scores

Criterion-referenced Test

A criterion-referenced test evaluates an individual's in a given situation with respect to specific characteristics expected in the performance. These tests measure the effectiveness of a programme or instruction. The main features are as follows:

- They provide specific information in individual level of performance with respect to objectives.
- Scores of individuals can be interpreted individually.

- The purpose is not to classify and rank learners, but to ensure development.
- The results can be used to evaluate student performance relative to specific performance level anticipated.
- The examiner or test constructor is not concerned with developing a test to maximise the variability of test scores.

Qualities of a Good Test

The qualities of a good test are:

- Validity: In order for the test to be valid, the test should measure what it claims to measure. If a test is reliable, it does not mean that is it is also valid.
- Reliability: It measures accurately and consistently, from one time to another. For the test to be reliable, it must be relatively objective. How can a researcher obtain consistency among the scores the students earn from one time to the next, if the examiner is biased are inconsistent?
- Objectivity: It yields a correct and clear score value for each performance. A test can only be objective, if—using the same scoring key—whoever scores the test arrives at the same score—assuming no clerical errors.

Summary

Data collection is an important aspect of any type of research study. Inaccurate data collection can impact the results of a study and ultimately lead to invalid results. The data collection strategy, the type of variable, the accuracy required, the collection point and the skill of the researcher influence the choice of a given method of data collection.

The chapter explained three key data collection methods, namely, observations, interviews and tests widely used in educational research and social sciences. It described how the participant observation method involves human or mechanical observations of what children actually do or what events take place during the teaching–learning process and how researchers can use tests to assess the impact of a novel educational intervention on an individual learning process. Finally, the chapter explained interviewing as a data collection method for evaluation. The chapter comprised a basic overview of these key methods: when to use, how to plan and conduct, and their advantages and disadvantages.

Chapter 10 presents a similar analysis for another three key techniques of data collection. The chapter looks at how to handle, analyse and present data using these techniques to a variety of audience.

Self-test Exercises

Exercise 9.1:

9.1.1. Observations are designed to gather information on which to base decisions, make recommendations, develop curriculum and plan activities and learning strategies. The purposes of observation include:

☐ To determine the cognitive, linguistic, social, emotional and physical development of children.
☐ To identify children's interests and learning styles.

- ☐ To provide information to parents.
- ☐ To determine the cognitive, linguistic, social, emotional and physical development of children.
- ☐ All of the above.

9.1.2. Indirect observation refers to:

- ☐ Situations where the observation work is contracted to an outside agency.
- ☐ Situations where mechanical, electronic or 'non-human' means of observation are used.
- ☐ The observation and record, not of the 'direct' or key events of study, but rather that of associated events.
- ☐ Inferential rather than direct observation.

Exercise 9.2:
9.2.1. What characterises unstructured interviews?

- ☐ The respondent (interviewee) controls the course of the interview.
- ☐ The respondents (interviewees) are limited in kind of answers they can give.
- ☐ They are the only type of interviews in which probing techniques are useful.
- ☐ They are most suited to learn about the respondents' (interviewees') view points.

9.2.1. Describe briefly about the theoretical framework you will use in developing your interview.

Exercise 9.3:
9.3.1. What kinds of questions will provide information on what students know and do not know, and where they need to go next?
9.3.2. In what ways focus group method offers the benefits of interaction and group dynamics? Explain briefly.

References

Archar, L., A. Halsall and S. Hollingworth. 2007. 'Class, Gender, (Hetero) Sexuality, and Schooling: Working-class Girls' Engagement with Education and Post-16 Aspirations'. *British Journal of Sociology of Education*, 28 (2): 165–180.

Career Explorer. http://www.careerexplorer.net/article/the-role-of-aptitude-tests-in-the-modern-world (accessed on 13 July 2017).

DeWalt, K.M. and B.R. DeWalt. 2011. *Participant Observation: A Guide for Fieldworkers*. Lanham, MD: Rowman and Littlefield.

Hewson, C. 2010. 'Internet-mediated Research and its Potential Role in Facilitating Mixed Methods Research'. In *Handbook of Emergent Methods*, edited by S.N. Hesse-Biber and P. Leavy. New York, NY: Guilford. Available at http://assets.cambridge.org/97805218/70016/excerpt/9780521870016_excerpt.pdf (accessed on 25 May 2017).

Kitzinger, J. 1994. 'The Methodology of Focus Groups: The Importance of Interaction Between Research Participants'. *Sociology of Health & Illness* 16 (1): 103–121.

Kvale, S. 1996. *Interviews: An Introduction to Qualitative Research Interviews*. Thousand Oaks, CA: SAGE Publications.

Lofland, John. 1971. *Analysing Social Settings: A Guide to Qualitative Observation and Analysis*. Belmont, CA: Wadsworth Publishing Company, Inc.

Marshall, C. and G.B. Rossman. 2011. *Primary Data Collection Methods Designing Qualitative Research*. Los Angeles, CA: SAGE Publications.

Marwat, A. 2010. *Methods of Data Collection—Education Awareness and Research*. Available at http://research-education-edu.blogspot.in/2010/03/methods-of-data-collection.html (accessed on 17 November 2017).
Mcintyre, C.W. 2012. *One-on-one Interviews: What are They and Why Use Them*. Available at http://strategicinitiatives.ca/blog/uncategorized/one-on-one-interviews-what-are-they-and-why-use-them/ (accessed on 24 May 2017).
Patten, M. 2007. *Understanding Research Methods: An Overview of the Essentials*, 6th ed. Glendale, CA: Pyrczak Publishing.
Patton, M.Q. 1990. *Qualitative Evaluation and Research Methods*, 2nd ed. Newbury Park, CA: SAGE Publications.
Phellas, C.N., A. Bloch and C. Seale. 2011. 'Structured Methods: Interviews, Questionnaires and Observation'. In *Researching Society and Culture* edited by C. Seale, 3ed. London: SAGE, 181–202.
Roulston, K. 2010. *Reflective Interviewing: A Guide to Theory and Practice*. Los Angeles, CA: SAGE Publications.
Seidman, I. 2013. *Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences*. New York, NY: Teachers College.
Silverman, D. 2000. *Doing Qualitative Research*. London: SAGE Publications.
Stewart, D.W. and P.M. Shamdasani. 1990. *Focus groups: Theory and Practice*. London: SAGE Publications.
Williams, T.M. 1986. *The Impact of TV Violence on Children and Adolescents*. Available at http://www.thesuccessfulparent.com/categories/children-and-media/item/the-impact-of-tv-violence-on-children-and-adolescents#.WopqAa6WbIU (accessed on 19 February 2018).
Wise GEEK. *What is an Achievement Test?* Available at http://www.wisegeek.com/what-is-an-achievement-test.htm (accessed on 17 November 2017).

Appendix 9.1. Example of an Interview Schedule

An Interview Schedule
(Undergraduate Faculty Enhancement Programme)
Name of the Interviewer ____________________________
Date ____
Name of Interviewee ________________________________

1 Opening

1.1 Establish rapport
[Shake hand]

My name is _______________ and as your mathematics teacher, I thought it would be a good idea to interview you, so that I can better inform the rest of the class about you.

1.2 Purpose

I would like to ask you some questions about your background, your education, some experiences you have had and some of your hobbies and interests in order to learn more about you and share this information with our class.

1.3 Motivation

I intend to use this information to help the class become more comfortable speaking to and with you by knowing you better.

1.4 Timeline

The interview should take about 10–15 minutes. Are you ready to respond to some questions at this time?

(Transition: Let me begin by asking you some questions about where you live and about your family.)

2 Body of the interview

Topic 1 General demographic information

- How long have you lived in Shanghai?
- Are you originally from Shanghai?
- I noticed from faculty/university records and from your bio-sheet that you previously lived in a village.
- What was it like living there? Why?
- Do you have a small or large family?
- How many siblings do you have?
- Please describe your relationship with your family.
- What type of activities do you do with your family?

(Transition to the next topic)

Topic 2 Education

- Why did you choose East China Normal University (ECNU) for your studies?
- Why did you choose to major in education?
- How many classes have you taken in education?
- Would you recommend to other student to major in education? If so, why?
- Why did you change your science major to education?
- Do you plan to pursue a career in the field of education? If so, where?
- What area of services would you like to specialise? Why?
- What type of people would you like to work with? Why?

(Transition to the next topic)

Topic 3 Experiences

- You stated in your bio-sheet that you have had the opportunity to do some public speaking. Where did you speak?
- What was/were the occasion(s) that you presented?
- How do you think that your previous public speaking experience(s) will help you in this class?
- On your bio-sheet, you stated three of your goals for the class this semester. Please elaborate on these.
- You stated on your bio-sheet that you could competently inform our class about the types of events involved in an inter-university debate competition and what it takes to win the competition.
- What topic will you speak on for your informative speech?
- You listed on your bio-sheet that you do teaching job in your spare time. How do you think this course will help you become a professional teacher?
- You declared on your bio-sheet that you have travelled to India.
    - Where did you travel in India?
    - Describe your first impressions of Indian culture.
    - Do you speak Japanese?
    - Would you recommend India to a fellow student as a travel destination? Why or why not?
    - What was your most enlightening experience in your travels to India?

(Transition to the next topic)
Topic 4 Activities/Hobbies/Interests

- What campus and/or community activities are you involved in besides your postgraduate programme?
- You listed on your bio-sheet that you love to windsurf and you compete professionally.
  - What do you like best about windsurfing?
  - What was your most memorable windsurfing experience?
  - Do you ever give windsurfing lessons?
  - Where is your favourite place to windsurf? Why?

(Transition: Well, it has been a pleasure finding out more about you. Let me briefly summarise the information that I have recorded during our interview).

3 Closing the session

Summarise

- You are very involved in _______________. You plan to pursue a career in _______________. Your hobbies and interests are _______________.

Maintain Rapport

- I appreciate the time you took for this interview. Is there anything else you think would be helpful for me to know so that I can successfully introduce you to our class?

Action to be taken

- I should have all the information I need. Would it be alright to call you at home if I have any more questions? Thanks again. I look forward to introducing you to the rest of our class.

10
CHAPTER

Data Collection: Questionnaires and Surveys

Introduction

Observations, interviews, tests, questionnaires and surveys are the most widely used data collection techniques in social sciences. In Chapter 9, we noticed how the information gathered through observation, interviews and tests tools provides descriptions of characteristics of individuals, institutions or other phenomena under study, how useful it is for measuring the various variables pertaining to the study and their interrelationships for testing the hypothesis or for exploring the content areas set by the research objectives.

This chapter is about questionnaires and surveys as data collection methods for research. The chapter includes basic overviews of questionnaires and surveys, when to use them, how to plan and develop them and their advantages and disadvantages.

The chapter also provides a compilation of do's and don'ts for good questionnaire and survey designs that will, if followed, make your data analysis easier and faster than otherwise. The survey process is outlined to place questionnaire design in context. The chapter will help you learn the key principles of designing effective questionnaires and conducting surveys. Precisely, you will learn about:

- How to formulate meaningful questions and design surveys accordingly?
- How to formulate the three most important types of questions for asking about behaviour, attitudes and classifying respondents?
- Key terms used in questionnaire design
- The link between the interviewer, the respondent and the questionnaire

Questionnaires

This tag presents the process of designing standardised questionnaires that are commonly used to undertake research studies in school systems involving students, teachers, school headmasters or principals. Once read carefully, you would be able to design a quality survey questionnaire that is suitable for addressing the research issue at hand. You will know the key principles of effective questionnaire design. The tag will enable you learn how to:

- Decide on the target population and the vital information required
- Make a tentative listing of the questions and the variables and indicators that will address the research issues and hypotheses on which data are to be collected
- Phrase and refine structured, semi-structured, unstructured open-ended quantitative and qualitative questions and contingency questions (questions to study a future event or circumstance that is possible but cannot be predicted with certainty)
- Design probe questions and interviewer's instructions on the questionnaire
- Conduct a pilot test of the questionnaire, and revise and adjust its final design in accordance to the results of the pilot test
- Prepare a codebook for data entry

Defining a Questionnaire

A questionnaire is a data collecting technique in which a respondent provides answers to a series of questions. It comprises a set of questions for gathering information from individuals.

The process of developing a questionnaire that will help you collect the data you want takes effort and time. However, by following a step-by-step approach to questionnaire development, you can come up with an effective means to collect data that will answer your unique research question.

In common parlance, a questionnaire is a formal list of questions designed to gather responses from the respondents on a given topic.

> 'A questionnaire is a list of a research or survey questions asked to respondents, and designed to extract specific information, It serves four basic purposes: (i) collect the appropriate data, (ii) make data comparable and amenable to analysis, (iii) minimise bias in formulating and asking question, and (iv) make questions engaging and varies' (*Business Dictionary*).

> 'A formal standardized questionnaire is a survey instrument used to collect data from individuals about themselves, or about a social unit such as a household or a school. A standardised questionnaire exposes each respondent to the same questions and the same system of coding responses' (Siniscalco and Auriat 2005).

A well-designed questionnaire is an efficient data collection mechanism when the researcher knows exactly what is required and how to measure the variables of interest. A questionnaire translates the research objective into specific questions. The answers to these questions provide the relevant data for the hypotheses testing (Collis and Hussey 2003).

There are several different ways such as e-mail, telephones, using face-to-face interviews and handouts or electronic means by which you can administer questionnaires. And the types of information that can be collected by means of a questionnaire can be facts, activities, level of knowledge, opinions, expectations and aspirations, membership of various groups, and attitudes and perceptions.

Because the procedure of designing an effective questionnaire is complex, it is important that you understand the salient features of three key components of a questionnaire, namely: (a) variables, (b) indicators and (c) questions.

Variables

Once you have identified a research problem and formulated a researchable hypothesis, the next important step that you have to follow is to prepare a tentative list of variables and indicators that you intend to use for measuring your specific research questions and hypotheses of interest. You would not be able to do much in research unless you know how to talk about variables. All research projects are based around variables.

The research variables of any scientific inquiry or research process are factors that can take on different values that can be manipulated and measured. A scientific variable takes on different values and influences the outcome of experimental research. For instance, gender, colour and country are all good, perfect and acceptable variables, because they are all inherently changeable.

> 'A variable is a characteristic that can assume two or more properties. If a property can change either in quantity or quality, then it can be regarded as a variable' (Postlethwaite 2005).

Knowing the types of variables you are investigating in your research inquiry is necessary for all types of quantitative research design no matter you are using an experimental, quasi-experimental, relationship-based or descriptive research design. When you carry out your investigation, you may need to measure, manipulate and/or control the variables you are investigating. Here we present and illustrate the different types of variables you may wish to use or come across in your dissertation. We explain first the main groups of variables and then highlight the difference between dependent and independent variables so as to provide you one of the foundations required to handle your research study based on a quantitative research design.

How can you identify variables in a questionnaire? As a researcher, you must always determine which variable needs to be manipulated to generate quantifiable results. For an education research survey, Postlethwaite (2005) lists the following types of variables you should include in your survey questionnaire:

Discrete Variables

Discrete variables are also known as 'qualitative' or 'categorical variables'. Discrete variables are those that take numbers or specific values. Its value is obtained by counting. For example, the number of students present in a class is a discrete variable. The class cannot have 10.51 students; it will have either 10 students or 11 students.

Discrete variables are further classified as being 'nominal', 'dichotomous' or 'ordinal' variables. Each of these types of categorical variable has what are known as categories or levels. These categories or levels are the descriptions that we give to a variable. The levels help us to explain how variables should be measured, manipulated and/or controlled.

Suppose we want to know the career choices of university students. We could ask university students a number of closed questions related to their career choices.

Example

Question 1: What is your planned occupation?
Question 2: What is the most important factor influencing your career choice?
Question 1: highlights the use of categories and Question 2 levels in each category.
Question 1: What is your planned occupation?
Question 2: What is the most important factor influencing your career choice?

On a scale of 1 to 5, how important are the following factors in influencing your career choice (1 = least important; 5 = most important)?

Other examples of discrete variables could be: number of teachers absent, number of heads when flipping three coins, students' grade level and so on.

What is important for you to note about the categories in Question 1 and the levels in Question 2 is that all of these will be formulated by you. Ideally, you will include these categories or levels based on some primary or secondary research. Finally, you will choose which categories or levels to include and how many categories or levels there should be.

Each of these types of categorical variables (i.e., nominal, dichotomous and ordinal) are explained further with associated examples:

Variables with Categories

Categories	
Architect	
Attorney	
Biochemist	
Engineer	
Dentist	
Doctor	✓
Entrepreneur	
Social Worker	
Teacher	
Others	

Variables with Levels

Levels	*Scale*
Career prospects	5
Nature of the work	2
Physical working conditions	3
Salary and benefits	4
Others	

Nominal Variables

Where more than two categories are used for a variable, the variable becomes a nominal variable. We use nominal scales for labelling variables, without any quantitative value. Thus, 'nominal scales' could simply be called 'labels'. It is important to note that all of these scales are mutually exclusive (no overlap) and none of them have any numerical significance. A good way to remember all of this is that 'nominal' is like a 'name'. Consider the following examples:

Examples

Question 1

What is your gender?
● M-Male
○ F-Female

Question 2

What type of property you are interested in?
○ 1. House
● 2. Apartment
○ 3. Bungalow

Question 3

What is your hair colour?
○ 1. Blond
○ 2. Brown
○ 3. Red
● 4. Black
○ 5. Other……..

Question 1:	
Nominal variable:	Gender
Category:	Male, Female
Question 2:	
Nominal variable:	Type of property
Category:	House, Apartment Bungalow
Question 3:	
Nominal variable:	Hair colour
Category:	Blond, Brown, Red, Black, Other

Thus, when we talk about nominal variables not having an intrinsic order, we mean that they can only have categories (e.g., black, blond, brown and red hair); not levels (e.g., a Likert scale from 1 to 5).

Dichotomous Variables

Dichotomous variables are nominal variables that have 'just two categories'. They have a number of characteristics:

1. Dichotomous variables are designed to give you an either/or response.

Example

You are either male or female. You either like watching television (i.e., you answer YES) or you do not (i.e., you answer NO).

2. Dichotomous variables can either be fixed or designed.

Example

Some variables (e.g., your sex) can only be dichotomous (i.e., you can only be male or female). They are, therefore, fixed.

3. Dichotomous variables are designed by the researcher.

Example

Consider the question:

Do you like watching television?

For this question, you can determine that the respondent would select only 'YES' (i.e., he or she likes watching television) or 'NO' (i.e., he or she does not like watching television). However, there is another researcher who could provide the respondent with more than two categories to this question (e.g., most of the time, sometimes, hardly ever).

Researchers generally use non-parametric statistical methods for nominal scales. The most likely ones are:

- Mode
- Cross-tabulation—with chi-square

There are also highly sophisticated modelling techniques available for nominal data.

Ordinal Variables

Like nominal variables, ordinal variables have two or more categories. However, unlike nominal variables, ordinal variables can also be ordered or ranked (i.e., they have levels).

Example

Question:	Do you like the education policies of the Congress Party?
Answer:	Not very much (or they are OK, or Yes, a lot).
Ordinal variable:	Opinions towards Democratic Party's policies.
Level:	Not very much. They are OK, Yes, a lot.

So if you asked someone if they liked the policies of the Congress Party and you categorised them as: (a) Not very much, (b) They are OK or (c) Yes, a lot; you have an ordinal variable. Why? Because you have three categories, namely, not very much, they are ok and yes, a lot and you can rank them from the most positive (yes, a lot), to the middle response (they are ok), to the least positive (not very much). However, whilst you can rank the three categories, you cannot place a value to them. For example, you cannot say that the response, 'they are ok', is twice as positive as the response, 'not very much'. In other words, for example, is the difference between 'ok' and 'unhappy' the same as the difference between 'very happy' and 'happy?' We cannot say.

Other examples of ordinal variables are:

Question:	In what year did you start university?
Answer:	I started in 2006 (or 2007, 2008, 2009, 2010)
Ordinal variable:	Year of university entry
Level:	2006, 2007, 2008, 2009, 2010

Question:	Do you like watching television?
Answer:	Most of the time (or sometimes or hardly ever)
Ordinal variable:	Opinion about watching television
Level:	Most of the time, sometimes, hardly ever

Question:	To what extent do you agree or disagree with the following statement? Going to university is important to get a good job (Based on a 5-point Likert scale of 1 = strongly agree, 2 = agree, 3 = neither agree nor disagree [neutral], 4 = disagree, 5 = strongly disagree)
Answer:	2 = I agree (or 1, 3, 4 or 5 on the 5-point Likert scale)
Ordinal variable:	The importance of university to getting a good job
Level:	1 = strongly agree, 2 = agree, 3 = neither agree nor disagree, 4 = disagree, 5 = strongly disagree

When it comes to Likert scales, as highlighted in the previous example, there can be some disagreement over whether these should be considered ordinal variables or continuous variables.

Continuous or Quantitative Variables

Continuous variables, which are also known as quantitative variables, can be further classified as being either interval or ratio variables. Both interval and ratio variables have numerical properties. These numerical properties are the values by which continuous variables can be measured, manipulated and/or controlled. These are variables that take all values within a particular range.

Ordinal data would use non-parametric statistics. These would include:

- Median and mode
- Rank order correlation
- Non-parametric analysis of variance
- Modelling techniques can also be used with ordinal data

Interval Variables

Interval variables have a numerical value and can be measured along a continuum.

Example

Interval variable:	Temperature (measured in degrees Celsius or Fahrenheit)
Explanation:	The difference between 20 degrees C and 30 degrees C is the same as 30 degrees C to 40 degrees C.

However, there is a problem with interval scales: they don't have a 'true zero'. For example, there is no such thing as 'no temperature'. You cannot compute ratios without a true zero. With interval data, you can add and subtract, but cannot multiply or divide. Consider this: 10 degrees + 10 degrees = 20 degrees. No problem there. But 20 degrees is not twice as hot as 10 degrees, however, because there is no such thing as 'no temperature' when it comes to the Celsius scale. Thus, interval scales are great, but we cannot calculate ratios.

For interval scale data, we use parametric statistical techniques:

- Mean and standard deviation
- Correlation—*r*
- Regression
- Analysis of variance
- Factor analysis

In addition, there are a whole range of advanced multivariate and modelling techniques. You can use non-parametric techniques with interval and ratio data. But non-parametric techniques are less powerful than the parametric ones.

Ratio Variables

Ratio variables are interval variables that meet an additional condition: they have value '0' (zero) as a measurement value. A ratio variable is like an interval variable except that it also has a clear definition of '0'. When the variable equals '0', it means there is none of that variable. Variables such as height, mass/weight and so on are examples of ratio variables.

Temperature, expressed in Fahrenheit (F) or Celsius (C), is not a ratio variable because a temperature of '0 degree' on either F or C scales does not mean 'no heat'. However, temperature in Kelvin is a ratio variable, because '0 degree' Kelvin really does mean 'no heat'.

Examples

Ratio variable:	Temperature measured in Kelvin
Explanation:	0 Kelvin, often called absolute zero, indicates that there is no temperature whatsoever

A temperature of 10 Kelvin is four times the temperature of 2.5 Kelvin.

Ratio variable:	Distance
Explanation:	If two houses are joined together (e.g., terraced housing), the distance between the adjoining walls is 0 (i.e., there is no distance whatsoever). On the other hand, a distance of 10 metres between the houses would be twice the distance of a 5-metre gap between the houses (i.e., a distance of 10 metres is twice the distance of 5 metres).

You can use the same statistics as for interval data.

Dependent and Independent Variables

'Dependent variables' (aka 'outcome variables') are those variables that the researcher tries to explain (e.g., student achievement). As the name suggests, a dependent variable depends on independent variable(s).

'Independent variables' are variables that cause, or explain, a change in the dependent variable. An independent variable (sometimes called an 'experimental' or 'predictor variable') is a variable that is manipulated in an experiment so as to observe the effect this has on a dependent variable.

Example

Imagine that a teacher asks 100 students to complete a maths test. The teacher wants to know why some students perform better than others. Whilst the teacher does not know the answer to this, he or she thinks that it might be because of two reasons:

1. Some students spend more time revising their test and
2. Some students are naturally more intelligent than others

Therefore, the teacher decides to investigate the effect of revision time and intelligence on the test performance of the 100 students. In this example, the dependent and independent variables for the study are:

Dependent Variable:	Test Mark (measured from 0 to 100)
Independent Variables:	Revision time (measured in hours)
	Intelligence (measured using IQ score)

Control Variables

Control variables are used to test for a spurious (fake) relationship between dependent and independent variables. In other words, control variables test whether an observed relationship between dependent and independent variables may be explained by the presence of another variable.

For instance, we know that language teaching and learning is a very complex process. For studying this complex process, it is not possible to consider every variable in this type of study. Therefore, any variable that is not measured in a particular study must be held constant, neutralised/balanced or eliminated so as to control its biasing effect on other variables. Variables that are controlled in this way are called control variables.

Example

Suppose that we want to measure how much water flow increases when we open a faucet. It is important to be certain that the water pressure (the controlled variable) is held constant. That is because both the water pressure and the opening of a faucet have an impact on how much water flows. If we change both of them at the same time, we cannot be sure how much of the change in water flow is because of the faucet opening and how much because of the water pressure. In other words, a test such as this would not be a fair test. Most experimental research studies have more than one controlled variable. Some experts refer to controlled variables as 'constant variables'.

In summary, nominal scales are used when the aim of research is to label a series of values. Ordinal scales provide us good information about the order of choices such as in a customer satisfaction survey. We can determine the order of values plus the ability to quantify the difference between each one by interval scale. Finally, ratio scales give us the ultimate order, interval values and the ability to calculate ratios since a 'true zero' can be defined.

Indicators in an Education System

You noticed in Chapter 8 that educational indicators are used to monitor the education system. They provide insight on current issues and challenges, trends and forecast impending changes. We have also noticed that the overriding purpose of indicators is to characterise the nature of a system through its components—how they are related and how they change over time. We described also how this information could then be used to judge progress towards some goal or standard, against some past benchmark, or by comparison with data from some other institution or country.

You also noticed that an 'indicator' is an empirical, observable, measure of a concept. When an indicator consists or composed of a combination of variables involving only simple calculations (such as addition, subtraction, division, multiplication or a combination of these) it is called a 'simple indicator'. When more complex analytical methods, such as factor analysis or regression are used to develop an indicator, the result is referred to as a 'complex indicator'. Examples of simple indicators are: number of school library books per pupil or PTR. An example of a complex indicator is a factor score entitled 'emphasis on phonics' in the teaching of reading formed from three variables: learning letter sound relationships; word attack skills and assessment of phonic skills (Siniscalco and Auriat 2005).

Education indicators are statistics that tell us the important aspects of the education system, but not all statistics about education are indicators. You should remember that statistics could be considered as indicators only if we can use them as standard for comparison when we are judging other people or things, that is, indicators must tell us a great deal about the entire system by reporting the condition of a few particularly significant features of it.

For example, the number of girl students enrolled in schools is an important fact, but it hardly tells us how well the education system is functioning. However, data on the proportion of university students who have successfully completed MBBS can provide considerable insight into the health of the system and can be appropriately considered an indicator. Thus, a good education indicator system is expected to provide accurate and precise information to illuminate the condition of education and contribute to its improvement.

It has been mentioned in Chapter 8 that indicators are widely used in an education system to undertake diagnostic studies and for monitoring purposes. As such, these indicators cannot be used to: (a) set goals and priorities; (b) evaluate programmes and (c) develop a balance sheet of the system. Sheldon and Parke (1975) believe that with the use of educational indicators:

> We will be able to describe the state of the society [education system] and its dynamics and thus improve immensely our ability to state problems in a productive fashion, obtain clues as to promising lines of endeavour, and ask good questions. The fruit of these social indicator efforts will be more directly a contribution to policy-makers' cognition than to their decisions.

Computation, application and use of educational indicators in a research enquiry were discussed in details in Chapter 8.

Question Structure

Questionnaires are one of the key tools in collecting necessary information from the respondents of a survey. By making the right choices on the type of survey questions, you can exclusively extract data that are related to the purpose or goal of your survey. There are two important aspects of questionnaire design: (a) 'structure of the questions' and (b) the 'decisions on the types of response formats' for each question.

Before constructing questions, it is important that you must be knowledgeable about each type of question used in your survey research. These basically include three structures: (a) 'closed-ended', (b) 'open-ended' and (c) 'contingency questions'.

Closed-ended or Multiple-choice Questions

Closed-ended questions limit the answers of the respondents to response options provided in the questionnaire by the researcher, for example, multiple-choice questions, scale questions and so on.

In closed questions, the researcher asks the respondent to choose, among a possible set of answers, the response that most closely represents his or her viewpoint. The respondent is usually asked to tick or circle the preferred answer. Questions of this kind generally offer simple alternatives such as 'Yes' or 'No'. They may also require that the respondent chooses among several answer categories or that he or she uses a frequency scale, an importance scale or an agreement scale (Postlethwaite 2005).

Example

How often do your parents ask you about your homework?
(Please encircle one answer only)

Never ________________________________ 1
1 or 2 times a week______________________ 2
3 or 4 times a week______________________ 3
Nearly every day_________________________ 4

As these types of questions used to generate statistics in quantitative research follow a set format, and most responses can be entered easily into a computer for ease of analysis, greater numbers of questionnaires can be distributed for the study.

Some closed questions may contain a dichotomous response format, which means only two mutually exclusive responses are provided.

Example

What is your sex?
(Please tick one box only)

☐ Male
☐ Female

Some examples of close-ended questions are:

- Dichotomous or two-point questions (e.g., Yes or No, Unsatisfied or Satisfied)
- Multiple-choice questions (e.g., A, B, C or D)

Advantage and Limitations of Closed Questions

Advantages

- Easier and quicker for respondents to answer
- Answers of different respondents are easier to compare
- Answers are easier to code and statistically analyse

- Response choices can clarify question meaning for respondents
- Respondents are more likely to answer about sensitive topics
- Fewer irrelevant or confused answers to questions
- Less articulate or less literate respondents are not at a disadvantage
- Replication is easier
- Respondent is restricted to a finite (and therefore more manageable) set of responses
- They have response categories that are easy to code and interpret
- They permit the inclusion of more variables in a research study
- They are ideal for quantitative type of research

Limitations

- They can suggest ideas that the respondent would not otherwise have
- Respondents with no opinion or no knowledge can answer anyway
- Respondents can be frustrated because their desired answer is not a choice
- It is confusing if many response choices are offered
- Misinterpretation of a question can go unnoticed
- Distinctions between respondent answers may be blurred
- Clerical mistakes or marking the wrong response is possible
- They force respondents to give simplistic responses to complex issues
- They force respondents to make choices they would not make in the real world

Open-ended Questions

Open-ended questions do not have pre-defined options or categories. The study participants supply their own answers. For open-ended question, there are no boxes to tick but instead there is a blank space or section for the respondent to write in his or her answer. Closed-ended questionnaires might be used to find out 'how' many people use a service, whereas open-ended questionnaires might be used to find out 'what' people think about a service.

Example

What are your favourite TV programmes?
(Please specify their titles)

____

____

What do you like most about your school?

____

____

As there are no standard answers to these questions, data analysis is more complex. As they are opinions which are sought rather than numbers, fewer questionnaires need to be distributed.

Some examples of open-ended questions include:

- Completely unstructured questions: These questions openly ask the opinion or view of the respondent. Questions tend to be open-ended and express little control over informants' responses.

Example: Who are your favourite romantic English poets?

- Word association questions: In this type of questions, the respondent states the first word that pops in his or her mind once a series of words are presented. Word association questions are objective type questions and answers. This is actually an analogy test, (i.e., a comparison between two or more words or things, typically for the purpose of explanation or clarification), two items or words bear a particular relation with each other, third item bears the same relation with one of the answer choices. The respondent is supposed to find relation and association between words.

Example

Question Pick up the common property of the following words:
SOBER, GENTLE, POLITE

Select appropriate answer from the following:

- ☐ Nice
- ☐ Civil
- ☐ Affable
- ☐ Wise

The right answer is 'Civil'.

- Thematic apperception test: In this test, a picture is presented to the respondent to explain his or her own point of view. It is a projective psychological test. Proponents of this technique assert that a person's responses reveal underlying motives, concerns and the way they see the social world through the stories they make-up about ambiguous pictures of people. Historically, it has been among the most widely researched, taught and used of such tests.

Example: Explain and give your own point of view of the following picture:

Sentence, story or picture completion: Here the respondent continues with an incomplete sentence or story, or writes on empty conversation balloons in a picture.

Example

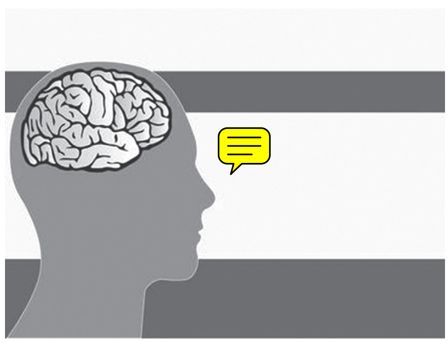
1. I love cats ***and*** I______ dogs.
   A. **Love**
   B. Play
   C. Hate
   D. Sleep

2. I love cats ***but*** I ______ dogs.
   A. Love
   B. Play
   C. **Hate**
   D. Sleep

Advantages and Limitations of Open-ended Questions

Advantages

Participants can respond to the questions exactly as how they would like to answer them and the researcher can investigate the meaning of the responses ideal for qualitative type of research.



Disadvantages

Time-consuming; responses are difficult to code and interpret

Combining Open- and Close-ended Questions

By combining open- and closed-ended questions, you can find out how many people use a service and what they think of the service in the same questionnaire. This type of questionnaire begins with a series of closed-ended questions, with boxes to tick or scales to rank and then finishes with a section of open-ended questions or more detailed response.

Scaled or Matrix Questions

'Scaled questionnaires' are basically used to measure attitudes and beliefs. The scaled questionnaires use a Likert scale in which respondents are given choices showing varying degrees of intensity. For example, educational researchers have developed scaled questionnaires to measure school culture using items such as the following:

Question: In this school, staff members are recognised when they do a task well.
Choose one: ☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

According to UNESCO-IIEP, scaled questionnaires have the same validity and reliability requirements as tests. For example, what is the evidence that a school culture scale is actually measuring school culture and not some other property or characteristic of the school, such as material wealth? How a scaled questionnaire used in a study also affects research validity? A scaled questionnaire developed to measure school culture might not have any relationship to leadership or student achievement, yet sometimes a researcher will make such unwarranted conclusions. The conclusions of a research study can be invalid despite the use of a valid data collection instrument if the conclusions extend beyond the limits of what was measured (Siniscalco and Auriat 2005).

Here is another example of how scaled questionnaires are developed:

A scaled questionnaire designed to measure school culture might ask teachers and administrators questions such as the following:

Question: I feel comfortable about discussing my concerns in this school.
Choose one: ☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

To develop a scaled questionnaire, you should ask a large sample of participants to respond to a large number of items you have judged to have content validity with regard to a particular concept. For example, you might verify with practitioners and other researchers that the items concern aspects of school culture. Next, you can, if necessary, reduce the number of questionnaire items through a statistical procedure called 'factor analysis', which results in a small number of factors (variables) that relate to school culture.

Contingency or Filter Questions

A filter question is a question you ask in order to determine if the person you are speaking to is a candidate for another question, that is, you ask the respondent one question to determine if he or she is qualified or experienced enough to answer a subsequent question.

For example, you want to know how many times the parents of a child have attended Pupil–Teacher Association (PTA) meetings. In order to ask this question, you may ask a filter question first: 'have you ever attended a PTA meeting?' Asking filter questions effectively avoids asking people questions that are not applicable to them. Take another example. Suppose you may want to ask one question if the respondent has ever smoked a cigarette and a different question if they have not. In this case, you would have to construct a filter question to determine whether they have ever smoked a cigarette.

In the following example, Question 1 is a closed-ended question. Question 2 is contingency or filter question which follows a closed-ended question.

Question 1: Have you ever smoked a cigarette?
Yes ☐ →
No ☐

Question 2: If Yes, how many times have you smoked cigarette?

Once	☐
2–5 times	☐
6–10 times	☐
More than 10 times	☐

You should remember that filter questions often get very complex. In situation such as this, you should formulate multiple filter questions in order to direct your respondents to the correct subsequent questions. Thus,

- You should always try to avoid having more than three levels (two jumps) for any question. Too many levels or jumps will confuse the respondent and may discourage him or her from continuing with the survey.
- For two levels, you should use graphic to jump (e.g., arrow and box). See example above. It shows how you can make effective use of an arrow and box to help direct the respondent to the correct subsequent question.
- If possible, jump to a new page. For instance, if you cannot fit the response to a filter question on a single page, then you instruct the respondent by saying something like 'If YES, please turn to page 6' rather than 'If YES, please go to Question 30' because the respondent will generally have an easier time finding a page than a specific question.

Developing a Questionnaire for Research

Developing your research questionnaire is a step-by-step systematic process to collect accurate facts and opinions and will be an enjoyable event for the respondent. The entire process involves three parts, namely:

- The designing the questionnaire
- Writing the questionnaire
- Distributing the questionnaire.

Designing the Questionnaire

We have seen above that a questionnaire is the key instrument in any research inquiry as it helps to extract data from respondents. It provides you guidelines and you, in turn, have to ask the questions in exactly the same way. Without these guidelines, questions would be asked in a haphazard way at your own discretion.

While designing a questionnaire for your study, you should consider the following five people:

- Client: who wants answers to his or her particular problem
- Researcher: who uncovers information and balances the needs of three groups of people—interviewer, respondents and client. Thus, he or she should ensure that (a) the interviewer can manage the questionnaire easily, (b) the questions are interesting for the respondent and (c) the questionnaire matches the client's needs
- Interviewer: who wants a questionnaire which is easy to follow and which can be completed in the time specified by the researcher
- Respondent: who feels that the questions are phrased in a simple and systematic way so that they can be answered truthfully. They should make the respondent to actually say what he or she thinks.
- Data processor: the questionnaire results in data which can be processed, analysed and interpreted efficiently and with minimum error.

Your questionnaire must comprise the following five sections:

- The respondent's identification data: This should include information on items such as their name, address, date of the interview, name of the interviewer. Each questionnaire would also have a unique number (questionnaire identity number) for purposes of entering the data into the computer.
- An introduction: This section should contain your (interviewer's) request for help. It normally provides the credentials of the research entity or organisation, the purpose of the study and any aspects of confidentiality.
- Instructions: You and the respondent both need to know how to go through the questionnaire such as which questions to skip and where to move to if certain answers are given.
- Information: This is the main body of the questionnaire and is made up of many questions and response codes.
- Classification data: These questions, sometimes at the front of the questionnaire, sometimes at the end, establish the important characteristics of the respondent, particularly related to their demographics characteristics such as sex, race, age, education and where the participant works or lives. The purpose of these questions is to describe subgroups of respondents. You should always limit the demographic questions to only those that are important for your analysis. For example, if you do not plan to compare the data by age, do not include age in the questionnaire.

There are seven steps in the design of a questionnaire:

Step 1: Identify the goal of your questionnaire
This is your starting point. It should contain the proposal and brief, and list all the objectives of the study. What kind of information do you want to gather with your questionnaire? What is your main objective? Is a questionnaire the most effective tool to gather this information?

- Start with the key research question(s). The key research question can be one question or several, but this should be the focal point of your questionnaire.
- Formulate one or several hypotheses that you want to test. The questions should help you testing these hypotheses systematically.

Step 2: Make a rough listing of the questions

Your aim at this stage should be as comprehensive as possible in the listing. You should not worry much about the phrasing of the questions. Phrasing comes next.

Depending on the information you wish to gather, there are several possible types of questions to include in your questionnaire, each with unique merits and demerits. The types of commonly used questions on a questionnaire are explained above.

Step 3: Refine the question phrasing

You must now develop questions close to the point where they make sense and will generate the right answers. The questions should be clear, concise and direct. This will guarantee that you get the best possible answers from your respondents. The following tips may help you in refining the question phrasing:

- Write questions that are succinct and simple. You should avoid complex statements or technical jargon, as it will only confuse your respondents and lead to incorrect responses.
- Ask only one question at a time. This will help avoid confusion.
- Exercise utmost caution while asking for 'private' or 'sensitive' information. This can be simple information related for example to age or weight, or something as complex as past sexual history.
- Determine if your questionnaire will include an answer such as 'I don't know' or 'Not applicable to me'. Questions such as these can give your respondents a chance of not answering certain questions; however, providing these options can also help you to know the missing data, which can be problematic during data analysis.
- Start with the most important questions at the beginning of your questionnaire. This can help you gather important data even if you sense that your respondents are distracting by the end of the questionnaire.

Step 4: Develop the response format

Every question needs a response. The responses could be a pre-coded list of answers or it could be open-ended to collect verbatim comments. Consideration of the responses is just as important as getting the questions right. In fact, considering the answers will help you get the questions right.

Step 5: Restrict the length of your questionnaire

You should always try to keep your questionnaire as short as possible. More people will be likely to answer a shorter questionnaire, so make sure that it is as concise as possible while still collecting the necessary information. If you can make a questionnaire that only requires five questions, do it.

- You should include only questions that are directly useful and related to your research objectives. You should always remember that a questionnaire is not an opportunity to collect all kinds of information about your respondents.
- Avoid asking redundant questions as it might frustrate those who are taking your questionnaire.

Step 6: Identify your target population

Before distributing the question, you should identify: (a) is there a certain group of people who you want to target with your questionnaire? If this is the case, it is best that you determine this before you begin to distribute your questionnaire.

- You should also decide whether you want that your questionnaire collect information from both men and women. Some of your studies only survey one sex.
- You should also consider all other aspects that would make a person a target for your questionnaire. Make sure you are very clear about this before you distribute your questionnaire.

Step 7: Ensure you can protect privacy
It is important that you make a plan to protect respondents' privacy before you begin recording your survey. This is one of the key ethical considerations for almost all research studies.

- You should design an anonymous questionnaire where you may not want to ask for names on your questionnaire. This is how you can prevent privacy. However, you should keep in mind that it is often possible to figure out a respondent's identity using other demographic information (such as age, physical features or postal address).
- For preventing privacy, you can think of de-identifying the identity of your respondents. It means, you assign to each questionnaire (and thus, each respondent) a unique number or word and only refer to this as an identifier.
- You should also keep in mind that you do not need to collect much demographic information to be able to identify someone. People may show caution about possible dangers or problems and thus may not like to provide this information. In this case, you should ask less demographic questions so as to get more respondents.
- Once you have completed your study, you must destroy all identifying information.

Writing the Questionnaire

While writing your questionnaire, you should keep in mind the following steps.

Step 1: Introduce yourself
You introduce yourself by explaining who you are and what your credentials are. You should clarify if you are the only one working for this study or as a part of a team. If you are collecting data for an academic institution or company, mention the name of the academic institution or company for whom you are collecting data. Following are some examples:

- My name is Varavara and I am the senior researcher fellow and designer of this questionnaire. I am part of the National Council of Educational Research and Training of the MOE, Thailand, where I am focusing in developing cognition in infants.
- I am Ms Manal, a third-year undergraduate student at Amman Open University, Jordan. This questionnaire is part of my final exam in statistics.
- My name is Natasha and I am a senior executive and marketing analyst in REDA Chemicals Ltd, Moscow. I have been working on questionnaire development to determine attitudes surrounding drug use in universities in South Asian countries for several years.

Step 2: Explain the purpose of the questionnaire
Many people will not answer a questionnaire without understanding what the goal of your questionnaire is. There is no need to write long explanation; instead, a few concise sentences will serve the purpose. Here are some examples:

- I am collecting data regarding the attitudes surrounding discipline in secondary schools. This information is being collected for my dissertation at the Sorbonne University, Paris, France.
- This questionnaire will ask you 15 questions about your reading, eating and exercise habits. I am attempting to make a correlation between healthy eating, frequency of exercise and incidence of obesity among graduate students.
- The purpose of this questionnaire is to ask you about your recent experiences with international air travel. There are three sections of questions that will ask you to recount your

recent trips and your feelings surrounding these trips, as well as your travel plans for the future. The purpose of the questionnaire is to gather information to understand how a person's feelings surrounding air travel impact their future plans.

Step 3: Reveal what will happen with the data you collect
When collecting data, you should be clear that it is for a class project, for a publication or will be used for educational research. Depending on what you intend to do with the data you collect from your questionnaire, there may be diverse requirements to which you need to pay particular attention to before distributing your survey questionnaire.

- A note of caution. If you are gathering information for a university or for publication, you must check with your Institutional Review Board (IRB) and seek its permission before administering the questionnaire. Most universities have a dedicated IRB staff and you can very easily trace this information from the university's website.
- This is necessary for reasons of transparency. It is important to be honest about what will happen with the data you collect.
- Include an informed consent form, if necessary. Note that you cannot guarantee confidentiality, but you will make all reasonable attempts to ensure that you protect their information.

Step 4: Estimate how long the questionnaire will take
Before someone sits down to take your questionnaire, it may be helpful for him or her to know whether the questionnaire will take him or her 10 minutes or 2 hours. You should inform the respondents at the very onset of your questionnaire. This is more likely to get you more complete questionnaires in the end.

- Time yourself taking the survey. Then consider that it will take some respondents longer than you and some respondents less time than you.
- Avoid a rigid time range. For example, it is better to say that this questionnaire will take between 15 and 30 minutes than to say it will take 15 minutes and have some respondents quit halfway through.
- Use this as a reason to keep your survey concise. You will feel much better asking people to take a 20 minutes survey than asking them to take a 3 hour-survey.

Step 5: Describe any incentives that may be involved
You should see your ability to offer an incentive to the respondents at the end of the questionnaire. Incentives can take the form of: monetary prizes, awards, gift certificates, candy and so on. However, there are both advantages and disadvantages of offering incentives.

- There is likelihood that you end up by attracting the wrong kind of respondent. You do not want to incorporate responses from people who rush through your questionnaire just to get the reward at the end. This is a potential danger of offering an incentive.
- Incentives can encourage people to respond to your survey who might not have responded without a reward. This is a situation in which incentives can help you reach your target number of respondents.
- Instead of directly paying respondents to take their surveys, you can offer 50 cents to the charity of their choice when a respondent fills out a survey. This may lessen the chances that a respondent will fill out a questionnaire out of pure self-interest.
- You should consider entering each respondent in to a drawing for a prize if they complete the questionnaire. You can offer a gift card to a restaurant or a ticket to a movie. This makes it less

tempting just to respond to your questionnaire for the incentive alone, but still offers the chance of a pleasant reward.

Step 6: Make sure your questionnaire looks professional
Because you want respondents to have confidence in you as a data collector, your questionnaire must have a professional look.

- You should always proofread the questionnaire and check for spelling, grammar and punctuation errors.
- The questionnaire should have very clear title. This will help the respondents to understand the focus of the survey as quickly as possible.
- You should express your gratitude by vote of thanks to your respondents for taking the time and effort to complete your survey.

Distributing the Questionnaire

Step 1: Do a pilot study
Ask a small group of people who are similar to your intended participants to take your questionnaire (they should not be included in any results stemming from this questionnaire), and be ready to revise it if necessary. You should include 5–10 people in the pilot testing of your questionnaire to get their feedback on your questionnaire by asking the following questions:

- Was the questionnaire easy to understand? Were there any questions that confused you?
- Was the questionnaire easy to access?
- Do you feel the questionnaire was worth your time?
- Were you comfortable answering the questions asked?
- Would you suggest any improvements in the questionnaire?

Step 2: Disseminate your questionnaire
You should determine the best way to disseminate the questionnaire. You may consider the following common ways to distribute the questionnaire:

- Using the mail: If you mail your survey, always make sure you include a self-addressed stamped envelope so that the respondent can easily mail their responses back. Your questionnaire must fit inside a standard business envelope.
- Conducting face-to-face interviews: This can be a good way to ensure that you are reaching your target demographic and can reduce missing information in your questionnaires.
- Using the telephone: While this can be a more time-effective way to collect your data, it can be difficult to get people to respond to telephone questionnaires.

Step 3: Include a deadline
Ask your respondents to have the questionnaire completed and returned to you by a certain date to ensure that you have enough time to analyse the results.

- Make your deadline reasonable. Giving respondents up to two weeks to answer should be more than sufficient. Anything longer and you risk your respondents forgetting about your questionnaire.
- Consider providing a reminder. A week before the deadline is a good time to provide a gentle reminder about returning the questionnaire. Include a replacement of the questionnaire in case your respondent has misplaced it.

Questionnaire Construction

While designing a questionnaire, you should keep the following points in mind (Postlethwaite 2005):

- Explain the purpose of the questionnaire to all respondents.
- Do not use jargon or specialist language (unless the recipients really prefer and understand it).
- Phrase each question so that only one meaning is possible.
- Avoid vague, descriptive words, such as 'large' and 'small'.
- Avoid asking negative questions, as these are easy to misinterpret.
- Include, if possible, questions which serve as cross-checks on the answers to other questions.
- Avoid questions which require participants to perform calculations.
- Avoid asking 'difficult' questions, for example, where the respondent may struggle to answer (people hate to look stupid by not knowing the 'answer').
- Decide which questionnaire to use: closed or open-ended; self or interviewer administered.
- Questions should be kept short and simple. Avoid double barrelled, that is, two questions in one—ask two Qs rather than one.
- Question should not contain prestige bias—causing embarrassment or forcing the respondent to give false answer in order to look good. Questions about educational qualification or income might elicit this type of response.
- Use indirect questions for sensitive issues. In indirect questions, respondents can relate their answer to other people.
- Using closed-ended questions, try to make sure that all possible answers are covered so that respondents are not constrained in their answer. 'Don't Know' category also needs to be added.
- Avoid leading question: Don't lead the respondent to answer in a certain way; for example, 'How often do you wash your car?' assumes that respondent has a car and he or she washes his or her car. Instead, ask a filter question to find if he or she has a car, and then, 'If you wash your car, how many times a year?'
- Keep the questionnaire as short as possible.
- Ask easy Qs which respondents will enjoy answering.
- If it is combined questionnaire, keep open-ended Qs for the end.
- Make Qs as interesting as possible and easy to follow by varying type and length of question.
- Group the Qs into specific topic as this makes it easier to understand and follow.
- Layout and spacing are important, as a cluttered questionnaire is less likely to be answered.
- Once you have constructed your questionnaire, you must pilot it. This means that you must test it out to see if it is obtaining the result you require. This is done by asking people to read it through and see if there are any ambiguities which you have not noticed. They should also be asked to comment about the length, structure and wording of the questionnaire. Alter the questions accordingly.

Questionnaire Formatting

The general format of the questionnaire is something that is just as important as the wording of the questions asked. If your questionnaire is poorly formatted, it can lead respondents to miss questions, confuse respondents or even cause them to throw the questionnaire away.

First, your questionnaire should be 'spread out and uncluttered'. Oftentimes you may fear that your questionnaire looks too long and therefore you try to fit too much onto each page. Instead, each question should be given its own line. You should not try to fit more than one question on a line because that could cause the respondent to miss the second question or get confused.

Second, you should never abbreviate words in an attempt to save space or make a questionnaire shorter. Abbreviating words can be confusing to the respondent and not all abbreviations will be interpreted correctly. This could cause the respondent to answer the question in a different way or skip it entirely.

At last, you should always provide an 'ample space' between questions on each page. Questions should not be too close together. Leaving a double space between each question is ideal as it will ensure where one ends and another begins.

Formatting Individual Questions

In many questionnaires, respondents are required to check and verify one response from a series of responses. You may find a square or circle next to each response for the respondent to check or fill or tick in, or the respondent might be instructed to circle their response. Whatever method you use, you should make sure that instructions are made clear and displayed prominently next to the question. If a respondent indicates their response in a way that is not expected, this could hold up data entry or cause data to be miss-entered.

You should also make sure that response choices are equally spaced. For instance, if your response categories are 'yes', 'no' and 'maybe', all three response categories should be equally spaced from each other on the same page. You should not write 'yes' and 'no' to be right next to each other, while 'maybe' is three inches away. This can lead to confusion, thus misleading the respondents. This may cause them to choose a different answer than expected.

Question Wording

You should pay utmost care in wording of questions and their corresponding response choices (options) in your questionnaire, as both of them are very important. If you are asking a question with the slightest difference in wording, it could result in a different answer or could cause the respondent to misinterpret the question.

Quite often, we make the mistake of making questions unclear and ambiguous. While writing questions, you should be sure that each question is clear and unambiguous. Researchers commonly overlook it. Often they are so grossly involved in the topic of the research and have been studying it for so long that opinions and perspectives seem clear to them when they might not be so to an outsider. Conversely, it might be a totally new topic and one that the researcher has only a superficial knowledge and understanding of, so the question might not be precise and specific enough. You should see that your questionnaire items, particularly the question and the response categories, are so precise that the respondent knows exactly what you are expecting and asking.

You have to exercise caution about asking respondents for a single answer to a question that actually has multiple parts. This is called a double-barrelled question. For example, let's say you ask respondents whether they agree or disagree with this statement: Do you like Obama and his political party? While many respondents might agree or disagree with this statement, many would be unable to provide an answer. They might think Donald Trump is a good president, but they may not like his political party.

Others might like his political party but not him. Therefore, the double-barrelled questions of this kind would be misleading the respondents.

As a general rule, whenever the word 'and' appears in a question or response category, it signifies that you are likely asking a double-barrelled question. You should ask multiple-choice questions instead.

Ordering Items in a Questionnaire

The ordering of questions (the sequence of asking the questions) is also important as it can affect responses. First, the appearance of one question can affect the answers given to later questions. For instance, if there are several questions at the beginning of a survey that asks about the respondents' views on terrorism and then following those questions is an 'open-ended question' asking the respondent what they believe to be dangers to humanity in general, terrorism is likely to be cited more than it otherwise would be. It would be better if you ask the open-ended question first before the topic of terrorism is 'put' into the respondents' head.

You should make efforts to order the questions in the questionnaire so they do not affect the responses of subsequent questions. You may sometimes find this quite hard and nearly impossible to do with each question; however, you should try to estimate what the various effects of different question orders would be and choose the ordering with the smallest effect.

Questionnaire Instructions

You should make sure that every questionnaire, no matter how it is administered, should contain very clear instructions as well as introductory comments when appropriate. The instruction should be short so as to help the respondent make sense of the questionnaire and make the questionnaire seem less chaotic. The instructions should also help put the respondent in the proper frame of mind for answering the questions.

You should write basic instructions for completing the questionnaire at the very beginning of the survey. You should tell the respondent exactly what is wanted from him or her: that they are to indicate their answers to each question by placing a check mark or X in the box beside the appropriate answer or by writing their answer in the space provided when asked to do so.

If there is one section on the questionnaire with closed-ended questions and another section with open-ended questions, for example, you should provide the instructions at the beginning of each section, that is, you should write instructions for both the closed-ended questions and the open-ended questions just above those questions rather than writing them all at the beginning of the questionnaire.

When to Use Questionnaires?

You have seen that the questionnaire is a well-established tool within educational and social science research for acquiring information on participant's social characteristics, present and past behaviour, standards of behaviour and reasons for action with regard to the topic under investigation. You can very conveniently use a questionnaire when:

- You have limited resources and you need data from many people. You can disseminate questionnaires relatively inexpensively.
- You rigorously want to protect the privacy of participants.

To sum up, a questionnaire is the link between the interviewer and the respondent. In a good interview, the respondent feels more like an interesting conversation than an interrogation. The success of your interview will depend on the combination of a good interviewer and a good questionnaire.

Surveys

If you have ever been sitting at a train station, a particular lecturer's classroom or in a public area and a person with a stack of papers in his or her hands approaches you out of the blue and asks if you have a few minutes to talk, then you have likely been asked to take part in a survey. A 'survey' is a data collection tool used to gather information about and learn from your respondents.

Surveys provide a comprehensive, representative summary of specific characteristics, beliefs, attitudes, opinions or behaviour patterns of a population. There are a number of different methods of collecting survey information such as in-person (face-to-face interview), telephone, mail, and e-mail questionnaires.

Surveys are always conducted in response to particular research questions, generally qualitative in nature. The aim of survey is to collect only information that might be relevant to the study at hand. It focuses primarily on factual information about individuals, or it aims to collect the opinions of the survey takers.

Surveys, though primarily quantitative in design and implementation, are always developed in response to a qualitative inquiry, unlike other forms of quantitative research (Lodico, Spaulding and Voegtle 2006). Surveys are not experimental, so data are not collected to test a hypothesis, but rather, to describe—both qualitatively and quantitatively, but primarily quantitatively—existing conditions and attitudes.

In an educational setting, researchers widely use surveys particularly in descriptive research studies. For instance, surveys have been used to gather information on test scores in order to identify patterns of low achievement, to form impressions of new teachers' attitudes towards teaching and to identify trends in student interests. Survey research has also been used (Mellard, Patterson and Prewett 2007) to collect information about adult students' reading patterns and about students' traits in order to identify ways to encourage adults to read.

Surveys in Quantitative Research

There are couple of different ways in which you can administer surveys. For instance, in the 'structured interview method', the interviewer asks each interviewee the questions. In the 'questionnaire', the interviewee fills out the survey on his or her own. Though neither format requires specific types of questions: interviews might require interviewees to choose their answers between previously determined categories, while in questionnaires the interviewer may ask for open-ended responses. Di Iorio (2005) suggests using open-ended questions for gathering information of a personal or sensitive nature and multiple-choice questions for easily quantifiable information.

Sometimes researchers use both questionnaires and interviews within a single survey study. For example, Mamlock-Naamam (2007) evaluated a workshop designed to encourage science teachers to create their own curricula through questionnaires—to get a rough quantitative approximation to teacher beliefs—as well as through interviews—to gauge teachers' initial reactions and to probe quantitative answers in more detail. Appendix 10.1 and Appendix 10.2 show a template of a survey questions.

Experts highly recommend survey approach for gathering descriptive information. Surveys can be both 'structured' and 'unstructured'. Structured surveys use formal lists of questions asked of all respondents in the same way. Unstructured surveys let the interviewer probe respondents and guide the interview according to their answers.

Survey research may be 'direct' or 'indirect'. In direct approach, we ask direct questions about behaviours and thoughts. For example, 'Why do you not attend the class?' In indirect approach, the researcher might ask: 'What kind of pupils attend the class?' For questions such as these, surveys provide information that enables the researcher to discover the reasons why the pupil did not attend the class. The approach may suggest all those reasons (factors) of which the consumer is not consciously aware.

Surveys are generally standardised to ensure that they have 'reliability' and 'validity'. Standardisation is also important so that the results can be generalised to the larger population.

Designing A Survey

The key to a good survey is its design. Although surveys might appear easy to design at first blush, there is more than meets our understanding the several steps involved in designing a survey are discussed above in the section dealing with questionnaire design). Thus, it is necessary to choose the survey items carefully to produce the data needed to answer your research questions. Survey items should be clear and precise and should not bias a respondent towards particular answers (such as socially desirable responses). When the survey is the main data collection instrument in a study, you should always include the survey design in an appendix or make it available upon request.

It is also important to keep in mind that the 'response rate' of a mailed questionnaire is generally low compared to the one administered in person (face-to-face interview). In case you are using mailed questionnaire for your investigation, you should always report the response rate and discuss the implications if it is low (i.e., less than 75%). Because a low response rate might not ensure the 'representativeness' of the group of persons to whom you mailed the questionnaire. It is particularly important in a comparative descriptive study to inform the reader whether the response rates in your study were different for the different groups.

The first step in the design of a survey study is the key step. It is the articulation of the research question, of the scope of the study, and of the targeted population. Following this step comes your research question. You should subdivide your research question into themes or subquestions that reflect different dimensions of the problem. Then you formulate the survey items and use 'scaling technique' in their selection (see further). Scaling is a method of generating questions and statements that adequately measure—quantitatively—the qualitative aspects of the attitude or belief under investigation.

After survey items are selected, and after the questionnaire or interview format are completed, your survey must undergo 'pilot testing' before you use it to collect and generalise information about the targeted population.

Scaling Techniques

We have seen above a classification system that is often used to describe the measurement of concepts or variables that are used in educational and behavioural research. This classification system categorises the variables as being measured either by a nominal, ordinal, interval or ratio scale. For example, when your question has the response choices of 'always', 'sometimes', 'rarely' and 'never', you should use this scale because the answer choices are rank ordered and have differences in intensity. Another example of scaling could be 'strongly agree', 'agree', 'neither agree nor disagree', 'disagree' and 'strongly disagree'.

We should keep in mind that scaling is not simply a process of assigning numerical values to qualitative statements such as a score of 5 for 'highly agree' or of 1 for 'highly disagree'. Scaling is the intricate, nuanced technique of developing qualitative statements that gauge respondents' beliefs about a particular issue; thus, scaling research always begins with the clear identification of a research question or aim. Researchers usually conduct experimental studies to determine appropriate qualitative statements for use on surveys that adequately gauge the desired features in the population of interest.

There are several formalised conceptual and experimental frameworks commonly used in scaling research; the most relevant to educational researchers are discussed next.

Likert Scale

Likert scales are one of the most commonly used scales in social science research. On a survey or questionnaire, the Likert scale typically has the following format:

- Strongly agree
- Agree
- Neither agree nor disagree (neutral)
- Disagree
- Strongly disagree

Individual questions that use this format are termed as Likert items, while the Likert scale is the sum of several Likert items. To create the scale, each answer choice is assigned a score, say 0–4, (some studies use 1–5) and the answers for several Likert items (that measure the same concept) can be summed together for each individual to get an overall Likert score.

For example, let us say that you are interested in measuring prejudice against women. A simple way to do that would be to develop a series of statements reflecting prejudice ideas against women, each with the Likert response categories listed above. One of the statements could be 'women should not be given right to vote', while another might be 'women cannot drive well compared to men'. You would then assign each of the response categories a score of 0 to 4, that is, 0 = 'strongly disagree', 1 = 'disagree', 2 = 'neither agree nor disagree Neutral' and so on. Then you sum up all the scores for each of the statements for each respondent to create an overall score of prejudice. If you had five statements and a respondent answered 'strongly agree' to each item, his or her overall prejudice score would be 20, indicating a very high degree of prejudice against women.

Bogardus Social Distance Scale

The Bogardus social distance scale is a technique for measuring the willingness of people to participate in social relations with other kinds of people.

Let us say we are interested in the extent to which Hindus in India are willing to associate with, say, SC Hindus. We might ask the following questions:

- Are you willing to live in India as SC?
- Are you willing to live in the same community as SC?
- Are you willing to live in the same neighbourhood as SC?

- Are you willing to live next door to a SC?
- Are you willing to let your child marry a SC?

The Bogardus scale signifies that scales can be important data reduction tools. By knowing how many relationships with SC a given upper class Hindu respondent will accept, we know which relationships were accepted. A single number can thus accurately outline five or six data items without a loss of information.

Thurstone Scale

The Thurstone scale is intended to develop a format for generating groups of indicators of a variable that have an empirical structure among them. For example, if you were studying discrimination, you would put together a list of items (e.g., 10 items) and then ask respondents to assign scores of 1 to 10 to each item. Basically, Thurstone scales are a method of ranking the items in order of: 'which is the weakest indicator of discrimination all the way to which is the strongest indicator'.

Once the respondents have scored the items, only then you examine the scores assigned to each item by all the respondents to determine which items the respondents agreed upon the most. If you develop the scale items adequately and record the score of scaling correctly, the economy and effectiveness of data reduction present in the Bogardus social distance scale would be apparent.

Semantic Differential Scale

In the semantic differential scale, you ask respondents of your questionnaire to choose between two opposite positions using qualifiers to bridge the gap between them. Let us look at an example. Suppose you want to obtain respondents' opinions about a new comedy television show. The first thing you have to do in this case is to decide what dimensions you wish to measure. And then you find two opposite terms that represent those dimensions. For example, 'enjoyable' and 'not enjoyable', 'funny' and 'not funny', 'relatable' and 'not relatable'. You should then develop a rating sheet for each respondent to indicate how he or she feels about the television show in each dimension. Your response sheet would look something like the one shown in Table 10.1.

Guttman Scaling

Researchers use Guttman scaling techniques for determining whether a set of indicators constitute a scale and, if so, how to score them. The technique enables the researcher derive a single dimension that can be used to position both the questions and the subjects. The position of the questions and subjects

TABLE 10.1 Dimensions of Responses

	Very Much	*Somewhat*	*Neither*	*Somewhat*	*Very Much*
Enjoyable		×			Not enjoyable
Funny				×	Not Funny
Relatable			×		Not relatable

Source: Adapted from Babbie (2001).

on the dimension can then be used to give them a numerical value (Abdi 2010). Psychologists and educational researchers widely use Guttman scaling.

Suppose that we test a set of children and that we assess their mastery of the following types of mathematical concepts:

1 = counting from 1 to 50
2 = solving addition problems
3 = solving subtraction problems
4 = solving multiplication problems
5 = solving division problems

In this particular study, you will notice that some children who would be unable to master any of these problems, and these children will not provide information about the problems, so you should ignore them. Some children will master counting but nothing more, some will master addition and we expect them to have mastered addition but no other concepts; some children will master subtraction and we expect them to have mastered counting and addition; some children will master multiplication and we expect them to have mastered subtraction, addition and counting. Finally, some children will master division problem and we expect them to have mastered counting, addition, subtraction and multiplication (Abdi 2010). What we do not expect to find, however, are children, for example, who have mastered division but who have not mastered addition or subtraction or multiplication. So the set of patterns of responses that we expect to find is well structured and is shown in Table 10.2. The pattern of data displayed in this figure is consistent with the existence of a single dimension of mathematical ability. In this framework, a child has reached a certain level of this mathematical ability and can solve all the problems below this level and none of the problems above this level.

The rows and the columns of the table can both be represented on a single dimension. The operations will be ordered from the easiest to the hardest and a child will be positioned on the right of the most difficult type of operation solved. So the data from Table 10.2 can be represented by the following order:

Counting = S_1 Addition = S_2 Subtraction = S_3 Multiplication = S_4 Division = S_5

TABLE 10.2 Listing of Responses in Guttman Scale*

	Problems				
Children	*Counting*	*Addition*	*Subtraction*	*Multiplication*	*Division*
S_1	1	0	0	0	0
S_2	1	1	0	0	0
S_3	1	1	1	0	0
S_4	1	1	1	1	0
S_5	1	1	1	1	1

*A value of 1 means that the child (row) has mastered the type of problem (column); a value of 0 means that the child has not mastered the type of problem.
Source: Abdi (2010).

This order can be transformed into a set of numerical values by assigning numbers with equal steps between two contiguous points. For example, this set of numbers can represent the numerical values corresponding to Table 10.2:

Counting S_1		Addition S_2		Subtraction S_3		Multiplication S_4		Division S_5	
1	2	3	4	5	6	7	8	9	10

This scoring scheme implies that the score of an observation (i.e., a row in Table 10.2) is proportional to the number of non-zero variables (i.e., columns in Table 10.2) for this row.

Advantages and Limitations of Surveys

Advantages

- Surveys allow researchers to collect a large amount of data in a relatively short period of time.
- Surveys are less expensive than many other data collection techniques.
- Surveys can be created quickly and administered easily.
- Surveys can be used to collect information on a wide range of things, including personal facts, attitudes, past behaviours and opinions.

Limitations

- Poor survey construction and administration can undermine otherwise well-designed studies.
- The answer choices provided on a survey may not be an accurate reflection of how the participants truly feel.
- While random sampling is generally used to select participants, response rates can bias the results of a survey.
- Participants may be reluctant to answer questions asked by unknown interviewers about things they consider private.
- Busy people may not want to take the time to answer questions; instead, they may try to help by giving pleasant answers.
- Participants may be unable to answer because they cannot remember or never gave a thought to what they do and why; they may still answer in order to appear smart or well informed.

Summary

The questionnaires are the primary tools in collecting necessary information from the respondents of a survey. By making the right choices on the type of survey questions, you will be able to extract only data that are related to the purpose or goal of the survey. The best questionnaires are constantly edited and refined until finally they have clear questions and instructions, laid out in a logical order.

Questionnaires and survey designs and techniques are used in an empirical research inquiry. They have advantages over many types of surveys because they are cheap and can easily target groups of interest in many ways and they can be used to target a large audience in a wide geographical area. No matter which approach is chosen to collect data, it is impossible to get 100 per cent response rates. All types of surveys have advantages and disadvantages; in some of them, the researchers have more control over the situation; for example, the respondents can be influenced or get clues from interviewers to respond in certain ways.

Finally, the questionnaires method is the most commonly used technique in research because they can be applied in many different areas; for example, education, science, health care surveys and so on; whereas interviews can be too expensive, difficult to arrange and time-consuming to be applied in a large geographical area as was cited in Chapter 8.

Chapter 11 presents several statistical methods and techniques of analysing the data collected from different tools discussed in Chapter 9 and in this chapter.

Self-test Exercises

Exercise 10.1:
10.2.1. How would you design a questionnaire to find out the following about yourself and your job?

- The racial or ethnic group you belong
- Your sex
- Your current annual income in US dollars before tax and other deductions
- Your date of birth
- Year you were registered in this university

10.2.2. Design a questionnaire to study the modern methods of documentation and library work in your university highlighting the following:

- When you last used the documentation centre/library
- Which type of materials (newspapers, books, periodicals and so on) you used
- How far these materials were useful for the purpose of your research work?

Exercise 10.2:
10.2.1. Write an introduction to a questionnaire that you think would be successful in winning your cooperation. The introduction should include all the necessary coverage of who is carrying out the survey (not necessarily who is commissioning it), promises of anonymity and confidentiality, how long it will take and a persuasive hook. See if you can use less than 100 words.

Exercise 10.3:
10.3.1. How many times per year do you visit the community primary school? _________________
10.3.2. Do you visit the school for inspection?
☐ Yes ☐ No
10.3.3. The school is accessible to all children of the community.
Please tick (☑) the appropriate box:
☐ Strongly Agree
☐ Agree
☐ Neutral
☐ Disagree
☐ Strongly Disagree
10.3.4. How would you rate your overall experience at the school?
☐ Highly satisfactory
☐ Satisfactory
☐ Neutral

- ☐ Unsatisfactory
- ☐ Highly Unsatisfactory

10.3.4.1. What could you suggest to make the school better?

____

____

____

Exercise 10.4:

10.4.1. While conducting a survey, the questions you ask are who, what, when, where, why and how. Which of these do you think is the most difficult for people to answer? Why is it the most difficult? Describe briefly.

10.3.2. How can you create a survey for students? Give an example.

References

Abdi, H. 2010. 'Guttman Scaling'. In *Encyclopaedia of Research Design*, edited by Neil Salkind. Thousand Oaks, CA: SAGE, 558–560 Publications.

Babbie, E. 2001. *The Practice of Social Research*, 9th ed. Belmont, CA: Wadsworth Thomson.

Collis, J. and R. Hussey. 2003. *Business Research: A Practical Guide for Under-graduate and Post-graduate Students*, 2nd ed. Basingstoke: Palgrave Macmillan.

Di Iorio, C.K. 2005. *Measurement in Health Behavior: Methods for Research and Evaluation*. San Francisco, CA: Jossey-Bass.

Lodico, M.G., D.T. Spaulding and K.H. Voegtle. 2006. *Methods in Educational Research: From Theory to Practice*. San Francisco, CA: Jossey-Bass.

Mamlock-Naamam, R. 2007. 'Involving Science Teachers in the Development and Implementation of Assessment Tools for "Science for All" Type Curricula'. *Journal of Science Teacher Education* 18 (4): 497–524.

Mellard, D., M.B. Patterson and S. Prewett. 2007, April/May/June. 'Reading Practices among Adult Education Participants'. *Reading Research Quarterly* 42 (2): 188–213.

Postlethwaite, N. 2005. *Educational Research: Some Basic Concepts and Terminology Module 1*. Paris: UNESCO International Institute for Educational Planning.

Sheldon, E.B. and R. Parke. 1975. 'Social Indicators'. *Science* 188: 693–699.

Siniscalco, M.T. and N. Auriat. 2005. *Quantitative Research Methods in Educational Planning: Questionnaire Design*, Module 8. Paris: UNESCO International Institute for Educational Planning.

Appendix 10.1. Example of a Questionnaire in an Educational Setting[1]

A. SCHOOL IDENTIFICATION

Name of the school ____

Address of the school ______________________________________

[1] Adapted from Siniscalco, M. T. and Auriat, N. (2005): *Quantitative Research Methods in Educational Planning,* Module 8: Questionnaire Design, UNESCO International Institute for Educational Planning (IIEP), Paris

B. STUDENT BACKGROUND

Gender and age

Q1. Are you a boy or a girl?
(Tick one box only.)
☐ Boy
☐ Girl

Q2. What is your date of birth?

Day	Month	Year
____	________	______

Q3. How old were you at your last birthday?
Age ________

C. SOCIO-ECONOMIC BACKGROUND: OCCUPATION, EDUCATION AND POSSESSIONS

Parent's occupation

Q4. Does your father work for pay now?

Yes, he now works full time	1	(GO TO Q.....)
Yes, he now works part-time..	2	(GO TO Q.....)
No, he is now looking for work	3	(GO TO Q.....)
No, he is not working at present (unemployed, retired)........	4	(GO TO Q.....)

Q5. What is the occupation of your father (or the male person responsible for your education)?
(Please, describe as clearly as possible.)

____

____

____

Q6. In your opinion, how can the occupation of your father be defined?
(Please tick only one box.)
☐ Professional and managerial
☐ Clerical and sales
☐ Skilled and blue-collar
☐ Semi-skilled and unskilled

Parent's education

Q7. What is the highest level of education that your father (or the male person responsible for your education) has completed?
(Please tick only one box.)
☐ Never went to school
☐ Completed some primary school
☐ Completed all of primary school
☐ Completed some secondary school
☐ Completed all of secondary school
☐ Completed some education/training after secondary school
☐ Don't know

Q8. How many years of academic education has your father/guardian completed?
__________ years of primary school
__________ years of secondary school
__________ years of post-secondary academic education

Possessions in the home

Q9. About how many books are there in your home?
(Do not count newspaper or magazines. Please tick one box only)
☐ None
☐ 1–10
☐ 11–50
☐ 51–100
☐ 101–200
☐ More than 200

Q10. Which of the following items can be found in your home?
(Please, circle one number for each line.)

	No	Yes
Radio	1	2
TV set	1	2
Video cassette recorder	1	2
Telephone	1	2
Refrigerator	1	2
Car	1	2
Piped water	1	2

D. TEACHER CHARACTERISTICS

Q11. How many years of academic education have you completed?
(Do not count grade repetition years. Part-time years should be converted into full-time years. For example, two half-years equals one full year.)
__________ years of primary school
__________ years of lower secondary school
__________ years of upper secondary school
__________ years of post-secondary academic education

Q12. How many years of pre-service teacher training have you received altogether?
(Please, encircle only one choice.)
(a) I did not receive any teacher training
(b) I have had a short course of less than one-year duration
(c) I have had a total equivalent of one year
(d) I have had a total equivalent of two years
(e) I have had a total equivalent of three years
(f) I have had a total equivalent of more than three years

E. SCHOOL LOCATION

Q13. What type of community does your school serve?
(Please tick only one box.)
☐ A village or rural community
☐ A small town community
☐ A large town
☐ A city

Q14. How many kilometres by road is it from your school to the places in the list below?
(a) the nearest health centre/clinic __________ kilometres
(b) the nearest asphalt/tarmac/tarred road __________ kilometres
(c) the nearest public library __________ kilometres

(d) the nearest secondary school __________ kilometres
(e) the nearest city __________ kilometres
(f) the nearest regional capital __________ kilometres

Q15. How often do you read these types of publications for personal interest and leisure?

Publication type	***Rarely***	***Less than once a month***	***One or two times a month***	***About once a week***	***Two or three times a week***	***Most days***
Mystery	1	2	3	4	5	6
Romance	1	2	3	4	5	6
Sport	1	2	3	4	5	6
Adventure	1	2	3	4	5	6
Music	1	2	3	4	5	6
Nature	1	2	3	4	5	6

F. TEACHER ACTIVITIES

Q16. During the school year, how often do you teach comprehension of each of the following kinds of texts?
(Circle one number per line.)

Kind of text	***Almost never***	***About 3–4 times a year***	***About once a month***	***At least once a week***	***Nearly every day***
a. Narrative text (that tells a story or gives the order in which things happen)	1	2	3	4	5
b. Expository text (that describes things or people, or explains how things work or why things happen)	1	2	3	4	5
c. Documents (that contain tables, charts, diagrams, lists and maps)	1	2	3	4	5

G. SCHOOL HEAD ACTIVITIES

Q17. Does your school have any special programmes or initiatives for reading outside normal classroom activities?
(You may tick more than one.)
Extra-class lessons in reading ☐
Extra-individual tuition in reading at school ☐
Special remedial reading courses ☐
Other (specify)……………… ☐
None ☐

Q18. Please rank the following activities in order of importance in your work as a school head, with ‘1’ being the most important activity and ‘6’ the least important activity.
Importance ranking
(a) Evaluating the staff __________
(b) Discussing educational objectives with teachers __________
(c) Pursuing administrative tasks __________
(d) Organizing in-service teacher training courses __________

(e) Organizing extra-class special programmes __________
(f) Talking with students in case of problems __________

H. ATTITUDES, OPINIONS AND BELIEFS

Q19. School is a place where I usually feel great
☐ Strongly agree
☐ Agree
☐ Neither agree nor disagree
☐ Disagree
☐ Strongly disagree

Q20. The teachers at my school are often unfair
☐ Strongly agree
☐ Agree
☐ Neither agree nor disagree
☐ Disagree
☐ Strongly disagree

Q21. I really like going to school
☐ Strongly agree
☐ Agree
☐ Neither agree nor disagree
☐ Disagree
☐ Strongly disagree

Q22. Going to school makes me feel important
☐ Strongly agree
☐ Agree
☐ Neither agree nor disagree
☐ Disagree
☐ Strongly disagree

Q23. School is a place where I sometimes feel depressed
☐ Strongly agree
☐ Agree
☐ Neither agree nor disagree
☐ Disagree
☐ Strongly disagree

Appendix 10.2. Example of a Survey Questionnaire

1. Specific Information Request
   In which year did you start the degree course? __________
2. Category
   Have you ever been or are you a student representative? (Please tick only one box.)
   ☐ Yes (currently)
   ☐ Yes (in the past)
   ☐ Never

3. Multiple Choices
   Do you view the money you have spent on your higher education as any of the following? If so, tick which.
   ☐ A luxury
   ☐ A necessity
   ☐ A gamble
   ☐ A burden
   ☐ A right
   ☐ None of these
4. Scale
   How would you describe your parents' attitude to higher education at the time you applied? Please tick one of the following options.
   ☐ Very positive
   ☐ Positive mixed/neutral
   ☐ Negative
   ☐ Very negative
   ☐ Not sure
5. Ranking
   What do you see as the main purpose(s) of your degree study? Please rank all those relevant in order from 1 (most important) downwards.
   ☐ Personal development
   ☐ Career advancement
   ☐ Subject interest
   ☐ Recreation
   ☐ To fulfil ambition
   ☐ Intellectual stimulation
   ☐ Other (give details) ____
6. Grid or Table
   How would you rank the benefits of your degree study for each of the following? Please rank each item: (For very positive = 1; Positive = 2; Neutral = 3; Negative = 4; Very negative = 5; Not sure = 6)
   ☐ You
   ☐ Your family
   ☐ Your employer
   ☐ The country
   ☐ Your community
   ☐ Your friends
7. Open Questions
   Please summarize the benefits of your degree study in the space below:
   ____
   ____

11 CHAPTER

Data Analysis: Classification and Tabulation

Introduction

Data analysis is one of the main steps of the research process. How to analyse the data is an important question that every researcher asks. It is an important step in answering an experimental question. Analysing data from a well-designed study helps the researcher answer questions. With these data, the researcher can also draw conclusions that further the research and contribute to future studies. Keeping your data well organised during the collection process will help you make the analysis step that much easier. The researcher collects the data using one of the qualitative or quantitative methods of data collection discussed in Chapters 7, 8 and 9. Data analysis highly depends on whether the data are qualitative or quantitative.

The process of analysing information calls for examining it in ways that disclose the relationships, patterns, trends and so on that can be found within it. That may mean putting it to statistical operations that can tell us not only what kinds of relationships seem to exist among variables but also to what level we can trust the answers we are getting. It may mean comparing our information to that from other groups (a control or comparison group, statewide figures and so on) to help draw some conclusions from the data. The point, in terms of our analysis, is to get an accurate assessment in order to better understand our work and its effects on those we are concerned with or in order to better understand the overall situation.

The purpose of this chapter is to review some of the fundamental concepts and terms used in data analysis that are shared across the social sciences. You should familiarise (or re-familiarise) yourself with this material before proceeding to the subsequent chapters, as most of the terms introduced here will be referred to again and again throughout the text. If you are currently taking your first course in statistics, Chapters 11 and 12 provide an elementary introduction to statistical methods commonly used in research. If you have already completed a course in statistics, they provide a quick review. The chapter is primarily practical, but also includes some theoretical concepts necessary for understanding many of the topics that will be covered in this chapter and in succeeding chapters.

Meaning of Data Analysis

Data analysis has multiple facets and approaches. It encompasses diverse techniques under a variety of names in different business, science and social science domains. For instance, data mining is a particular data analysis technique that focuses on modelling and knowledge discovery for predictive rather than purely descriptive purposes. Business intelligence, on the other hand, covers data analysis that relies heavily on aggregation, focusing on business information.

In statistical applications, some researchers divide data analysis into descriptive statistics (DS), exploratory data analysis (EDA) and confirmatory data analysis (CDA). DS are used to describe the basic features of the data in a study. They provide simple summaries about the sample and the measures. EDA helps researcher to discover new features in the data, and CDA focuses on confirming or falsifying existing hypotheses.

Similarly, predictive analytics focuses on application of statistical or structural models for predictive forecasting or classification, while text analytics applies statistical, linguistic and structural techniques to extract and classify information from textual sources, a species of unstructured data. All are varieties of data analysis.

Whatever may be the analytical purpose of researchers, data analysis is the process of scanning, examining and interpreting data available in a tabulated form. It is the procedure of evaluating data using analytical and logical reasoning to examine each component of the data provided.

> 'Data analysis is a systematic search for meaning. It is a way to process qualitative data so that what has been learned can be communicated to others. Analysis means organizing and interrogating data in ways that allow researchers to see patterns, identify themes, discover relationships, develop explanations, make interpretations, mount critiques, or generate theories. It often involves synthesis, evaluation, interpretation, categorization, hypothesizing, comparison, and pattern finding' (Hatch 2002).

In its simplest form, data analysis is, thus, a process of inspecting, cleaning, transforming and modelling data with the goal of discovering useful information, suggesting conclusions and supporting decision-making. It is the process by which large amounts of raw data are reviewed in order to determine conclusions based on that data.

Data analysis is just one of the many steps that must be completed when conducting a research experiment. You should always review the data that you have gathered from various sources, and then analyse it to form some sort of finding or conclusion. There are several specific methods of data analysis, some of which include data mining, text analytics, business intelligence and data visualisations.

Why to Analyse Data

The underlying purpose of data analysis is to understand the nature of the data and reach a conclusion. In fact, data analysis provides answers to the research questions or research problems that you have formulated. Without analysing the data, you cannot draw any conclusion and inferences. A clear understanding of how to properly handle the data analysis will allow you to get the most from your data and make the right decisions. The data analysis by the analyst is instrumental in understanding the current situation so we then know which steps to be taken to grow. After analysing data, we get an organised and well-examined form of data that can help us know whether our hypothesis is accepted or rejected.

In all research proposals, the researcher discusses how he or she will conduct an analysis of his or her data. By the time he or she gets to the analysis of his data, most of the really difficult work has been done. It is much more difficult to define the research problem and implementing a sampling plan, develop a design structure and determine the measures. If a researcher has done this work well, the analysis of the data is usually a fairly straightforward affair.

The prime objective of analysing data is to obtain usable and useful information. The analysis, regardless of whether the data are qualitative or quantitative may assist you to:

- Describe and summarise the data
- Identify relationships between variables
- Compare variables
- Identify the difference between variables
- Forecast outcomes

With the process of data analysis, we bring order, structure and meaning to the mass of collected data. It is a messy, ambiguous, time-consuming, creative and fascinating process. It does not proceed in a linear fashion; it is not neat. Qualitative data analysis is a search for general statements about relationships among the categories of data (Marshall and Rossman 2006).

Earlier in Chapters 3 and 4, we distinguished between qualitative and quantitative research. It is highly unlikely that a research study will be purely one or the other—it will probably be a 'mixture' of the two approaches.

For example, you intend to do an ethnographic research (a rigorous investigation of a situation or problem in order to generate new knowledge or validate existing knowledge), which is qualitative. In your first step, you will select a small sample (normally associated with qualitative research) but then execute a structured interview or use a questionnaire (normally associated with quantitative research) to ascertain people's attitudes to a particular phenomenon (qualitative research). It is therefore likely that your mixed approach will take a qualitative approach some of the time and a quantitative approach at others, depending on the needs of your investigation.

A source of confusion for many researchers is the belief or confidence that qualitative research generates just qualitative data (text, words, opinions and so on) and that quantitative research generates just quantitative data (numbers). Often this is the case, but both types of data can be generated by each approach. For instance, a 'questionnaire' (quantitative research) will often gather factual information such as age, salary, length of service (quantitative data) and so on—but may also gather information on opinions and attitudes (qualitative data).

When it comes to data analysis, some experts and researchers believe that statistical techniques are only applicable for quantitative data. This is not so. There are many statistical techniques that can be applied to qualitative data, such as ratings scales, that have been generated by a quantitative research approach. Even if a qualitative study uses no quantitative data, there are many ways of analysing qualitative data. For example, once you have conducted an interview, your first stage of data analysis is transcription and organisation of data. This would then be followed by a systematic analysis of the transcripts—grouping together comments on similar themes and attempting to interpret them and draw conclusions.

Collecting 'quantitative data'—information expressed in numbers—and subjecting it to a visual inspection or formal statistical analysis can tell you whether your work is having the desired or anticipated effect and may be able to tell you why or why not as well. It can also single out connections (correlations) among variables and call attention to factors you may not have considered.

On the other hand, collecting and analysing 'qualitative data'—interviews, descriptions of environmental factors or events and circumstances—can provide insight into how participants experience the issue you are addressing, what barriers and advantages they experience and what you might change or add to improve what you do.

It should be pointed out that merely analysing data is not sufficient from the point of view of making a decision. How does one interpret from the analysed data is more important. Thus, 'data analysis is not a decision-making system, but decision-supporting system'.

Types of Data Analysis

There are several different ways of data analysis, all geared towards the nature of the data being analysed. Generally speaking, there are two most widely used categories of data analysis: (a) quantitative analysis and (b) qualitative analysis.

'Qualitative analysis' handles the data that are categorical in nature. In simple terms, in qualitative analysis data are not described in numerical values, but rather by some sort of descriptive context such as text.

Qualitative analysis serves the following three basic principles (Seidel 1998):

- Notice things
- Collect things
- Think about things

'Quantitative analysis' is the process by which numerical data are analysed and often involves DS. The following statistical methods are widely used in quantitative data analysis:

- Statistical models
- Analysis of variables
- Data dispersion
- Analysis of relationships between variables
- Contingence and correlation
- Regression analysis
- Statistical significance
- Precision
- Error limits

Table 11.1 highlights the difference between qualitative and quantitative data analysis.

TABLE 11.1 Comparison of Qualitative and Quantitative Data Analysis

Qualitative data	*Quantitative data*
• Data are observed • Involves descriptions • Emphasis is on quality • Examples are colour, smell, taste	• Data are measured • Involves numbers • Emphasis is on quantity • Examples are volume, weight and so on

Source: Authors.

Benefits of Data Analysis

How can you improve your processes and identify problematic issues if you are not willing to look at the data? The simple answer, of course, is that you hardly make any reliable and accurate improvements without analysing your data.

Having said that, data analysis can offer the following benefits. It

- Allows meaningful insights from the data set
- Highlights critical decisions from the findings
- Allows a visual view leading to faster and better decisions
- Offers better awareness regarding the habits of potential customers
- Structures the findings from survey research or other means of data collection
- Breaks a macro picture into a micro one
- Rules out human bias through proper statistical treatment

Nature of Statistical Data: Variables and Attributes

There are two kinds of data a researcher is apt to be working with, although not all research will necessarily include both. We have seen earlier that quantitative data refer to the information that is collected as, or can be translated into, numbers, which can then be displayed and analysed mathematically. Qualitative data, on the other hand, are collected as descriptions, anecdotes, opinions, quotes, interpretations and so on and are generally either not able to be reduced to numbers or are considered more valuable or informative if left as narratives. As expected, quantitative and qualitative information should be analysed differently.

Thus, data must either be numeric in origin or transformed by researchers into numbers for analysis. For instance, statistics could be used to analyse percentage scores English students receive on a grammar test: the percentage scores ranging from 0 to 100 are already in numeric form. Statistics could also be used to analyse grades on an essay by assigning numeric values to the letter grades, for example, A = 4, B = 3, C = 2, D = 1 and E = 0.

It is important to remember that statistics is much more than just the tabulation of numbers and the graphical presentation of these tabulated numbers. Statistics is the science of gaining information from numerical and categorical data. Statistical methods can be used to analyse data to find answers to the questions like:

- What kind and how much data need to be collected?
- How should we organise and summarise the data?
- How can we analyse the data and draw conclusions from it?
- How can we assess the strength of the conclusions and evaluate their uncertainty?

That is, statistics provides methods for the following:

- Design: Planning and carrying out research studies
- Description: Summarising and exploring data
- Inference: Making predictions and generalising about phenomena represented by the data

Statistics is also frequently used for the purpose of prediction (forecast). Prediction relies on the concept of generalisability: if you collect enough data and compiled a particular context (e.g., students studying writing in a specific set of classrooms), the patterns revealed through your analysis of the data collected about that context can be generalised or predicted to occur in similar contexts. All of these important elements have been discussed further in detail.

In the course of writing your dissertation or thesis, one of the first terms that you encounter is the word 'variable'. Failure to understand the meaning and the usefulness of variables in your study will prevent you from doing good and meaningful research. What then are variables and how do you use variables in your study? We explain in the following paragraph the concept with lots of examples on variables commonly used in research.

In social research, both variables and attributes represent social concepts. You would not be able to do very much in research unless you know how to talk about attributes and variables.

Variables

'Variables' can be defined as those simplified portions of the complex phenomena that a researcher intends to study. The word variable comes from the root word 'vary', meaning changing in amount, volume, number, form, nature or type. Variables should always be measurable, that is, you can count them or put them subject to a scale.

A variable is a 'data item'. Examples of variables could be age, sex, business income and expenses, country of birth, capital expenditure, class grades, eye colour and vehicle, and many more. The reason why we call them variable is because the value may vary between data units in a population and may change in value over time.

For example, 'test score' is a variable that can vary between pupils in a given class (i.e., pupils being studied may not have the same scores) and can also vary over time for each pupil (i.e., tests scores can go up or down).

'A variable is any individual item or entity that can take on different values'. Variables have what social researches call attributes (or categories or values). Anything that varies can be considered a variable. For example, age or income can be considered variables because both can take different values for different people or for the same person at different times. Likewise, coffee and tea are two distinct variables because both of them can be assigned values.

When analysing your data, you should keep in mind that variables are not always 'quantitative' or numerical. For instance, the variable 'city' can consist of text values like 'Beijing' or 'Paris'. For our analysis, if required, we can assign quantitative values (Beijing = 1, Paris = 2) instead of or in place of the text values, but we do not have to assign number in order for something to be a variable. You should also keep in mind that variables are not only things that we measure in the traditional sense. For instance, in educational research and in programme evaluation, we consider the treatment or programme consisting of one or more variables. Seen from this angle, an educational programme can have varying amounts of 'time on task', 'classroom settings' and 'PTR', and so on. It means that even an education programme can be considered a variable which can be made up of a number of sub-variables.

Attributes

An 'attribute' is defined as a characteristic or quality of a variable (something or someone). An attribute refers to a specific value on a variable. To simplify the concept, let us consider, for example, 'male' and 'female'. These are attributes of the variable 'sex' or 'gender'. In other words, the variable sex or gender is composed of these two attributes. Similarly, the variable occupation is composed of attributes such as agriculturist, fisherman, teacher, doctor, driver, engineers, manager and so on. Social class is a variable as it can compose of a set of attributes such as upper class, middle class and lower class. 'Attributes refer to the categories that make-up a variable'. In science and research, attribute is a typical characteristic of an object (person, thing and so on). While an attribute is usually intuitive, a variable is an operationalised way in which the attribute is represented for further data processing. In data processing, data often constitute a combination of items (objects organised in rows) and multiple variables (organised in columns).

In business, an attribute reflects a characteristic or feature of a product that is thought to appeal to customers. Attributes basically represent a manufacturer's or a seller's perspective and not necessarily that of a customer. Attributes of instant coffee, for example, may include its aroma, flavour, colour, caffeine content, packaging and presentation, price, shelf life, source and so on. An attribute assumes only 'two possible ratings' (negative or positive) expressed as acceptable or unacceptable, desirable or undesirable, good or bad and so on. Similarly, the variable 'agreement' might be defined as having five attributes:

1 = Strongly disagree
2 = Disagree
3 = Neutral (neither agree nor disagree)
4 = Agree
5 = Strongly disagree

Age is a variable that can be operationalised in many ways. It can be dichotomised so that only two values—'old' and 'young'—are allowed for further data processing. In this case, the variable 'age' is operationalised as a 'binary variable'. If more than two values are possible and they can be ordered, the attribute is represented by 'ordinal variable', such as 'young', 'middle age' and 'old'.

The 'social class' variable can be operationalised in similar ways as age, including 'lower', 'middle' and 'upper class', and each class could be differentiated between upper and lower, thus changing the three attributes into five like (a) upper-upper, (b) upper-middle, (c) middle-middle, (d) middle-lower and (e) lower-lower. A selected number of variables and their possible attributes are shown further.

Variables	*Attributes*
Age	Young, middle, old
Gender	Male, female
Occupation	Doctor, farmer, teacher
Social class	Upper, middle, lower

A variable uses numerical values to measure an attribute. It is a quantity that expresses a quality in numbers to allow more precise measurement.

You must remember that qualitative research focuses primarily on the meaning of subjective attributes of individuals or groups, whereas quantitative research focuses on the measurement of objective variables that affects individuals or groups.

There are many different types of variables.

Independent variables: These variables constitute the presumed cause. They are introduced under controlled conditions during the experiment as treatments to which experimental groups are exposed. The independent variables are just those variables that may influence or affect the other variable, that is, the dependent variable. The possible relevant and meaningful independent variables for your study on the topic on the phenomenon of 'poor performance of students in college entrance exams' could be as follows:

1. Entrance exam score
2. Number of hours devoted to studying
3. PTR
4. Number of students in the class
5. Educational attainment of teachers
6. Teaching style
7. The distance of school from home
8. Number of hours devoted by parents in providing homework support

In this case, variables 2 to 8 are independent variables, whereas variable 1 'entrance exam score' is a dependent variable. An understanding of the relationship between dependent and independent variables is highly relevant when you are investigating cause-effect relationship.

- Dependent variables: These are the presumed effect. These variables are measured before and after the treatment to see whether any change occurred. It is a variable you are interested in.
- Background variables: These are antecedents that affect the situation prior to the study. They can be observed and measured, but they usually do not change.
- Intervening variables: These are events between the treatment and the post-test measurement that might affect the outcome. They appear in more complex causal relationships.
- Extraneous variables: These are variables that can be observed and which might affect the outcome during the study, but which cannot be controlled. They are the variables other than independent and dependant variables.

In research terminology change, variables are called independent variables, whereas outcome/effect variables are known as dependent variables; the unmeasured variables influencing the cause-and-effect relationship are called extraneous variables and the variables that link a cause-and-effect relationship are called intervening variables.

Categorical variables are measured on nominal or ordinal measurement scales. For continuous variables, the measurements are made either on an interval or a ratio scale (see Chapter 10). Categorical variables are divided into three categories:

- Constant
- Dichotomous
- Polytomous

When a variable can have only two categories as in yes/no, good/bad and rich/poor, it is known as a 'dichotomous variable'. When a variable can be divided into more than two categories, for example, religion (Christian, Muslim and Hindu); political parties (labour, liberal, democrat) and attitudes (strongly favourable, favourable, uncertain, unfavourable, strongly unfavourable), it is called a 'polytomous variable'.

'Continuous variables', on the other hand, have continuity in their measurement. Variables such as age, income and attitude score are continuous variables as they can take on any value on the scale on which they are measured. Income can be measured in dollars and cents, and age can be measured in years, month and days, and so on.

An important question now for you to reflect: How will you know that one variable may cause the other to behave in a certain way? To answer this question or for finding out the relationship between variables, you have to do a thorough 'review of the literature'. A thorough review of the relevant and reliable literature will enable you to find out which variables influence the other variable. You should not just simply guess relationships between variables. You have to understand the whole process which is basically the essence of research.

Parametric and Non-parametric Data

One of the potential sources of confusion in working out what statistics to use in analysing data is whether our data allow for parametric or non-parametric statistics. The importance of this issue should not be neglected. If you get it wrong, you risk using an incorrect statistical procedure or you may use a less powerful procedure.

'Non-parametric' statistical procedures are less strong or powerful because these variables use less information in their calculation. For example, a parametric correlation uses information about the mean and deviation from the mean, while a non-parametric correlation uses only the ordinal position of pairs of scores.

The basic distinction for parametric versus non-parametric is:

If our measurement scale is nominal or ordinal, then we use non-parametric statistics. On the other hand, if we are using interval or ratio scales we use parametric statistics. The other considerations which you have to take into account are:

- You have to carefully observe the distribution of your data. If you find the possibility of your data to take parametric statistics, you should check that the distributions are approximately normal.

- The best way to do this is to check the skewness and Kurtosis measures from the frequency output (e.g., by using SPSS software). For a relatively normal distribution:
  Skewness ~ = 1.0
  Kurtosis ~ = 1.0

If a distribution deviates markedly from normality, then you take the risk that the statistic will be inaccurate. The safest thing to do is to use an equivalent non-parametric statistic. As Table 11.2 shows, parametric data has an underlying normal distribution which allows for more conclusions to be drawn as the shape can be mathematically described. Anything else is non-parametric.

Based on Table 11.2, Table 11.3 highlights the several dimensions of data and its analysis methods.

While using any of these two types of tests, you should keep in mind the following:

- Non-parametric tests are more conservative, but less powerful. Often used for small sample sizes (< 30), non-normally distributed data and ordinal or nominal data.
- If your data truly 'do' allow for parametric approaches, you should use a parametric test (more power, more likely to see a true significant result). However, if you are unclear (small sample size, on the fence about normality and so on), it is less likely to wrongly conclude that there is a significant result if you use a non-parametric alternative.
- You should verify that the assumptions are met when using a parametric test. This may include testing for equal variances (usually by the *F*-test).

TABLE 11.2 Choosing a Parametric or Non-parametric Test

	Parametric	*Non-parametric*
Assumed Distribution	Normal	Any
Assumed Variance	Homogeneous	Any
Typical Data	Ratio or interval	Ordinal or Nominal
Data Set Relationships	Independent	Any
Usual Central Measure	Mean	Median
Benefits	Can draw more conclusions	Simplicity; less affected by outliers
TESTS		
Choosing	Choosing Parametric Test	Choosing a Non-parametric Test
Correlation Test	Pearson	Spearman
Independent Measures, 2 Groups	Independent Measures *t*-test	Mann–Whitney Test
Independent measures, > 2 Groups	One-way Independent Measures ANOVA	Kruskal–Wallis Test
Repeated Measures, 2 Conditions	Matched-pair *t*-test	Wilcoxon Test
Repeated Measures > 2 Conditions	One-way, Repeated Measures ANOVA	Friedman's Test

Source: Authors.

TABLE 11.3 Dimensions of Data and Data Analysis Methods

	Experimental Design	
Descriptive		**Inferential**
	Population	
	• Parameter	
	Sample	
	• Random	
	• Bias	
	• Statistic	
	Type	
Variables		**Graphs**
• Maserment Scales		• Bar Graph
▪ Normal		• Histogram
▪ Ordinal		• Box Plot
▪ Interval		• Scatterplot
▪ Ratio		
• Qualitative		
• Quantitative		
• Independent		
• Dependent		
	Measures of	
Central Tendency	**Spread (Dispersion)**	**Shap**
• Mean	• Range	• Skewness
• Median	• Variance	• Kurtosis
• Mode	• Standerd Deviation	
	Tests of	
Association		**Infereance**
• Correlation	Central Limit Theorem	• Chi-Square
• Regression		• *t*-test
▪ Slope		▪ Independent Sample
▪ Y-intercept		▪ Correloted Sample
		• Analisis of Variance

Source: http://bobhall.tamu.edu/FiniteMath/Module8/Introduction.html (accessed on 5 December 2017).

Classification of Data

After the data are collected with the help of a questionnaire, interview schedule or observation and so on, they need to be properly classified, tabulated and presented. This process will help you eliminate the unnecessary details and keep only the relevant part of the whole data that you have collected. The procedure adopted for this purpose is known as the method of classification and tabulation of data.

'Classification' is the process of arranging data in homogeneous groups or classes on the basis of resemblances and common characteristics. Classification is the grouping of related facts into classes. It is the first step in tabulation.

Objectives of Classification

The principal objectives of classifying data are:

- To condense the mass of data in such a manner that similarities and dissimilarities can be easily observed and readily apprehended. Millions of figures can thus be grouped and arranged in a few classes having common features
- To facilitate comparison
- To identify the most significant features of the data at a glance
- To recognise and give prominence to the important information gathered while dropping out the unnecessary elements
- To allow a statistical treatment of the material collected

Methods of Classification of Data

Data having common characteristics are placed in one class and in this way, you can divide the entire data into a number of groups or classes. Classification of data can be done on the basis of either of the following two types:

- Classification on the basis of attributes
- Classification on the basis of class intervals

Classification on the Basis of Attributes

You have seen above that attributes refer to the particular characteristics of the population. These attributes may be descriptive. The chief characteristic of the descriptive attributes is qualitative. They cannot be measured in any numerical terms.

In this type of classification, researchers classify data on the basis of some attributes of quality such as sex, religion, occupation and so on. Such attributes cannot be measured along with a scale. In this type of classification, there is either the presence or an absence of the attributes in a given variable.

Classification on the basis of attributes can be further divided into the following two categories:

- Simple, dichotomy or twofold classification
- Manifold classification

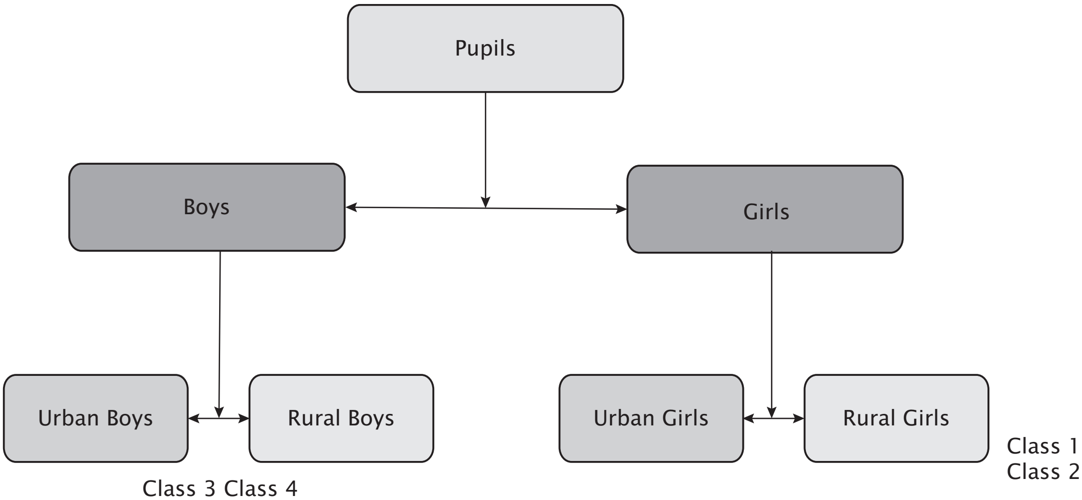


FIGURE 11.1 Manifold Classification

Source: Authors.

In 'simple classification', only one attribute is considered and the universe is divided on that basis. For example, you want to classify pupils' data in a class in a primary school in terms of one attribute, say 'gender'. The only classification that you can do is to divide them as boys and girls. Similarly, you can also classify pupils in the same class of a primary school by rural or urban schools on the basis of another attribute called 'location of school'.

But in the case of 'manifold classification', the universe may be classified into several groups on the basis of more than one attribute. First, we divide the universe (population) into two classes (categories) on the basis of one attribute then each class is further divided into two sub-classes on the basis of second attribute. If we are considering third attribute, each sub-class is further classified into two sub-classes. In this case, we will first divide pupils with respect to gender into boys and girls then each of these classes will be further classified into rural and urban on the basis of location. It means pupils will be classified into four classes as shown in Figure 11.1.

Classification on the Basis of Class Intervals

Unlike descriptive characteristics, the numerical characteristics refer to quantitative phenomena which can be measured through some statistical units such as age, height, weight, prices, production, income, expenditure, sales, profits and so on. When the goal of data classification is to arrange a set of data into a useful form, 'frequency distribution' provides a general approach to data classification.

In frequency distribution, raw data are shown by distinct groups. These groups are termed as 'classes'. The number of observations in each class is called the 'class frequency'. In this way, the data set which may be very large is condensed into a smaller and more manageable set of numbers.

The main methods of such classification are as follows:

Geographical classification: This type of classification is based on geographical or locational differences between various items in the data, for example, such as states, cities, regions, zones and so on.

The average score of primary school graduate for different geographical regions in 2016 may be presented as follows:

Average score of primary school graduates by Geographical regions in country X (Marks scored out of 100 marks)					
Region	*North*	*East*	*South*	*West*	*Central*
Average Score	52	49	74	70	83

Chronological classification: Here data are classified with respect to different periods of time (hour, day, week, month, year and so on). This classification is also known as chronological or temporal classification. For instance, the population of India for different decades may be presented as follows:

Population of India (in millions)					
Year	*1961*	*1971*	*1981*	*1991*	*2001*
Average Output	43.9	54.7	68.5	84.4	102.7

Variable classification: The quantitative phenomenon under study is known as variable and classification on this basis is known as variable classification. Variables are of two kinds:

- Continuous variable: Continuous variables are those that take all the possible values in a given specified range.
- Discrete Variable: Those variables which cannot take all the possible values within a given specified range are termed as discrete (discontinuous) variables.

Classification on the Basis of Discrete Values		*Classification on the Basis of Continuous Values*	
Height	*No. of Students*	*Income (in Baht)*	*No. of Employees*
152	12	1,000–1,500	15
153	13	1,500–2,000	33
154	2	2,000–2,500	22
155	13	2,500–3,000	18
156	10	3,000–3,500	12
TOTAL	50	**TOTAL**	100

When you are classifying your data, there are certain technical terms you need to know. These terms are as follows:

- Class limits: The boundaries of classes are called class limits, for example, 1,000–1,500. In this case, 1,000 employees is the lower limit, whereas 1,500 employees is the upper class limit.
- Class frequency: The number of observations corresponding to a particular class is known as the frequency of the class, for example, the frequency of the above class is 15 employees.
- Magnitude of the class: The difference between the upper class limit and lower class limit of a class is called the size or length of the class, for example, 1,500 − 1,000 = 500 is the length of the class.
- Mid-points: The centre of the limits of a class is called the mid-point or mid-values. Mid-point of a class is defined as:

$$\text{Mid-point} = 1/2\ (\text{lower class limit} + \text{upper class limit})$$

How to Construct Continuous Series?

In continuous series, measurements are only approximations. They are expressed in class intervals, that is, within certain limits. In a continuous frequency distribution, the class intervals theoretically continue from the beginning of the frequency distribution to the end without break.

If you are planning to construct continuous series, you have to follow the steps listed further.

Step 1—Determine the range: You can obtain range by subtracting the size of the smallest item from the size of the largest item of the observed values. For example, if the smallest and the largest values in a class are 350 and 640, the range would be 640 − 350 = 290.

Step 2—Determine the number of classes: As such there are no hard and fast rules to determine the number of classes in any given observations. However, the following factors need to be kept in mind:

- The total frequency (i.e., total number of observations)
- The nature of data (i.e., the size or magnitude)
- The accuracy aimed at
- The ease of computation of various descriptive measures of data analysis

However, you should not forget that number of classes should be neither too small nor too large as it may result very broad and rough classification. Too many frequencies will be concentrated in the class which in turn will obscure some important features and characteristics of the data. This will result in the loss of important information. On the other hand, too many classes will not allow the researcher to handle the data effectively, and data analysis would become cumbersome and tedious. As a principle, statisticians suggest that the number of classes should range between 20 and 5. While determining the number of classes, you may use the following formula.

$$k = 1 + 3.2222 \text{ Log N}$$

Where
k = number of classes and
Log N = Logarithm of observations

Example

Suppose the total number of observations is 15, then the number of classes would be:

$$k = 1 + 3.2222 \text{ Log } 15 = 5 \text{ classes}$$

Determinants of Class Intervals

A possible source of confusion in research is the determination of class interval in continuous series. Statisticians use exclusive and inclusive methods for determining the class intervals in a continuous series.

Exclusive method: within this method, while counting the observations, researchers include the lower limit and exclude the upper limit. For example, suppose that your class interval is 20–40. This class interval will include all the values from 20 to 39 excluding the value of 40 which will be included in the next class interval, that is, 40–60. In this method, the upper class limit of the preceding class will become the lower limit of the succeeding class. The class intervals are formed continuously without any gap or break.

Inclusive class interval: In this method, both the limits are included while counting the observations. For example, the class interval 20–29 will include all the values from 20 to 29, both inclusive. The next value 30 will be included in the next class interval 30–39 and so on. It is important to remember that fractional values between 20 and 29 should not be accounted for in such a classification. This method should be used for those variables whose values are integers.

You can also classify continuous series as follows:

Cumulative frequency series: Cumulative frequency of a given variable or class represents the total frequency of a given variable. There are two types of cumulative frequencies series 'less than' and 'more than'. Both of them can be understood from the following examples:

Example

Less than Cumulative Frequency

Class interval	0–10	10–20	20–30	30–40	40–50
Frequency	4	5	12	9	15

Less than Cumulative Frequency	
UPPER LIMIT	CUMULATIVE FREQUENCY
Less than 10	4
Less than 20	(4 + 5) 9
Less than 30	(4 + 5 + 12) 21
Less than 40	(4 + 5 + 12 + 9) 30
Less than 50	(4 + 5 + 12 + 9 + 15) 45

Similarly, we can also compute more than cumulative frequency series as follows:

More than Cumulative Frequency

Class interval	0–10	10–20	20–30	30–40	40–50
Frequency	4	5	12	9	15

Less than Cumulative Frequency	
UPPER LIMIT	CUMULATIVE FREQUENCY
More than 0	(4 + 5 + 12 + 9 + 15) 45
More than 10	(5 + 12 + 9 + 15) 41
More than 20	(12 + 9 + 15) 36
More than 30	(9 + 15) 24
More than 40	15

Open-end classes series: In this series, the lower limit of the first class interval and the upper limit of the last class interval are not specified.

Example

Open-end Classes Series

Class interval	Less than 10	10–20	20–30	30–40	40–50
Frequency	5	15	60	43	10

Bivariate frequency distribution: When data are classified with reference to two basis or criteria simultaneously, we can call such distributions as bivariate distribution. The following shows the bivariate.

Example

Here we have two variables (a) age of husbands and (b) age of wives. The table also shows that age of husbands take the values from 24 to 28, whereas those of wives these values are 17 to 20.

Age of husbands (Years)	Age of wives (Years)				Total
	17	18	19	20	
24	1	1	–	–	2
25	3	1	1	–	5
26	1	3	2	1	7
27	–	1	2	1	4
28	–	–	1	1	2
Total	5	6	6	3	N = 20

Rules of Classification of Data

While making classification of your data, the following general rules should be followed so that the objectives of a classification may be properly meted out.

- Exhaustive: Your classification should made in an exhaustive manner so that each and every item of the data must belong to any one of the classes without leaving any item to be shown under any class. You should remember that provision for a residual class always signifies a weak point in the classification.

- Exclusive: The classification of data should be made in such a manner that the different classes stand mutually exclusive from each other without leaving any room for overlapping of any class. This means that every item should be confined to a particular class and no item should be included in another class or vice versa.
- Homogeneity: In the classification of data, you should always see that all the data included in a particular class are of homogeneous, or similar nature, and accordingly each of the different classes must include homogeneous data. If a class includes heterogeneous type of data, it will call for another sub-classification of the data included in the same class.
- Flexibility: The manner of classification of data should be such that it should permit changes in time and situation. The old and outdated classes should allow for changes in time and situation. For this, the whole data should be classified into some major classes, and the detailed subdivisions of the classes should be left to be done from time to time taking note of the changes in the situation. In this way, both consistency and flexibility of classification can be sustained without any difficulty.
- Appropriate: The basis of classification decided should be appropriate to the nature of the data otherwise the very purpose of the classification may be defeated. For instance, if a study were designed to determine the economic condition of a section of the people it would be of no use to classify the number of people on the basis of their region. In such a case, you should classify them on the basis of their income or expenditure range.

In summary:

- There should not be any ambiguity in the definition of classes. It will eliminate all doubts while including a particular item in a class.
- All the classes should preferably have equal width or length. Only in some special cases, we use classes of unequal width.
- The class limits (integral or fractional) should be selected in such a way that no value of the item in the raw data coincides with the value of the limit.
- The number of classes should preferably be between 10 and 20, that is, neither too large nor too small. The classes should be exhaustive, that is, each value of the raw data should be included in them.

The classes should be mutually exclusive and non-overlapping, that is, each item of the raw data should fit only in one class.

- The classification must be suitable for the object of inquiry.
- The classification should be flexible and items included in each class must be homogeneous.
- Width of class interval is determined by first fixing the number of class intervals and then dividing the total range by that number.

Tabulation

Both classification and tabulation are methods widely used in research for summarising data, which makes further analysis of data to draw inferences from the data. In this section, we discuss in detail the tabulation method of summarising the data and distinguish between classification and tabulation of data.

We have seen above that the classification of data is the process of separation of data into several classes or groups using properties in the data set. For example, the mathematics test results of a class can be divided into two distinct groups on the basis of gender. Such a classification condenses the raw data into suitable forms and enables statistical analysis and also removes complex data patterns. It highlights the core representatives of the raw data. After classification, comparisons are made and inferences are drawn. Classified data also provide relationships or correlative data patterns.

There are three different ways in which you can present your data:

- Textual presentation: This type of presentation is a descriptive form. Consider the following example.

Example

> 'There were 265 million adult illiterates in Country X in 2010. The adult literacy rate for females for the same year was reported as 55 per cent, whereas this rate for male stood at 73 per cent.
>
> As a renewed EFA commitment of the government, the total number of illiterates decreased to 210 million in 2015. But the adult literacy rate for female increased 65 per cent as against 2 per cent decrease in male literacy rate. Even after 5 years, the gender disparities in adult literacy rates continue to persist in the country'.

Though the textual method of presentation is useful, it has some disadvantages: the method is too lengthy, repetitive of words, difficult to make comparisons and difficult to get a concrete idea to take appropriate action.

- Graphical presentation: Quantitative data may also be presented graphically by using bar charts, pie diagrams, pictographs, line diagrams and so on.
- Tabular presentation: Tabulation means a systematic arrangement of statistical data in columns and rows. It is done for making meaningful analysis, comparison and classification of the data and is designed to summarise lots of information in a simple manner. This section describes the salient features and characteristics of tabulation process with examples. This kind of presentation is a common feature of all research studies, no matter your data are qualitative or quantitative.

Definition of Tabulation

In statistics, tabulation means summarising data using a systematic arrangement of data into rows and columns. It shows the data in concise and attractive form which can be easily comprehended and used to compare numerical figures. Tabulation of data is done with the aim of carrying out investigation, for comparison, identifying errors and omissions in data, studying a prevailing trend, to simplifying the raw data, using the space economically and using it as future reference.

According to Tuttle

> A statistical table is the logical listing of related quantitative data in vertical columns and horizontal rows of numbers, with sufficient explanatory and quantifying words, phrases and statements in the form of titles, headings and footnotes to make clear the full meaning of the data and their origin.[1]

Tabulation is a common way of expressing the associations between two or more variables. Tables are the most common method we use in an investigation for presenting our analysed data. They offer a common means of presenting huge amounts of detailed information in a small space.

[1] See https://homework1.com/statistics-homework-help/tabulation-of-data/ (accessed on 17 November 2017).

A significant advantage of a table is that extensive data can be fitted into it and the precise figures are stored. A disadvantage is that a large table is not illustrative: it seldom reveals more than the most obvious regularities or interdependencies of the data.

You should be clear that data classification and data tabulation refer to two concepts. In classification, we separate and group data on the basis of a property of the data common to all values contained in the table. In tabulation, we organise and arrange data into columns and rows on the basis of characteristics/properties or indicators. The emphasis of tabulation is on the presentation aspects of the data, while classification is used as a means of sorting of data for further analysis.

Objectives of Tabulation

The main objectives of tabulation are as follows. Tabulation:

- Simplifies complex data: In the process of tabulation of data, unnecessary details are avoided and data are presented systematically in columns and rows in a concise form. All tabular data are presented in such a manner that data become more meaningful and can be easily understood by a common man.
- Facilitates comparison: Data presented in rows and columns facilitate comparison. Since a table is divided into various parts and for each part, separate subtotals and totals are given; relationship between various items of the table can be easily comprehended.
- Ensures economies of space: Tabulation ensures the economy of space, as all unnecessary details and repetitions are avoided without sacrificing quality and utility of the data.
- Depicts trend and pattern of data: Tabulation of data depicts the trend of the information under study and reveals the patterns within the figures which cannot be understood in a descriptive form of presentation.
- Helps reference: When data are arranged in tables with titles and table numbers, they can be easily identified and made use of as source reference for future studies.
- Facilitate statistical analysis: After classification and tabulation, statistical data becomes fit for analysis and interpretation. Various statistical measures such as averages, dispersion, correlation and so on can be calculated easily from the data which are systematically tabulated. Also, it can be used for the graphical presentation of data.
- Detects errors: In the process of collection of data, various errors of omission and commission may creep in. These errors cannot be detected easily unless data are properly tabulated. Thus, to detect the errors, if any, with the collected data, tabulation is made as a step of testing the accuracy of the data.

Components of a Table

In general, a statistical table consists of the following eight parts. They are as follows:

Table number

A. Heading
   1. Table Number
   2. Title of the table
   3. Designation of units

B. Body
   1. Stub Head: heading of all rows or blocks of stub items
   2. Body Head: headings of all columns or main captions and their sub captions
   3. Field/body: the cells in rows and columns
C. Notations
   1. Unit of measurement
   2. Footnotes, wherever applicable
   3. Source, wherever applicable

Part I: Table Number

Each table must be given a number. Table number helps in distinguishing one table from other tables. Usually tables should be numbered according to the order of their appearance in a chapter. For example, the first table in the first chapter of a book should be given number 1.1 and second table of the same chapter be given 1.2. Table number should be shown at the top of the table or towards the left of the table.

Part II: Title of the Table

Every table should have a suitable title. It should be short and clear. Title should be such that one can know the nature of the data contained in the table as well as where and when such data were collected. It is either placed just below the table number or at its right.

Part III: Caption

Caption refers to the headings of the columns. It consists of one or more column heads. A caption should be brief, concise and self-explanatory. Column heading is written in the middle of a column in small letters.

Part IV: Stub

Stub refers to the headings of rows. They are at the extreme left of the table.

Part V: Body

The body of the table contains the numerical information. This is the most vital part of the table. It contains a number of cells. Cells are formed due to the intersection of rows and column. Data are entered in these cells.

Part VI: Head Note

The head note should contain the unit of measurement of data. It is a brief explanatory statement applying to all or a major part of the material in the table and is placed below the title and enclosed in brackets.

Part VII: Foot Note

Anything in a table that the reader may find difficult to understand form the title, captions and stubs should be explained in footnotes. If footnotes are needed, they are placed directly below the body of the table. In most cases, footnotes are used to mention the source of data, especially in case of secondary data. Footnotes may be keyed to the title or to any column or to any row heading. It is identified by symbols such as *, +, @, £ and so on.

Part VIII: Source Note

The source note shows the source of the data presented in the table. Reliability and accuracy of data can be tested to some extent from the source note. It shows the name of the author, title, volume, page, publisher's name, year and place of publication of the book or journal from which data are complied.

Figure 11.2 shows the key elements of a table.

Table 11.4 shows an example of a table comprising one variable (global literacy status).

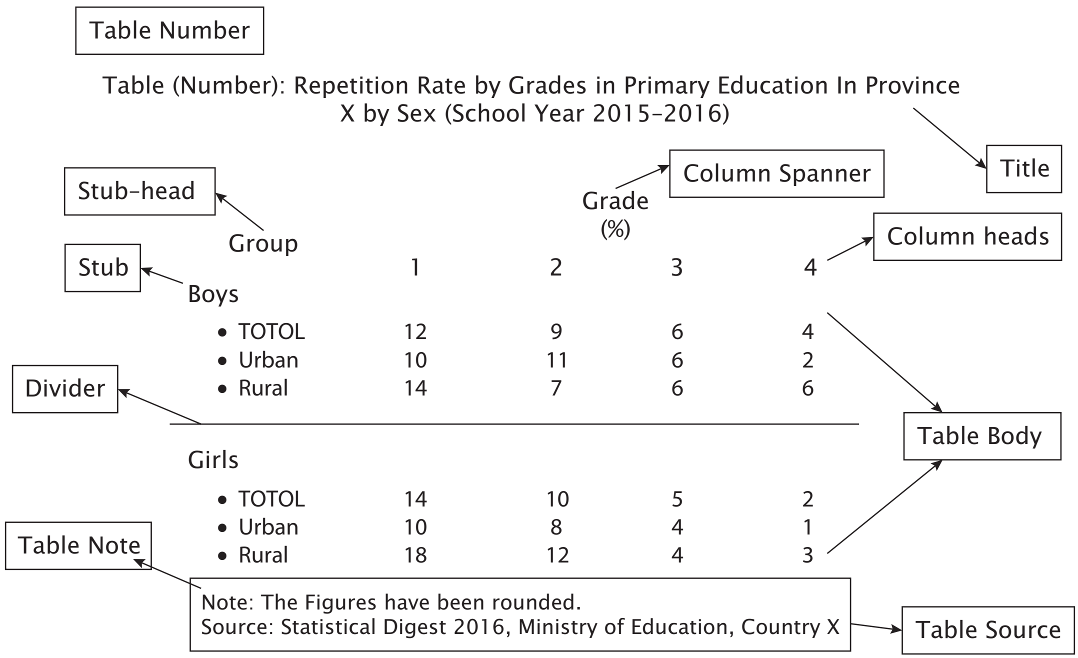


FIGURE 11.2 Elements of a Statistical Table

TABLE 11.4 Global Literacy Rates and Population Numbers for Adults and Youth, 2012

Adult[a] Literacy Rate (Total)	84.3%
Adult Literacy Rate (Male)	88.6%
Adult Literacy Rate (Female)	80.2%
Adult Illiterate Population (Total)	781 million
Adult Illiterate Population (Female Share)	63.5%
Youth[b] Literacy Rate (Total)	89.4%
Youth Literacy Rate (Male)	92.1%
Youth Literacy Rate (Female)	86.9%
Youth Illiterate Population (Total)	126 million
Youth Illiterate Population (Female Share)	61.9%

Source: *UNESCO Institute for Statistics, September 2014.*
Notes: a. Population aged 15 years and older; b. Population between the ages of 15 to 24 years.

In short, you should always keep in mind the following 'principles of table construction' when preparing a table from the raw data that you have collected for your research enquiry.

- Each table should have a title and it should be placed above the body of the table. The title should reflect and represent a succinct description of the contents of the table. The title of the table should be clear and concise.
- Every table should have a number. The number should be centred above the title. The table should run in a consecutive serial number.
- The column heading should be clear and brief. The units of measurements under each heading must always be indicated.
- Any explanatory footnotes concerning the table itself should be placed directly beneath the table.
- If the data in a series of tables have been obtained from different sources, you should indicate the specific sources in a place just below the table.
- Lines are always drawn at the top and the bottom of the table and below the captions.
- Columns may be numbered to facilitate reference. All column figures should be properly aligned.
- Columns and rows that are to be compared to one another should be brought close together.
- Totals of rows should be placed at the extreme right column and totals of columns at the bottom. Different kinds of type spacing and identification can be used to emphasise the relative significance of certain categories.
- Abbreviations and ditto marks should be avoided in a table.
- Sources of data (information) should be clearly indicated at the bottom of the table.

Advantages and Disadvantages of Tabulation

Advantages

- Tables can be useful to show comparison between certain types of data, such as what services each division/department of the ministry of education provides.
- The large mass of confusing data is easily reduced to reasonable form, that is, understandable to kind.
- The data once arranged in a suitable form give the condition of the situation at a glance or gives a bird's eye view.
- From the table, it is easy to draw some reasonable conclusion or inferences.
- Tables give grounds for analysis of the data.
- Errors, and omission, if any, are always detected in tabulation.

Therefore, the importance of a carefully drawn table is vital for the preparation of data for analysis and interpretation.

Disadvantages

One of the major disadvantages of tabular presentation of data is that it does not give a detailed view of the table, unlike textual (descriptive) presentation. Tabular presentation of data may be less effective if trying to show a trend over time, such as how many children entered Grade 1 of primary level over a period of 2010 to 2015.

Types of Tables

Depending upon the number of variables about which information is displayed, tables can be categorised as follows:

- Simple or one-way table: This type of table shows only one characteristic of the data. It is the simplest table which contains data of one characteristic only. The table is easy to construct and simple to follow. For example, tabulation of data on primary level enrolment in country classified by grades (one characteristic) like religion is example of simple tabulation. ~~The~~ Table 11.5 shows the format of a simple or one-way table.

TABLE 11.5 Enrolment by Grade in Primary School X in a Locality, 2015

Grade	*Enrolment*
Grade 1	
Grade 2	
All Grades (Total)	

Source: Authors.

- Two-way table: When the data are tabulated according to two characteristics at a time, it is said to be double tabulation or two-way tabulation. It is a table that contains information on two variables. In such tables, either stub or caption is divided into two coordinate parts. These tables are frequently used for cross tabulation. For example, tabulation of data on the occupation of pupil's parents classified by two characteristics father and mother is double tabulation. This is shown in Table 11.6.

TABLE 11.6 Occupation of Pupils' Parents, 2015

	Parents		
Occupation	*Father*	*Mother*	*Total*
Teacher			
Farmer			
Total			

Source: Authors.

- Multivariate table: This type of table contains information concerning more than two variables. For instance, we can classify primary enrolment by urban–rural category and by sex (Table 11.7).

TABLE 11.7 Students' Parents by Occupation and Sex, 2015

	Primary Enrolment				
	Urban		*Rural*		
Province	***Boys***	***Girls***	***Boys***	***Girls***	***Total***
Northern					
Eastern					
Total					

Source: Authors.

The following example will help you comprehend in a better way the concept of tabulation. The examples represent the following information in a suitable tabular form with proper ruling and headings.

The total number of dropouts in primary education in Province X in Country A in 2010 was 3,500, and it decreased by 300 in 2011 and by 700 in 2012. The total number of dropouts in Province X showed a progressive increase from 2010 to 2012. It was 245 in 2010, 346 in 2011 and 428 in 2012. In Province X, the number of boys who dropped out was 49 in 2010, 77 in 2011 and 108 in 2012. The corresponding figures for girls in Province Y were 2,867, 2,587 and 2,152, respectively, in these three years (Table 11.8).

TABLE 11.8 Tabular Presentation of the Number of Dropouts

	Number of Dropout								
	Pupils (2010)			*Pupils (2011)*			*Pupils (2012)*		
Province	***Boys***	***Girls***	***Total***	***Boys***	***Girls***	***Total***	***Boys***	***Girls***	***Total***
Prov. X	49	196	245	77	269	346	108	320	428
Prov. Y	388	2,867	3,255	267	2,587	2,854	220	2,152	2,372
TOTAL	**437**	**3,063**	**3,500**	**344**	**2,856**	**3,200**	**328**	**2,472**	**2,800**

Source: Authors.

Frequency Distribution Table

A frequency table is a table that lists items and uses tally marks to record and show the number of times they occur. Frequency tables are the normal tabular method of presenting distributions of a single variable.

The statistical data collected are generally raw data or ungrouped data. Let us consider the following example.

Row Data

1	0	3	2	1	5	6	2
2	1	0	3	4	2	1	6
3	2	1	5	3	3	2	4
2	2	3	0	2	1	4	5
3	3	4	4	1	2	4	5

⇩

Frequency Distribution of the Number of Children

Number of Children	Tally Marks	Frequency	Cumulative Frequency
0	ǀǀǀ	3	3
1	卌 ǀǀ	7	10
2	卌 卌	10	20
3	卌 ǀǀǀ	8	28
4	卌 ǀ	6	34
5	ǀǀǀǀ	4	38
6	ǀǀ	2	40
	TOTAL	**40**	

FIGURE 11.3 Raw Data and the Frequency Distribution

Source: Authors.

In a survey of 40 families in a village, the number of children per family was recorded. Figure 11.3 depicts both the raw data obtained and the frequency distribution of the number of children. The figure shows the discrete or ungrouped frequency distribution.

There are two types of frequency distributions: (a) discrete or ungrouped frequency distribution and (b) continuous frequency distribution.

Discrete or Ungrouped Frequency Distribution

In this form of distribution, the frequency refers to discrete value. Here the data are presented in a way that exact measurement of units is clearly indicated.

You will notice that there is definite difference between the variables of different groups of items. Each class is distinct and separate from the other class. Non-continuity from one class to another class exists. Examples of discrete data are number of classrooms in primary schools, number of trained teachers in secondary schools, number of children in a family and so on.

This form of distribution refers to groups of values. There are some basic technical terms when continuous frequency distribution is formed or data are classified according to class intervals. These terms are as follows:

Class limits: The class limits are the lowest and the highest values that can be included in the class. For example, take the class 30–40. The lowest value of the class is 30 and highest class is

40. The two boundaries of class are known as the lower limits and the upper limit of the class. The lower limit of a class is the value below which there can be no item in the class. The upper limit of a class is the value above which there can be no item to that class.
Class interval: The class interval may be defined as the size of each grouping of data. For example, 50–75, 75–100, 100–125… are class intervals. Each grouping begins with the lower limit of a class interval and ends at the lower limit of the next succeeding class interval.
Mid-value or mid-point: The central point of a class interval is called the mid-value or mid-point (Figure 11.3). This can be calculated by adding the upper and lower limits of a class and then dividing the sum by 2:

$$\text{Mid-value} = L + U / 2$$

Where
L = the lower limit of a class interval and
U = the upper limit of a class interval
For example, if the class interval is 20–30, then

$$\text{Mid-value} = 20 + 30 / 2 = 25$$

Continuous Frequency Distribution

Let us consider the frequency distribution of marks in statistics test of 420 Grade 12 students in a secondary school as shown in Table 11.9.

TABLE 11.9 Frequency Distribution (Grouped Data)

Marks in Statistics (Out of 100 marks)	*Number of Students*	*Cumulative Frequency*
30–40	25	25
40–50	53	78
50–60	77	155
60–70	95	250
70–80	80	330
80–90	60	390
90–100	30	420
TOTAL	420	

Source: Authors.

In Table 11.9, the class frequencies are 25, 53, 77, 95, 80, 60 and 30. The total frequency is 420. There are three methods of classifying the data according to class intervals, namely:

- Exclusive method: When the class intervals are so fixed that the upper limit of one class is the lower limit of the next class; it is known as the exclusive method of classification. The data in Table 11.10 are classified on this basis.

TABLE 11.10 Frequency Distribution (Exclusive Method)

Expenditure (Indian Rupees)	*Number of Families in the Community*
0–5,000	60
5,000–10,000	95
10,000–15,000	122
15,000–20,000	83
20,000–25,000	40
TOTAL	400

Source: Authors.

It is clear that the exclusive method ensures continuity of data as much as the upper limit of one class is the lower limit of the next class. In the above example, there are 6o families whose expenditure is between ₹0 and ₹4,999.99. A family whose expenditure is ₹5,000 would be included in the class interval 5,000–10,000. This method is widely used in practice.

- Inclusive method: In this method, the overlapping of the class intervals is avoided. Both the lower and upper limits are included in the class interval. This type of classification may be used for a grouped frequency distribution for discrete variable like number of pupils in Grade 1 in a primary school, number of professors in a university and so on, where the variable may take only integral values. It cannot be used with fractional values such as age, height, weight and so on. This method may be illustrated as shown in Table 11.11.

TABLE 11.11 Frequency Distribution (Inclusive Method)

Class Interval	*Frequency*
5–9	7
10–14	12
15–19	15
20–29	21
30–34	10
35–39	5
TOTAL	70

Source: Authors.

Thus, to decide whether to use the inclusive method or the exclusive method, it is important to determine whether the variable under observation is a continuous or discrete one. In case of continuous variables, the exclusive method must be used. The inclusive method should be used in case of discrete variable.

TABLE 11.12 Frequency Distribution (Open-ended Classes)

Salary Range of Teachers in a Secondary School ($)	*Number of Teachers*
Below 2,000	10
2,000–4,000	8
4,000–6,000	7
6,000–8,000	4
8,000 and above	1
TOTAL	30

Source: Authors.

Open-ended classes: Either a class limit is missing at the lower end of the first class interval or at the upper end of the last class interval or both are not specified (Table 11.12). The necessity of open-ended classes arises in a number of practical situations, particularly relating to economic data when there are few very high values or few very low values which are far apart from the majority of observations.

You have seen that the best representation for your data is a tabular presentation. It is up to you to sort through all the data and find a way to best represent it in tabular form. To accomplish this task, you will need to be familiar with the basic rules for tabular presentation:

- Limit your table to data that are relevant to the hypotheses in the experiment.
- Be certain that your table can stand alone without any explanation.
- Make sure that your table is supplementary to your text and does not replicate it.
- Refer to all tables by numbers in your text, for example, Tables 1, 2, 3....
- Describe or discuss only the table's highlights in your text.
- Always give the units of measurement in the table headings.
- Align decimal places.
- Round numbers as much as possible. Try to round to two decimal places unless more decimals are needed.
- Unless using a specific format style that requires that you place tables separately at the end of the report, place the tables near the text that refers to them.
- Decide on a reasonable amount of data to be represented in each table, not too little, so that the reader does not understand your results, but not too much, so that the reader is overwhelmed and confused.
- Only include the necessary number of tables in your paper; otherwise, it may be redundant or confusing to the reader.
- Do not use tables if you only have two or fewer columns and rows. In such cases, a textual description is enough.
- Organise your tables neatly so that the meaning of the table is obvious at first glance. If the reader spends too much time deciphering your table, then it is too complicated and not efficient.

- Remember that too many rows or columns could make it difficult for the reader to understand the data. You may need to reduce the amount of data or separate the data into additional tables.
- If you have identical columns or rows of data in two or more tables, combine the tables.
- Provide column/row totals or other numerical summaries that can make it easier to understand the data.
- Be consistent with your tabular presentation. Use consistent table, title and heading formats.

In short, although some of the rules are framed which actually guide you in tabulating the data, there cannot be any rigidity about these rules. The table constructed should be such which suits the needs and requirements of the problem under investigation.

Computation of Rates and Ratios

Ratios

In almost all research inquires, researchers often are required to calculate proportions, rates, ratios and percentages. Let us talk about all of them.

When we talk about the speed of a car or an airplane, we measure it in miles per hour. This is called a 'rate' and is a type of ratio. A 'ratio' is a comparison by division of two numbers. It is a way to compare two quantities by using division as in miles per hour where we compare miles and hours.

A ratio can be written in three different ways and all are read as 'the ratio of x to y'.

$$X \text{ to } Y \quad \text{or} \quad X{:}Y \quad \text{or} \quad X/Y$$

Ratios are used frequently for comparison. For example, we may want to know that which class has the highest ratio of boys to girl. In this case, we would need to first calculate the ratios of boys to girl in each class and then compare them. To compare the ratios, we express them as fractions and then cross-multiply.

Examples

Suppose you have ratios of boys to girl for Grade 1 and Grade 2 of primary school as 3:4 and 5:7, respectively. You have to compare these two ratios. In this case, first you express the ratios as fractions. The ratios become 3/4 and 5/7. Then you cross-multiply to get $3 \times 7 = 21$ and $4 \times 5 = 20$. Since $21 > 20$, 3/4 is greater the 5/7; thus, 3:4 is greater than 5:7.

In an educational setting, several types of enrolment ratios are deployed to see the coverage of the education system of a country. Similarly, PTRs are used as crude indicators of assessing the quality of education.

Proportion

A proportion on the other hand is an equation that says that two ratios are equivalent. For instance, if one package of cookie mix results in 20 cookies than that would be the same as to say that two packages will result in 40 cookies.

$$20/1 = 40/2$$

A proportion is read as 'x is to y as z is to w'

$$X/Y = Z/W \text{ where } Y, W \neq 0$$

If one number in a proportion is unknown, you can find that number by solving the proportion.

Example

Suppose that to make 20 pancakes you have to use 2 eggs. How many eggs are needed to make 100 pancakes?

	Eggs	Pancakes
Small amount	2	20
Large amount	x	100

$$\frac{\text{eggs}}{\text{pancakes}} = \frac{\text{eggs}}{\text{pancakes}} \text{ or } \frac{\text{pancakes}}{\text{eggs}} = \frac{\text{pancakes}}{\text{eggs}}$$

If we write the unknown number in the nominator, then we can solve this as any other equation.

$$x/100 = 2/20$$

Multiply both sides with 100.

$$100(x/100) = 100(2/20)$$

$$x = 200/20 \text{ or } x = 10$$

If the unknown number is in the denominator, we can use another method that involves the cross product. The cross product is the product of the numerator of one of the ratios and the denominator of the second ratio. The cross products of a proportion are always equal.

If we again use the example with the cookie mix used above

$$20/1 = 40/2$$
$$1(40) = 2(20) = 40$$

It is said that in a proportion if

$$x/y = z/w \text{ where } y, w \neq 0$$
$$xw = yz$$

If you look at a map, it always tells you in one of the corners that 1" (inch) of the map correspond to a much bigger distance in reality. This is called a 'scaling'. We often use scaling in order to depict various objects. Scaling involves recreating a model of the object and sharing its proportions, but where the size differs. One may scale up (enlarge) or scale down (reduce).

For example, the scale of 1:4 represents a fourth. Thus, any measurement we see in the model would be 1 / 4 of the real measurement. If we wish to calculate the inverse, where we have a 20 ft high wall and wish to reproduce it in the scale of 1:4, we simply calculate:

$$(20)(1):4 = (20)\frac{1}{4} = 5$$

In a scale model of 1:*X* where *X* is a constant, all measurements become 1 / *X* of the real measurement. The same mathematics applies when we wish to enlarge. Depicting something in the scale of 2:1, all measurements then become twice as large as in reality. We divide by 2 when we wish to find the actual measurement.

Rates

In an educational research, the most commonly used rates are simple rates and growth rates. These rates are very important to assess the changes in demographic variables, which affect access to and participation of children in school in future years. Here we present the most important ones.

Calculating (Straight-line) Simple Rates

The average annual rate of change from one period to another period is calculated from the formula:

$$R_n = \left(\frac{\frac{P_n - P_0}{P_0}}{N} \right) \times 100$$

where
R_n = Average Annual Rate of Change in Year n
P_n = Value in Year n
P_0 = Value in Year o
N = Number of years

The average annual rate of change is simply the absolute change between the two periods divided by the number of years elapsed.

Example

In 2005, the population of a district was 250,000. This grew to 280,000 in 2015. What is the average annual rate of change for the district?

R_{2015} = ?
P_{2015} = 280,000
P_{2005} = 250,000

Absolute Change in Population$_{2005-2015}$ = 280,000 − 250,000 = 30,000
Average Annual Absolute Change$_{2005-2015}$ = 30,000 / 10 = 3,000

$$R_n = \left(\frac{\frac{280{,}000 - 250{,}000}{250{,}000}}{10} \right) \times 100$$

$$R_n = \frac{\frac{30{,}000}{250{,}000} \times 100}{10}$$

$$R_n = \frac{12}{10} = 1.2\%$$

The population of district increased 12 per cent between 2005 and 2015 or at an average annual rate of 1.2 per cent.

Calculating Average Annual Growth Rate (Compound Growth Rate)

Another common method of calculating rates of change is the average annual compound growth rate (AAGR). AAGR works the same way that a typical savings account works. Interest is compounded for some period (usually daily or monthly) at a given rate. Here is how it works:

Let us say you had $100 dollars and invested it in an account that paid 5 per cent annually. Table 11.13 shows how your interest accumulates.

TABLE 11.13 Concept of Compound Growth

Year	*0*	*1*	*2*	*3*	*4*	*5*
Rate	+5%	+5%	+5%	+5%	+5%	+5%
Increase	NA	$5	$5	$6	$6	$6
Amount	$100	$105	$110	$116	$122	$128

Source: Authors.

This can also be written in a formula as

$$P_n = P_0\,(1 + r)^n$$

Where
P_n = Present amount
P_0 = Initial amount
r = Rate of increase
n = Number of years

To calculate your future value in the above example, the formula would be:

$$\$100\,(1 + 0.05)^5 = \$128$$

The formula to calculate a 'growth rate' given a beginning and ending value is:

$$r = \sqrt[n]{\frac{P_n}{P_0}} - 1$$

Let us use the district population example to illustrate how this works.

$$r = \sqrt[10]{\frac{280{,}000}{250{,}000}} - 1 = \sqrt[10]{1.12} - 1 = 0.011 \times 100 = 1.1\%$$

You can use excel or SPSS programme to calculate these growth rates.

In addition to these rates, there are other rates such as access, participation rates (intake rates, admission rates, attendance rates), students flow rates (repletion, dropout, promotion, completion, graduation, survival rates), teachers attrition rates and so on which are rigorously used in measuring the efficiency and coverage of an education system.

Percentages

The term percentage or symbol % is used frequently in everyday language and life. For example, it is common to see 48 per cent girls' share in primary enrolment, sales with 20 per cent discount or restaurant bills with 10 per cent service charges, as well as reports in newspapers discussing tax, unemployment and other values in percentage terms. Percentages are used in these ways as a simplified means of conveying size or scale, or value.

Percentage means parts out of 100 and is the same as a fraction with a denominator (bottom) of 100. Therefore,

17 per cent means 17 parts out of 100 and is the same as the fraction 17 / 100

A further way of expressing parts out of 100 is using a decimal and so percentages can also be expressed as decimals:

17 per cent is the same as 0.17 or 17 / 100

There are three main ways in which percentages are frequently used:

- To enable data with different sample sizes or totals to be compared
- To quantify the amount of change over time
- To express an increase or reduction relative to initial size

Using Percentages to Compare Information

Whilst researching for an essay or dissertation, you may come across many sources of data in tables, graphs or reports which you would like to incorporate into your work. However, this can be difficult if they do not share a common baseline. Percentages are useful for comparing information where the sample sizes or totals are different. By converting different data to percentages, you can readily compare them.

For example, the table 11.14 shows enrolment (in million) by levels of education in Country X in 2014 and 2015.

Because the total number of pupils in 2014 and 2015 were different, it is difficult to compare the data for the two years and to determine whether or not there was any notable change in the format in which pupils were enrolled. However, if the enrolment in level of education is expressed as a percentage of the total enrolment, then it is easier to compare the data for the two years.

TABLE 11.14 Enrolment by Levels of Education in 2014 and 2015

Year	*Pre-primary*	*Primary*	*Lower Secondary*	*Higher Secondary*	*Total*
			(in million)		
2014	2.4	159.7	78.3	46.2	286.6
2015	2.5	158.8	87.0	36.6	284.9

Source: Authors.

For example, the conversion from the actual enrolment at primary level in 2014 to a percentage can be done in the following way:

Determine the fraction of primary enrolment in 2014:
159.7 out of a total enrolment of 286.6 pupils
Rewrite this as a fraction: 159.7 / 286.6
Convert the fraction to a decimal by dividing 159.7 by 286.6
159.7 ÷ 286.6 = 0.5572 (this is the decimal format explained above)
Convert the decimal to a percentage by multiplying by 100
0.5572 × 100 = 55.72

The result indicates primary level enrolment accounted for 55.72 per cent total enrolment in 2014. These four stages can be summarised by the following formula:

$$(159.7 / 286.6) \times (100 / 1)$$

This process can be repeated for all levels of enrolment in 2014 and 2015 to produce Table 11.15.

TABLE 11.15 Percentage Share of Enrolment by Levels of Education, 2014–2015

Year	*Pre-primary*	*Primary*	*Lower Secondary*	*Higher Secondary*	*Total*
			(%)		
2014	0.84	55.72	27.32	16.12	100.0
2015	0.88	55.74	30.54	12.84	100.0

Source: Authors.

Converting the secondary level enrolment to percentages makes it easier to compare the enrolment according to the level format. For example in 2014, 16.12 per cent or 16.12 out of every 100 students were in higher secondary schools, but in 2015 this had fallen to 12.84 per cent.

Calculating Percentage Change

Percentages are also very useful if you wish to quantify change. This is because they provide a result in the form of parts per hundred that is usually more readily understandable and comparable than when the information is presented as raw values.

Using the information presented in Table 11.15, the decrease in higher secondary level enrolment between 2014 and 2015 can be calculated as a percentage.

Determine the difference between the 2014 and 2015 enrolment:

46.2 million (2014 enrolment) – 36.6 million (2015 enrolment) = 9.6 million

Express this difference as a fraction of the starting value (the 2014 enrolment of 46.2 million) and multiply the result by 100:

$$(9{,}600{,}000 / 46{,}000{,}000) \times (100 / 1) = 20.78\%$$

This means that there was a fall of 20.78 per cent in enrolment in 2015 as compared to 2014.

When calculating the percentage change between two values, it is important to determine the correct base from which to calculate the percentage change (i.e., the appropriate starting value). This is because the percentage change from a low number to a higher number is not the same as the percentage change from the same higher number to the same lower number. For example,

$$\text{The percentage increase from 50 to } 75 = (25/50) \times (100/1) = 50.0\%$$

However,

$$\text{The percentage decrease from 75 to } 50 = (25/75) \times (100/1) = 33.3\%$$

Calculating Percentage Increase or Decrease

In the previous example, the percentage change between two values was calculated. The reverse process is where the actual amount represented by a particular percentage is to be calculated. Typical examples of the use of percentages in this way are in shop sales, where the prices have been reduced by a certain percentage, or bank charges where interest rates are expressed as a percentage. Again the use of the term percentage expresses the amount in terms of parts per hundred.

Example

A bank interest rate of 6 per cent is the same as saying the interest on each €1.00 invested is 6 parts out of 100 or 6 cents, and a sale reduction of 12 per cent means that an item will be reduced by 12 cents for every €1.00 it originally cost.

There are two ways of calculating the new value of the item following a percentage increase or reduction. For example, you may wish to calculate the sale price of a TV normally priced at €180.00 that has been reduced by 15 per cent.

- The first method is to calculate the amount that the item has been reduced by:

$$15\% \text{ of } €180 = (15/100) \times €180 = €27$$

This amount is then subtracted from the original price:

$$€180 - €27 = €153$$

- The second method is to calculate the percentage of the original price you are now paying for the item:

$$100\% \text{ (full price)} - 15\% \text{ (discount)} = 85\% \text{ (discounted price)}$$

This value is multiplied by the original price:

$$0.85 \times €180 = €153 \text{ (this uses the decimal format explained above)}$$

The result (€153) is the same whichever method you use.

You can use the same methods to calculate the value of percentage increases. For example, if you had savings of €112.72 a year ago that you put in an account earning 4.3 per cent interest, you can calculate the current value of your savings. Again, there are two methods.

Either you calculate the actual amount by which your savings have increased:

$$4.3\% \text{ of } €\,112.72 = (€4.3/100) \times €112.72 = €4.85$$

and add this to the original value of the savings:

€112.72 + €4.85 = €117.57

or calculate the percentage of your original savings you now have:

100% + 4.3% = 104.3%

and multiply this by the original value of your savings:

1.043 × €112.72 = €117.57
(This uses the decimal format explained above)

Again the two methods produce the same result.

Row, Column and Total Percentages

Row percentage: calculated from the total of all the sub-categories of one variable that are displayed along a row in different columns, in relation to only one sub-category of the other variable.

Column percentage: calculated from the total of all the sub-categories of one variable that are displayed in columns in different rows, in relation to only one sub-category of the other variable.

Total percentage: this standardises the magnitude of each cell, that is, it gives the percentage of respondents who are classified in the sub-categories of one variable in relation to the sub-categories of the other variable.

There are three types of percentages: row, column and total as shown in Table 11.16.

TABLE 11.16 Row, Column and Total Percentages

Teachers / Schools	*Public School*	*Col. (%)*	*Private School*	*Col. (%)*	*Aided School*	*Col. (%)*	*Total*	*Col. (%) Total*
Trained Teachers	230	56.8	210	71.2	55	33.3	495	57.2
Row (%)	46.5	–	42.4	–	11.1	–	100.0	–
Untrained Teachers	175	43.2	85	28.8	110	66.7	370	42.8
Row (%)	47.3	–	23.0	–	29.7	–	100.0	–
Total	405	100.0	295	100.0	165	100.0	865	100.0
Row (%)	46.8	–	34.1	–	19.1	–	100.0	–

Source: Authors.

Let us consider first the row percentages. In Table 11.16, there are 495 trained teachers of which 46.5 per cent, 42.4 per cent and 11.1 per cent are in public, private and aided schools, respectively. Similarly, of the total 370 untrained teachers, 47.3 per cent, 23.0 per cent and 29.7 per cent are in public, private and aided schools, respectively. Now let us see the column percentages. Of the total 405 teachers in public schools, 56.8 per cent are trained and 43.2 per cent are untrained. These percentages in private school are 71.2 per cent and 28.8 per cent, respectively.

Thus, percentages are related to fractions and decimals, and express values in the common currency of parts per hundred. Expressing values in the form of percentages in your work will enable you to readily compare information from different sources, quantify change over time and find the amount by which something has increased or decreased following a percentage change. Whilst you will come across data that are already presented as percentages, remember that you can also convert data from tables and graphs that you find in books, journals and other publications into percentages in order to make useful comparisons for your research studies.

The ability to interpret data accurately and to communicate findings effectively are important skills for a researcher. For accurate and effective interpretation of data, you may need to calculate other measures such as percentages, cumulative percentages, ratios or rates. It is also sometimes important to apply other statistical procedures to data. The use of percentages is a common procedure in the interpretation of data.

Summary

This chapter showed that classification of data is the way in which a researcher arranges his or her raw data in different classes in order to give a definite form and a coherent structure to the data collected, facilitating their use in the most systematic and effective manner. The process of classification involves grouping the statistical data under various understandable homogeneous groups for the purpose of convenient interpretation. A uniformity of attributes is the basic criterion for classification and the grouping of data is made according to similarity. Classification becomes vital and necessary when there is diversity in the data collected for meaningful presentation and analysis. However, in respect of homogeneous presentation of data, classification may be unnecessary.

We have also noticed that researchers arrange the classified in tabular forms (tables) in columns and rows. Tabulation is the simplest way of arranging the data, so that anybody can understand it in the easiest way. It is the most systematic way of presenting numerical data in an easily understandable form. It facilitates a clear and simple presentation of the data, a clear expression of the implication and an easier and more convenient comparison. There can be simple or complex tables and general purpose or summary tables.

Classification and tabulation are interdependent events in a research. First, the data are classified and then they are presented in tables, the classification and tabulation in fact goes together. So classification is the basis for tabulation. Tabulation is a mechanical function of classification because in tabulation classified data are placed in row and columns. Classification is a process of statistical analysis, whereas tabulation is a process of presenting the data in suitable form.

Clearly, statistics are a tool, not an aim. Simple inspection of data, without statistical treatment, by an experienced and dedicated analyst may be just as useful as statistical figures on the desk of the disinterested. The value of statistics lies with organising and simplifying data to permit some objective estimate showing that an analysis is under control or that a change has occurred. Equally important is that the results of these statistical procedures are recorded and can be retrieved. Chapter 11 attempts to provide simple statistical methods to analyse the tabular data and more importantly to answer your research question(s) and the validity of your research hypothesis.

Self-test Exercise

Exercise 11.1:

11.1.1. Discrete variables and continuous variables are two types of:

A. Open-end classification
B. Time series classification
C. Qualitative classification
D. Quantitative classification

11.1.2. In a set of observations, amount of variation can be shown in form of figures with help of:

A. Absolute measures
B. Uniform measures
C. Non-uniform measures
D. Exploratory measures

11.1.3. Numerical methods and graphical methods are specialised procedures used in:

A. Social statistics
B. Business statistics
C. DS
D. Education statistics

Exercise 11.2:

11.2.1. Describe briefly the difference between a variable and an attribute. Identify three variables and their possible corresponding attributes commonly used in educational research.

Exercise 11.3: Study the following table showing the annual expenditure (in thousand US$) of the faculty of education of University X during the period 2011–2015 and answer the following questions.

Expenditure of Faculty of Education (in thousand US$) of University X by Major Items (2011–2015)

Year	*Salary*	*Subsistence Allowance*	*Transport*	*Purchase of Books and Publications*	*Student Vouchers*	*Total Expenditure*
2011	288.00	98.00	3.00	23.40	83.00	495.40
2012	342.00	112.00	2.52	32.50	108.00	597.02
2013	324.00	101.00	3.84	41.60	74.00	544.44
2014	336.00	133.00	3.68	36.40	88.00	597.08
2015	420.00	142.00	3.96	49.40	98.00	713.36

11.3.1. What is the average expenditure on books and publications per year which the company had to pay during this period?

A. US$ = 32.43 thousand
B. US$ = 33.72 thousand

C. US$ = 34.18 thousand
D. US$ = 36.66 thousand

Answer: Option D
Hint
Average expenditure on books and publications during the period

$$\bar{X} = \left[\frac{23.4 + 32.5 + 41.6 + 36.4 + 49.4}{5}\right]$$

11.3.2. The total expenditure on transport during the given period is approximately what per cent of the total amount of salary paid during this period?
A. 0.10 per cent
B. 0.50 per cent
C. 1.00 per cent
D. 1.25 per cent

Answer: Option C (approximately 1%)
Hint

$$\text{Required Percentage} = \left[\frac{3.00 + 2.52 + 3.84 + 3.68 + 3.96}{288 + 342 + 324 + 336 + 420}\right] \times 100$$

11.3.3. Total expenditure on all these items in 2022 was approximately what per cent of the total expenditure in 2015?
A. 62 per cent
B. 66 per cent
C. 69 per cent
D. 71 per cent

Answer: Option C (approximately 69.45%)
Hint

$$\text{Required percentage} = \left[\frac{288.00 + 98.00 + 3.00 + 23.40 + 83.00}{420.00 + 142 + 3.96 + 49.40 + 98.00}\right] \times 100$$

11.3.4. The ratio between the total expenditure on student voucher and subsistence allowance for all the years and the total expenditure on subsistence allowance for all the years, respectively, is approximately?
A. 4:70
B. 10:13
C. 5:18
D. 5:80

Answer: Option B (10.13)
Hint

$$\text{Required sratio} = \left[\frac{83.00 + 108.00 + 74 + 88 + 98}{98.00 + 112.00 + 101.00 + 133.00 + 142.00}\right]$$

11.3.5. Calculate the average annual growth rate in total expenditure during the period 2011–2015.

Answer:

Hint

$$\text{Average Annual Growth Rate} = \sqrt[n]{\frac{P_n}{P_0}} - 1$$

$$\text{Average Annual Growth Rate} = \sqrt[5]{\frac{713.36}{495.40}} - 1$$

References

Hatch, J.A. 2002. *Doing Qualitative Research in Education Settings*. Albany, NY: SUNY Press.

Marshall, C. and G.B. Rossman. 2006. *Designing Qualitative Research*. Thousand Oaks, CA: SAGE Publications.

Seidel, J.V. 1998. *Qualitative Data Analysis*. Ethnograph V5, Appendix E. Available at http://bobhall.tamu.edu/FiniteMath/Module8/Introduction.html (accessed on 17 November 2017).

12 CHAPTER

Data Analysis: Descriptive Methods and Index Numbers

Introduction

We have seen that a statistic is a numerical representation of information. Whenever our aim is to quantify or apply numbers to data to organise, summarise or better understand the information, we use statistical methods. These methods can span from somewhat simple computations such as determining the mean of a distribution to highly complex computations such as determining factors or interaction effects within a complex data set.

There are two major types of statistics. The branch of statistics dealing with the summarisation and description of data is called 'descriptive statistics' and the branch of statistics that covers and uses sample data to make an inference about a population is referred to as 'inferential statistics'. When a researcher analyses data, such as the marks achieved by 100 students for a piece of coursework, there is a possibility that he or she can use both descriptive and inferential statistics in his or her analysis of students' marks. Typically, in most research conducted on groups of people, the researcher will use both descriptive and inferential statistics to analyse his or her results and draw inferences or conclusions. So what do we mean by descriptive and inferential statistics and what are their differences? Descriptive and inferential statistics each provides different insights into the nature of the data gathered. One alone cannot give the whole picture. Together, they give a powerful tool for both description and prediction. For example, we can use both types of statistics when we analyse data on the marks achieved by 50 students in English Language test in our analysis of their marks. Typically, in most research conducted on group of people, researchers can use both descriptive and inferential statistics to analyse their results and draw conclusions.

This chapter presents an overview of statistical methods in order to better understand research results. The chapter presents important formulas for computations together with their usefulness and domains of application in your research analysis. The inferential methods of data analysis are presented in Chapter 13.

Meaning of Descriptive Methods

Before we define descriptive methods, it should be remembered that descriptive and inferential statistics are interrelated. It is almost necessary to use methods of descriptive statistics to organise and summarise the information obtained from a sample before methods of inferential statistics can be applied to make more detailed and thorough analysis of the topic of research under investigation. Furthermore, the preliminary descriptive analysis of a sample often reveals the main features offering clear insight and indication to the choice of the appropriate inferential method to be subsequently used.

Often the researchers collect data from the whole population. In that case, he or she can possibly perform a descriptive study both on the population as well as on the sample. The study becomes inferential only when an inference is made about the population based on information obtained from the sample.

In any research project, we collect a lot of quantitative and qualitative data. If we present these raw data as such, it would be very difficult to visualise and describe what the data are showing or telling. Descriptive statistics provide information that describes the data in some manner. For example, suppose a kiosk sells books, magazines, newspapers and novels. If a total of 100 items are sold and 40 out of the 100 were books, then one description of the data on the items sold would be that 40 per cent were books. The same kiosk may conduct a study on the number of newspapers sold each day for one month and determine that an average of 30 newspapers were sold each day. This average is an example of descriptive statistics.

Descriptive analysis is the term given to the analysis of data that helps describe, show or summarise data in a meaningful way such that, for example, patterns might emerge from the data. Descriptive statistics do not, however, permit the researcher to draw conclusions beyond the data he or she has analysed or reach conclusions regarding any hypotheses he might have made. Descriptive research attempts to show what already exists in a group or population. An example of this type of research would be an opinion poll to ascertain or determine which presidential candidate people plan to vote for in the next election. 'Descriptive studies do not measure the effect of a variable; they seek only to describe'.

> 'Descriptive statistics consist of methods for organizing and summarizing information' (Weiss 1999).

Descriptive statistics provides a quick method to make comparisons between different data sets and to spot the smallest and largest values and trends or changes over a period of time. If a bookseller wanted to know what types of books were purchased most by students during the school year, a graph might be a good medium to compare the number of each type of book sold and the months of the school year.

There are some other measurements in descriptive statistics which help the researcher to answer questions such as 'How widely dispersed are these data?' 'Are many of the values the same?' 'What value is in the middle of data?' 'Where does a particular data value stand with respect to other values in the data set?' A researcher has collected data on the number of out-of-school children in 40 districts of Northern Province in Country X. He or she is interested in studying the distribution or spread of the out-of-school children. Descriptive statistics allow him or her to do this.

A graphical presentation of data is another technique or method of descriptive statistics. Examples of this visual representation are histograms, bar graphs and pie graphs, to name a few. The graphical presentation of data is discussed in Chapter 13. Using these methods, the data are described by compiling them into a graph, table or other visual representation.

Descriptive statistics comprise the construction of graphs, charts and tables, and the computation of various descriptive measures such as averages, measures of variation and percentiles. Table 12.1 shows the several methods commonly used in descriptive statistics.

TABLE 12.1 Methods of Descriptive Statistics Analysis

<table>
<tr><th></th><th>Nominal Scale</th><th>Ordinal Scale</th><th>Interval Scale</th><th>Ratio Scale</th></tr>
<tr><td>Methods of Presenting Data</td><td colspan="4">– Tabulation; Graphical Presentation –</td></tr>
<tr><td rowspan="3">Averages</td><td colspan="4">– The Mode –</td></tr>
<tr><td>–</td><td colspan="3">–The Median–</td></tr>
<tr><td>–</td><td>–</td><td colspan="2">– The Arithmetic Mean –</td></tr>
<tr><td rowspan="3">Measures of Dispersion</td><td>–</td><td colspan="3">– The Quartile Deviation –</td></tr>
<tr><td>–</td><td colspan="3">– The Range –</td></tr>
<tr><td>–</td><td>–</td><td colspan="2">– Standard Deviation –</td></tr>
</table>

Source: Authors.

Descriptive statistics are used to describe and explain sets of numbers such as test scores. Data are organised in sets of scores into tables and graphs termed as frequency distributions.

For example, the following numbers represent students' scores on a reading test: 18, 22, 20, 18, 30, 24, 18, 20, 22, 24, 26, 20 and 28.

Table 12.2 shows a frequency table comprising the distribution or number of students who achieved a particular score on the reading test. In this example, three students scored 18.

TABLE 12.2 Frequency Distribution of Test Scores

Reading Score	*Frequency*	*Per cent*	*Percentile (Cumulative)*
18	3	21.3	21.3
20	3	21.3	42.6
22	3	21.3	63.9
24	2	14.2	78.1
26	1	7.3	85.4
28	1	7.3	92.7
30	1	7.3	100.0
TOTAL	14	100.0	–

Source: Authors.

Based on the data given in Table 12.2, we can show the same data in the form of a figure. This is presented in Figure 12.1 which shows the distribution or number of students who achieved a particular score.

Plotting data in a frequency distribution highlights the general shape of the distribution and gives a general sense of how the numbers are bunched. You can use several statistics to represent the 'centre' of the distribution. These statistics are commonly termed as 'measures of central tendency'.

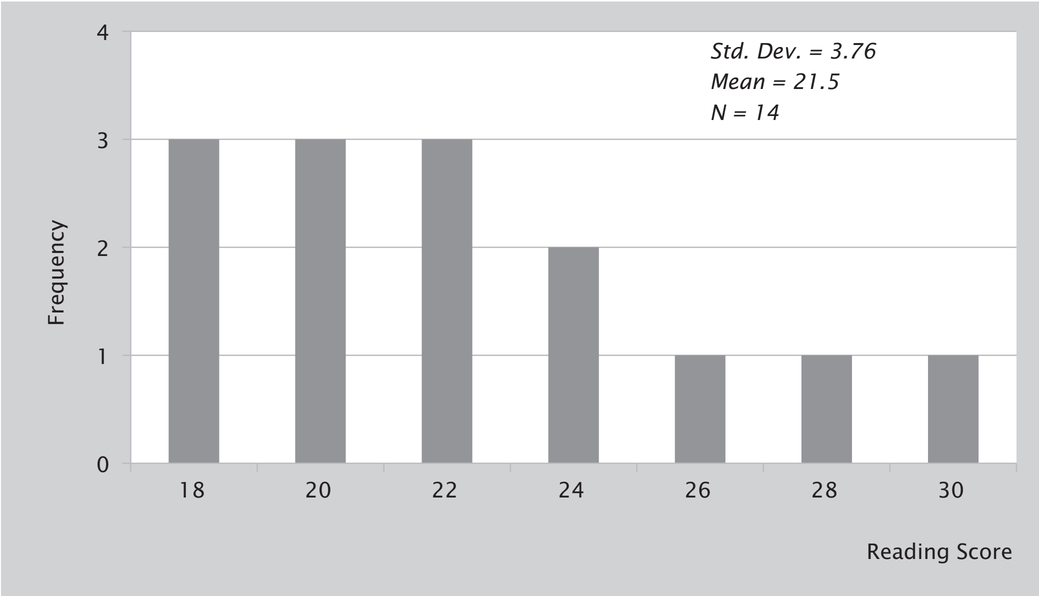


FIGURE 12.1 Frequency Distribution of Test Scores

Source: Authors.

Measures of Central Tendency

Tables and graphical presentations give a general description of data. However, the population that we study in our research comprises large number of observations. If we look at all the observations, we cannot grasp any idea about all the observations. Thus, we have to look at or find out a single number that must be good representative one for all the observations so that we can have a clear picture of that characteristic. Such representative number represents a central value of all observations and is called a 'measure of central tendency' or an 'average' or a 'measure of location'. There are three averages called mean, median and mode which are widely used in research. In addition to these measures, researchers also use quartiles and percentile in data analysis.

> A measure of central tendency is a typical value around which other figures congregate. An average stands for the whole group of which it forms a part yet represents the whole. One of the most widely used sets of summary figures is known as measures of location.

Arithmetic Mean or Average

The mean is the most common measure of central tendency and the one that can be mathematically manipulated.

> 'The arithmetic mean is the balance point in a distribution such that if you subtract each value in the distribution from the mean and sum all of these deviation score, the result will be zero.

The arithmetic mean is a mathematical representation of the typical value of a series of numbers, computed as the sum of all the numbers in the series divided by the count of all numbers in the series'.[1]

The mean of a variable is defined as the sum of observations divided by the number of total observations. If the variable X assumes N values $X_1, X_2, ..., X_n$, then the mean, $\overline{X}$, is given by the following.

Arithmetic Mean for Ungrouped or Raw Data or Individual Observation

$$\overline{X} = X_1 + X_2 + X_3 + \cdots + X_n / N = \sum X / N$$

Example

A student's test scores (out of 10 marks) in five subjects are 2, 4, 6, 8 and 10. The mean score ($\overline{X}$) is then

$$\overline{X} = 2 + 4 + 6 + 8 + 10 / 5 = \sum 30 / 5 = 6 \text{ marks.}$$

Mean for Grouped Data: Direct Method

The mean (or average) of observations, as we know, is the sum of the values of all the observations divided by the total number of observations. If $x_1, x_2, \ldots, x_n$ are observations with respective frequencies $f_1, f_2, \ldots, f_n$, then this means observation x_1 occurs f_1 times, x_2 occurs f_2 times and so on.

Now the sum of the values of all the observations $= f_1 \times 1 + f_2 \times 2 + \cdots + f_n \times n$, and the number of observations $= f_1 + f_2 + \cdots + f_n$.

So the mean ($\overline{X}$) of the data is given by

$$\overline{X} = \frac{f_1 x_1 + f_2 x_2 + \cdots f_n x_n}{f_1 + f_2 + \cdots f_n},$$

which can be written more briefly as follows:

$$\overline{X} = f_i x_i / \sum f_i$$

Example

The following table shows the marks obtained by 30 students of Class 11 of a certain secondary school in a statistics test consisting of 100 marks.

Marks obtained (x_i)	Number of Students (f_i)	$f_i x_i$
10	1	10
20	1	20
36	3	108
40	4	160
50	3	150
56	2	112

[1] See http://www.investopedia.com/terms/a/arithmeticmean.asp (accessed on 17 November 2017).

Marks obtained (x_i)	Number of Students (f_i)	$f_i x_i$
60	4	240
70	4	280
72	1	72
80	1	80
88	2	176
92	3	276
95	1	95
TOTAL	$\sum f_i = 30$	$\sum f_i x_i = 1{,}779$

The mean marks obtained by the students can be calculated as follows:

$$\bar{X} = f_i x_i / \sum f_i = 1{,}779 / 30 = 59.3 \text{ Marks}$$

Therefore, the mean marks obtained are 59.3 marks.

Mean for Grouped Data: Assumed Mean Method

Sometimes when the numerical values of x_i and f_i are large, finding the product of x_i and f_i becomes monotonous and time-consuming. In cases such as this, you can think of using a method of reducing these calculations. You can calculate the mean by using the following formula:

$$\bar{X} = A + \sum fd' / \sum f \times C,$$

$$d' = X - A / C,$$

Where
A = Assumed mean or any value in x;
$\sum$f = Total frequencies and
C = CI.

Frequency Distribution of Test Scores

Marks X	*f*	*Mid-value*	*(d' = X − A)/C*	*fd'*
0–10	6	5	–3	–18
10–20	8	15	–2	–16
20–30	10	25	–1	–10
30–40	12	35	0	0
40–50	7	45	1	7
50–60	4	55	2	8
60–70	3	65	3	9
	50			–20

$$Mean\ \bar{X} = A + \sum fd / \sum f \times C = 35 + [-20/50] \times 10 = 35 - 4 = 31 \text{ marks.}$$

This tells us that, on an average, the number of marks obtained by these 50 students in statistics test is 31 marks.

Advantages and Limitations of Arithmetic Mean

Advantages

An important and perhaps the biggest benefit of using the arithmetic mean, as a statistical measure, is its simplicity. Anyone capable of computing simple addition followed by division can calculate the arithmetic mean of a data set. Of all the measures of central tendency, the arithmetic mean or average is least affected by fluctuations when multiple data sets are extracted from a larger population.

Other advantages of arithmetic mean are as follows:

- Easy to understand and calculate
- Rigidly defined
- Single value represents the whole data set
- Computation considers all the values in the series
- Suitable for algebraic treatment
- Least affected by sampling fluctuations
- Useful in further statistical analysis, for example, useful in the computation of SD, correlation, coefficient of skewness and so on

Limitations

In data sets that are skewed or where outliers are present, computation of the arithmetic mean often provides a misleading result. Assume a situation where 10 people are sitting at a restaurant table. Nine of them are teachers earning annual incomes of US$4,500, while the 10th is a businessman, who earns US$100,000 per year. The arithmetic mean of their annual incomes is US$14,050. You can safely say that this figure is not what the typical person at the table earns.

In case your data sets are not normally distributed as represented by the bell curve, it is better that you compare the arithmetic mean with other statistical measures, such as the median. In the earlier example, the median income at the table—the income at which half the people studied are above it and half are below it—is US$4,500. This figure is better as it represents the group as a whole than the arithmetic mean does. When there is a noticeable difference between the values of median and mean, as they are in this example, it shows that the data are skewed in the direction of the mean. In short, the other important limitations of the arithmetic means are as follows:

- Seriously affected by the presence of a few extremely large or small values of the variable
- Cannot be calculated in case a single item is missing
- Cannot be calculated unless some assumptions are made regarding the sizes of the classes in a grouped frequency distribution
- No significance of its own
- May give misleading conclusions for non-homogenous data

Weighted Arithmetic Mean

If all the items in a data are not of equal importance, we may compute a weighted arithmetic mean, where the values are weighted according to their importance. A weighted arithmetic mean is a more accurate measurement of scores that are of relative importance to each other. This is what happens with investment portfolios, grade scoring and other statistics. For example, in order to have an idea of the change in cost of living of a certain group of people, the simple average of the prices of the commodities consumed by them will not do and will not serve the purpose because all the commodities are not equally important, for example, rice, wheat and pulses are more important than tea, chocolate and so on. In cases such as these, you should calculate the weighted arithmetic average which will help you in finding out the average value of the series after giving proper weight to each group.

> 'The weighted average is an average resulting from the multiplication of each component by a factor reflecting its importance' (*Oxford Dictionary*).
>
> 'The weighted average is a mean calculated by giving values in a data set more influence according to some attribute of the data. It is an average in which each quantity to be averaged is assigned a weight, and these weightings determine the relative importance of each quantity on the average'.[2]

Weighted arithmetic mean is used to:

- Compare the results of two or more universities where number of students differ
- Construct weighted index numbers
- Compute standardised death and birth rates

You can compute the weighted average by

$$\text{Weighted Mean} = \bar{X}w = \frac{W_1X_1 + W_2X_2 + \cdots + W_nX_n}{W_1 + W_2 + \cdots + W_n} = \frac{\sum W_iX_i}{\sum W_i}$$

where
W = weight and
X = score.

Example

Weighted Test Scores by Subjects

Subject	Scores (X)	Weight (W)	Weighted Scores ($X \times W$)
English Language	55	2	110
Mathematics	40	3	120
Science	65	3	195
Social Studies	75	1	75
Arts and Drawing	70	1	70
	$\Sigma X = 305$	$\Sigma W = 10$	$\Sigma WX = 570$

$$\text{Weighted Mean} = 570 / 10 = 57 \text{ Marks}$$

[2] See www.mathwords.com/w/weighted_average.htm (accessed on 15 July 2017).

Let us consider another example.

Example

If you are trying to figure out a grade in a subject during the academic year, you should identify what you were graded on each exam. This is shown further.

Weighted Test Scores by Subjects

Exam	Grade (X) (%)	Weight (W) (%)	$X \times W$
First Term	90% = 0.90	25% = 0.25	0.225
Second Term	75% = 0.75	50% = 0.50	0.375
Third Term	87% = 0.87	25% = 0.25	0.2175

$$\text{Weighted Mean} = 0.225 + 0.375 + 0.2175 = 0.8175 = 81.75\%$$

You can find out the grade you need to receive on a test by plugging a variable in the weighted averages formula. For example, if you need to find out what grade you need to receive an 80 per cent grade in our example before, write 0.9(0.25) + 0.75(0.50) + X(0.25) = 0.80. Solve for X. You would need an 80 per cent on the test to get an 80 per cent in the class.

You should remember that the weights should always total 100 if you are using percentages. You should also remember that a weighted average is not the same as the mean. If you compute the mean of 90, 75 and 87 per cent scores, you would arrive at an answer of 84 per cent, an incorrect answer when weights of 25, 50 and 25 per cent are to be factored in. The answer should be 81.75 per cent.

Harmonic Mean

The harmonic mean of a set of observations is the reciprocal of the arithmetic mean of the reciprocal of the observations. Harmonic mean is defined only for non-zero positive values and is used for averaging while keeping one variable constant:

$$\text{H.M.} = \frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} + \cdots + \frac{1}{a_n}}$$

where

n is the number of values in the set of numbers, and a_1, a_2, a_3, … are the values in the set.

Example

A typist is given three tests. In the first test, he or she types 400 words in 50 minutes, in second test, he or she types the same words (400) in 40 minutes and in the third test, he or she takes 30 minutes to type 400 words. What is his or her average speed for the three tests?

$$\text{H.M.} = \frac{3}{\frac{1}{50} + \frac{1}{40} + \frac{1}{30}} = \frac{3}{0.078} = 38.3 \text{ minutes}$$

In this case, the calculation of the harmonic mean is most appropriate because the set of numbers contains outliers that might skew the result.

Geometric Mean

For inflation, price escalations, rates of return and population growth, the geometric mean is the appropriate single point representing an average across time periods. In an education setting, the geometric mean is usually used to calculate the growth rate of school-going children and the rate of return to education across several time periods.

Geometric mean is defined as the *n*th root of the product of *n* numbers, and the formula for its calculation is as follows:

Geometric Mean = *n*th root of (X_1) (X_2) ... (X_n)

or

$$\text{G.M.} = \sqrt[n]{(X_1)(X_2)(X_3)\cdots(X_n)}$$

or

$$\text{G.M.} = (X_1 \times X_2 \times \cdots \times X_n)^{1/n}$$

where X_1, X_2 and so on show the individual data points, and n is the total number of data points used in the calculation.

Example

Consider this example. Suppose you wanted to calculate the geometric mean of the numbers 5, 3 and 12.

You can solve this simple example without the help of even a simple calculator—just in your head. First, take the product; 5 times 3 times 12 is 180. Since there are three numbers, the *n*th root is third root (or cubic root), and the third root of 180 is 5.65. Therefore, the geometric mean of 5, 3 and 12 is 5.65.

$$\text{G.M.} = (5\times3\times12)^{\frac{1}{3}} = \sqrt[3]{180} = 5.65$$

Median

One of the serious problems with using the mean is that it often does not show the typical outcome. We have noticed earlier that if there is one outcome that is very far from the rest of the data, then the mean will be strongly affected by this outcome. Such an outcome is called an 'outlier'. An alternative measure is the 'median'. The median is the middle score. If you have an even number of events, you just take the average of the two middle events. The median is best suited for describing the typical value. In research studies, median is often used for variables such as income and prices.

> 'The median is the value separating the higher half of a data sample, a population, or a probability distribution, from the lower half. In simple terms, it may be thought of as the "middle" value of a data set'.[3]

Median for Ungrouped or Raw Data or Individual Observations

The given values are ordered in terms of magnitudes—either in ascending or descending order. If the numbers of values are 'odd', median is the middle value. If the numbers of values are 'even', median is the mean of middle two values.

[3] See http://www.investopedia.com/terms/p/p-value.asp (accessed on 17 November 2017).

When Observations are in Odd Numbers

$$\text{Median} = \text{Md} = \left\{\frac{(n+1)}{2}\right\}\text{th item},$$

where

Md = Median and

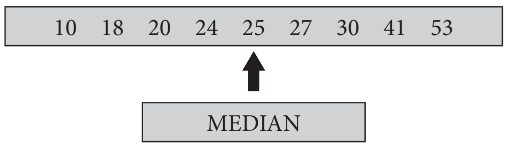
n = number of items.

Example

The test scores of 9 students out of 100 marks in statistics are as follows.

Test scores of 9 students = 18, 25, 27, 10, 53, 30, 20, 41 and 24. First, arrange the scores either in ascending or descending order as follows:

10, 18, 20, 24, 25, 27, 30, 41, 53



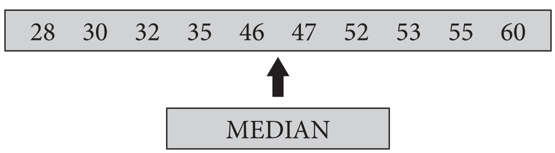
$$\text{Median} = \text{Md} = \left\{\frac{(9+1)}{2}\right\}\text{th item} = 5\text{th item} = 25 \text{ marks.}$$

When Observations Are in Even Numbers

Test scores of 10 students: 8, 18, 25, 27, 10, 53, 30, 20, 41 and 24.

$$\text{Median} = \left(\frac{n+1}{2}\right)\text{th item} = \left(\frac{10+1}{2}\right)\text{the item} = 5.5\text{th item}$$

$$= \frac{\text{Value of 5th item} + \text{Value of 6th item}}{2}$$



$$\text{Md} = 46 + 47 / 2 = 46.5 \text{ marks.}$$

Median for Grouped Data

Values in a grouped distribution are associated with frequencies. Grouped data can assume two forms. These can be a discrete frequency distribution or a continuous frequency distribution. No matter what type of distribution you have, you have to calculate cumulative frequencies so as to know the total number of items.

Discrete Series

Example

The following table shows the number of books in the library of 12 senior secondary schools in a district. How can we compute the median size of library of a secondary school in this district?

Distribution of Secondary School Libraries by Number of Books in District X in 2015

Number of Books (*X*)	Number of Libraries in Secondary Schools (*f*)	Cumulative Frequency (*cf*)
50	10	10
100	30	40
150	50	90
200	60	150
250	70	220
300	100	320
350	90	410
400	50	460
450	30	490
500	20	510
550	20	530
600	10	540
TOTAL	540	

$$\text{Median} = \text{size of } (N + 1/2)\text{th item}$$

$$= \text{size of } (540 + 1/2)\text{th item}$$

$$= 270.5\text{th item}$$

The cumulative frequency just greater than 270.5 is 320 and the value of X (number of books in the library) corresponding to 320 is 300. Hence, the median size is 300 books per library.

Continuous Series

The following table gives the frequency distribution of 325 students of a secondary school according to their monthly stipend in 2015. How can we compute the median level of stipend?

Distribution of Secondary School Students by Monthly Stipend: 2015

Stipend Group (US$)	*Number of Students (Frequency)*	*Cumulative Frequency (cf)*
Below 10	1	1
10–15	20	21
15–20	42	63

Stipend Group (US$)	Number of Students (Frequency)	Cumulative Frequency (cf)
20–25	55	118
25–30	62	180
30–35	45	225
35–40	30	255
40–45	25	280
45–50	15	295
50–55	18	313
55–60	10	323
60 and above	2	325
Total	325	

First, we have to calculate the cumulative frequencies. Next we have to find out $(N/2)$ which in this case is $(325/2)$, that is, 162.5. See in the cumulative frequency the value greater than 162.5. This class then becomes our median class. Then apply the following formula:

$$\text{Median} = L + \frac{\frac{N}{2} - M}{F} \times C,$$

where

L = Lower limit of the median class;
M = Cumulative frequency preceding the median class;
C = CI of the median class;
F = Frequency of the median class and
N = Total frequency.

Here $L = 25$, $N = 325$, $F = 62$, $C = 5$ and $M = 118$

$$\text{Median} = 25 + \left(\frac{\frac{325}{2} - 118}{62}\right) \times 5 = 25 + 3.58 = 28.5$$

Merits of Median

- Simplicity: Very simple measure of the central tendency of the series. In the case of simple statistical series, you can locate the median just at a glance.
- Free from the effect of extreme values: Unlike arithmetic mean, median value is not affected by the extreme values of the series.
- Certainty: Certainty is another merit of median. Median value is always a certain specific value in the series.

- Real value: Median value is real value and is a better representative value of the series compared to arithmetic mean, the value of which may not exist in the series at all.
- Graphic presentation: Besides algebraic approach, the median value can be also estimated through the graphic presentation of data.
- Possible even when data are incomplete: Median can be estimated even in the case of certain incomplete series. It is enough if one knows the number of items and the middle item of the series.

Demerits of Median

Following are the various demerits of median:

- Lack of representative character: It fails to be a representative measure in case of such series the different values of which are wide apart from each other. Also, median is of limited representative character because it is not based on all the items in the series.
- Unrealistic: When the median is located somewhere between the two middle values, it remains only an approximate measure, not a precise value.
- Lack of algebraic treatment: Median is not capable of further algebraic treatment. For example, a multiplication of median with the number of items in the series will not yield the sum total of the values of the series.

However, median is quite a simple method for finding an average of a series. It is quite a commonly used measure in the case of such series which are related to qualitative observation such as health of the student.

Quartiles

When data are arranged in ascending or descending order, it can be divided in various parts such as quartiles, deciles and percentiles. These values are collectively called quantiles and are the extension of median formula which divides a distribution of events into two equal parts. Since the basic purpose of these partition values is to divide data into different parts therefore a relationship exists between them. This relationship is shown in Table 12.3 and is elaborated with the help of a simple problem.

Let us discuss all of them for better understanding.

'Quartile' simply means 'one fourth of something'. Thus, we can safely say that a quartile value is a certain fourth of a data set. When a given data are ordered in the ascending descending manner, that is,

TABLE 12.3 Relationship Between Quartile, Deciles and Percentiles

First Quartile $Q_1 = P_2$	First Decile $D_1 = P_1$
Second Quartile $Q_2 = P_{10}$	Second Decile $D_2 = P_{20}$
Third Quartile $Q_3 = P_{75}$	Fifth Decile $D_5 = P_{50}$ and so on
Second Quartile = Fifth Decile = 50th Percentile = Median $Q_2 = D_5 = P_{50}$ = Median	

Source: Authors.

from the lowest value to the highest value and these data are divided into groups of four, we get what we call 'quartiles'.

> 'The quartile is a statistical term describing a division of a data set into four defined intervals based upon the values of the data and how they compare the entire data set' (Franklin 2007).

In quartiles statistics, there are three quartiles:

- First lower quartile Q_1: It lies in the 25 per cent of the bottom data.
- The second quartile Q_2: It is the same as 'median' of the data that divides the data in the middle and has 50 per cent of the data below it and the other 50 per cent of the data above it. It is denoted as Q_2.
- Third quartile Q_3: It has 75 per cent of the data below it and 25 per cent of the data above it.

Interpreting Quartiles

We have seen earlier that the median is a measure of the central tendency of the data but says nothing about how the data are distributed in the two arms on either side of the median. Quartiles help us measure this.

Thus, if the first quartile (Q_1) is far away from the median, while the third quartile (Q_3) is closer to it, it means that the data points that are smaller than the median are spread far apart while the data points that are greater than the median are closely packed together.

Deciles

'Deciles' are a simple statistical technique which measure inequality (Dikhanov 2005).[4] The main idea is that deciles show how income is distributed, how much of the income in a country is earned by lower wage earning groups and how much of the total income is earned by higher wage earning income groups. If the people in the top and bottom income groups earn the same proportion of the income, then there is income equality. If the top groups earn a much higher per cent of the total income, while people in the bottom groups earn much lower per cent of the total income, then there is inequality.

> The deciles of a distribution are the 9 values that split the data set into 10 equal parts.
>
> Decile is a system of dividing the given random distribution of the data or values in a series into 10 groups of similar frequency.

For example, you can use decile rating about schools and education. A decile, in this case, is a group into which similar schools in a country are placed. Schools are grouped in a way that reflects the average family situations and socio-economic backgrounds of the students at that school. In other words, the decile rating in a school is given to the economic and social factors of the community.

There are 10 deciles starting with decile 1 and moving through to decile 10. A decile is a statistical term, meaning that a group or population has been divided into 10 equally sized groups, giving 10 deciles.

[4] See http://hdr.undp.org/en/media/globalincometrends.pdf (accessed on 17 November 2017).

Schools in decile 1 have the highest proportion of students from low socio-economic backgrounds, while schools in decile 10 have the highest proportion of students from high socio-economic backgrounds.

Percentiles

Percentiles generalise the idea of a quartile and allow us to split our data set into many pieces. The *n*th percentile of a set of data is the value at which *n* per cent of the data is below it.

> 'Percentile is one of the values of a variable that divides the distribution of the variable into 100 groups having equal frequencies' (Croxton and Cowden 1946).

The median, first quartile and third quartile can all be expressed in terms of percentiles. Since half of the data is less than the median, and one half is equal to 50 per cent, we could call the median the 50th percentile. One fourth is equal to 25 per cent, and so the first quartile is the 25th percentile.

It is important to note that the words 'percentile' and 'percentage' are not the same. A 'percentage score' shows the proportion of a test that someone has completed correctly. A 'percentile score' tells us what per cent of other scores is less than the data point we are investigating. As seen in the earlier example, these numbers are rarely the same.

Percentile scores have a variety of uses. One common application of percentiles is for use with tests to serve as a basis of comparison for those who took a given test. For example, a score of 80 per cent initially sounds good. However, this does not sound as impressive when we find out that it is the 20th percentile—only 20 per cent of the class scored less than an 80 per cent on the test.

Percentiles are widely used in maintaining children's growth charts. In addition to a physical height or weight measurement, paediatricians typically state this in terms of a percentile score. A percentile is used in this situation in order to compare the height or weight of a given child to all children of that age. This allows for an effective means of comparison.

Computations of Quartiles, Deciles and Percentiles

Following are the formulas for calculating quartiles, deciles and percentiles.

Quartile

$$Q_1 = (n+1/4)\text{th}_{\text{item}} \text{ and } Q_3 = 3(n+1/4)\text{th}_{\text{item}}$$

Decile

$$D_1 = (n+1/10)\text{th}_{\text{item}} \text{ and } D_5 = (5(n+1)/10)\text{th}_{\text{item}}$$

Percentile

$$P_1 = (n+1/100)\text{th}_{\text{item}} \text{ and } P_{15} = (15(n+1)/100)\text{th}_{\text{item}}$$

Example

For odd numbers of values

Following are the numbers of defective chemistry laboratory test made in last 23 months by the university's chemistry department.

30, 36, 26, 16, 21, 33, 40, 32, 14, 10, 29, 23, 39, 17, 11, 18, 34, 19, 24, 21, 35, 42, 37

First thing that we have to do is to arrange the data in ascending order as follows:

10, 11, 14, 16, 17, 18, 19, 21, 21, 23, 24, 26, 29, 30, 32 33, 34, 35, 36, 37, 39, 40, 42

$$Q_1 = (23 + 1/4)\text{th}_{\text{item}} = 6\text{th item}$$

$$Q_3 = 3(23 + 1/4)\text{th}_{\text{item}} = 18\text{th item}$$

For even numbers of values
Following are the numbers of defective chemistry laboratory test made in last 24 months by the university's chemistry department.

10, 11, 14, 16, 17, 18, 19, 21, 21, 23, 24, 26, 29, 30, 32 33, 34, 35, 36, 37, 39, 40, 42, 45

$$Q_1 = (n + 1/4)\text{th}_{\text{item}} = (24 + 1/4) = (6.25)\text{th}_{\text{item}}$$

$$= 6\text{th}_{\text{item}} + (3/4) \times (7\text{th}_{\text{item}} - 6\text{th}_{\text{item}})$$

$$Q_1 = (18 + 3/4) \times (19 - 18) = 18.75$$

$$Q_3 = 3(n + 1/4)\text{th}_{\text{item}} = 3(24 + 1/4) = (18.75)\text{th}_{\text{item}}$$

$$= 18\text{th}_{\text{item}} + (1/4) \times (19\text{th}_{\text{item}} - 18\text{th}_{\text{item}})$$

$$Q_3 = (35 + 3/4) \times (36 - 35) = 35.75$$

Using the earlier formulas, we can easily compute the values of all deciles and percentiles for both odd and even numbers of observations.

Mode

The score that occurs most frequently in the distribution is known as 'mode'. However, for some frequency distributions, mode may not be the most frequent value. It is that value of the variate around which other items tend to concentrate most heavily. You may come across a situation where for some distributions the mode may not exist and even if it exists, it may not be unique as there may be more than one mode.

A distribution having only one mode is called 'unimodal'; a distribution having two modes is called 'bimodal' and a distribution having more than two modes is called 'multimodal'.

> The mode of a distribution is the value at the point around which the items tend to be most heavily concentrated. It may be regarded at the most typical of a series of values (Croxton and Cowden 1946).

It depicts the centre of concentration of the frequency in around a given value. Therefore, where your purpose is to know the point of the highest concentration you should use it. It is, thus, a positional measure.

Mode is often used in business. Mode is more suitable than mean or median in many situations. For instance, when we speak of 'most common wage'. We mean modal wage—the wage that the largest numbers of workers receive. In the case of a shopkeeper who sells shirts, he or she is interested in knowing the size of shirts which are commonly demanded. In such a situation, the mean would indicate a size that may not fit any person. Mode will give most common size of shirt which is most usually purchased by the customers.

Computation of the Mode

Ungrouped or Raw Data or Individual

For ungrouped data or a series of individual observations, mode is often found by mere inspection.

Example

2, 7, 10, 15, 10, 17, 8, 10, 2, 10

$$\text{Mode} = \text{Mo} = 10$$

In some cases, the mode may be absent while in some cases there may be more than one mode.

Example

Case 1: 12, 10, 15, 24, 30, 11, 9, 16, 8
For this distribution, there is no mode.
Case 2: 7, 10, 15, 12, 7, 14, 24, 10, 7, 20, 10
This distribution is bimodal as it contains two modes: 7 and 10.

Grouped or Raw Data

For discrete distribution, the highest frequency and its corresponding value is mode. However, for

$$\text{Mode} = \text{Mo} = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times c,$$

where
l = Lower limit of the modal class;
f_1 = frequency of the modal class;
f_0 = frequency of the class preceding the modal class;
f_2 = frequency of the class succeeding the modal class and
c = CI.

Example

The following table shows the size of primary schools in a province. The following steps will help you understand the computation of mode.
Number of Primary Schools by Level of Enrolment: 2015

School Size (Number of Pupils) Class Interval (CI)	Number of Schools (f)
0–50	5
50–100	14
100–150	40
150–200	91
200–250	**150**
250–300	87
300–350	60
350–400	38
400 and above	15

The highest frequency in the table is 150 and the corresponding CI is 200–250, which is modal class.
Here $l = 200, f_1 = 150, f_0 = 91, f_2 = 87, c = 50$

$$\text{Mode} = \text{Mo} = 200 + \frac{150-91}{2\times 150 - 91 - 87} \times c = 200 + \frac{2{,}950}{122}$$

Mode = 200 + 24.18 = 224.18

Merits

Following are the various merits of mode:

- Simple and popular: It is very simple measure of central tendency.
- Less effect of marginal values: It is less affected by marginal values in the series as compared to mean. It is determined only by the value with highest frequencies.
- Graphic presentation: It can be located graphically, with the help of histogram.
- Best representative: Since mode is that value which occurs most frequently in the series, it is the best representative value of the series.
- No need of knowing all the items or frequencies: The computation of mode does not require knowledge of all the items and frequencies of a distribution. In simple series, you should just know the items with highest frequencies in the distribution.

Demerits

Following are the various demerits of mode:

- Uncertain and vague: It is an uncertain and vague measure of the central tendency.
- Not capable of algebraic treatment: It is not suitable for further algebraic treatment.
- Difficult: With identical frequencies of all items, it is difficult to identify the modal value or modal class.
- Complex procedure of grouping: The computation of mode involves cumbersome procedure of grouping the data.
- Ignores extreme marginal frequencies: The calculation of mode does not take account of extreme marginal frequencies.

In conclusion, measures of central tendency describe how the data cluster together around a central point. They are useful for making comparisons between groups of individuals or between sets of figures. They reduce a huge data to a single figure, and thus make comparisons easy.

Unfortunately, there is no single, standard procedure for determining central tendency. The problem is that no single measure will always produce a typical, representative value in every situation. Therefore, we use different ways to measure central tendency.

Which Measure to Choose?

The mode should be used when calculating measure of centre for the qualitative variable. When the variable is quantitative with symmetric distribution, then the mean is proper measure of centre. In a case of quantitative variable with skewed distribution, the median is good choice for the measure of centre.

You should keep in mind that the mean could be highly influenced by extreme value or an observation that falls far from the rest of the data, called an outlier. You should also note that the sample mode, the sample median and the sample mean of the variable in question have corresponding population measures of centre, that is, we can assume that the variable in question have also the population mode, the population median and the population mean, which are all unknown. Then the sample mode, the sample median and the sample mean can be used to estimate the values of these corresponding unknown population values.

Measures of Dispersion

We have discussed various measures of central tendency. These measures serve to locate the centre of the distribution and help us to represent the entire mass of the data by a single value. But these measures do not describe the data fully and adequately. In order to understand it, let us consider the following example.

Example

Consider the following marks in statistics (100 marks maximum) of seven students of Grade 12 in School A and School B:

School A:	35	45	50	65	70	90	100
School B:	60	65	65	65	66	64	70

Here we observe that in both the schools the mean of the data is the same, namely 65.

However,

- In school A, the observations are much more dispersed (scattered) from the mean.
- In school B, almost all the observations are concentrated around the mean.

It is evident that the two groups differ even though they have the same mean. It is therefore necessary to differentiate between the two schools. We need some other measures which are related with the measures of dispersion in a distribution. It is for this reason that we study what is known as ‘measures of dispersion’.

Dispersion

In statistics, the term ‘dispersion’ refers to the variation of a random variable or its probability distribution. It shows how far away the data points lie from the central value. To express this quantitatively, measures of dispersion are used in descriptive statistics.

> ‘Dispersion (also called variability, scatter, or spread) denotes how stretched or **squeezed a distribution** (theoretical or that underlying a statistical sample) is (Aliaga and Gunderson 2000)’.
>
> ‘The extent to which scores in a data set differ from each other. Such a measure is called the dispersion of a distribution’ (Dictionary Definition)

To understand the meaning of dispersion, let us consider the following example.

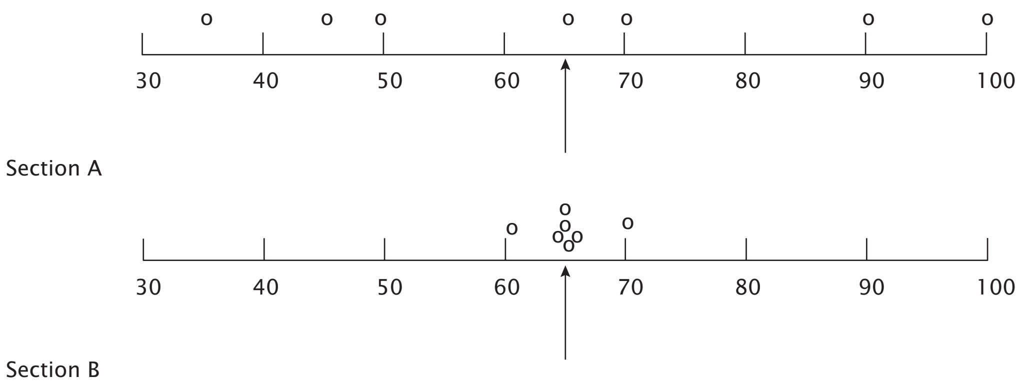


FIGURE 12.2 Dispersion in a Distribution

Example

Two sections of seven students each in class 12 in a certain school were given a common test in physics (100 maximum marks). The scores of the students are given as follows:

Section A:	35	45	50	65	70	90	100
Section B:	60	65	65	65	66	64	70

The average score in Section A is 65.
The average score in Section B is 65.

Let us construct a dot diagram, on the same scale for Section A and Section B (Figure 12.2). The position of arithmetic mean for both the sections is marked by arrow.

It is clear that the extent of dispersion of the data in Section A is different compared to Section B. The measurement of the scatter of the given data about the average is referred to as a measure of dispersion or scatter. The smaller the value of any measure of dispersion, the more consistent the data are.

Dispersion Graph

Since dispersion is a measure of data variability, it affects the confidence that a researcher can have in the representativeness and reliability of central location measures (Figure 12.3). A dispersion graph depicts the relationship between two variables. The graph gives a simple illustration of how one variable can influence the other. When constructing a dispersion graph, you should clearly define the variables that are to be evaluated. Then, plot data pairs using the horizontal axis for probable cause and using the vertical axis for probable effect. A dispersion graph locates individual data values along a number line, thereby representing the position of each data value in relation to all the other data values.

There are four main measures of dispersion frequently used in a research study. They are:

- Range
- Interquartile range
- Mean absolute deviation from mean
- Variance
- SD

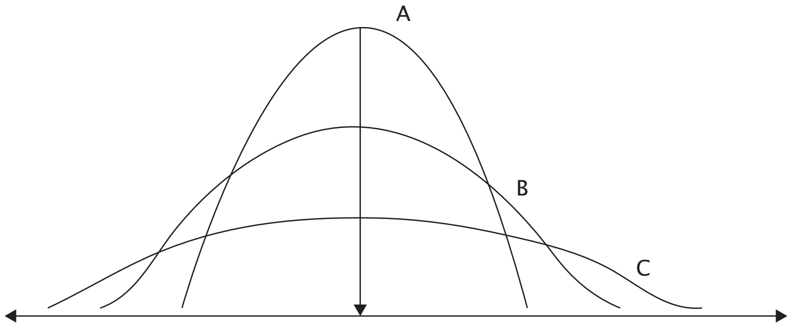


FIGURE 12.3 Dispersion Graph

Range

'Range' is the simplest of all our methods for measuring dispersion. It is the difference between the highest value and the lowest value in the data set. While being simple to compute, the range is often an unreliable measure of dispersion since it is based on only two values in the set.

The calculation of the range is very simple and straightforward. All you need to do is to find the difference between the largest data value in your set and the smallest data value. The formula you can use is:

$$\text{Range} = \text{Maximum Value} - \text{Minimum Value}$$

Example

12, 25, 27, 29, 36, 38, 40, 43, 50, 54, 62

Range = 62–12 = 0

Coefficient of Range

$$\text{Coefficient of Range} = \frac{L - S}{L + S}$$

In individual observations and discrete series, L and S are easily identified. In continuous series, the following two methods should be followed.

Method 1

L = Upper boundary of the highest class
S = Lower boundary of the lowest class

Method 2

L = Mid-value of the highest class
S = Mid-value of the lowest class

Interquartile Range

The interquartile range is another range used to measure dispersion (spread) in the data set. The difference between the upper and lower quartiles ($Q_3 - Q_1$), which is called 'the interquartile range', also indicates the dispersion of a data set. The interquartile range spans 50 per cent of a data set and eliminates the influence of outliers because, in effect, the highest and lowest quarters are removed.

$$\text{Interquartile Range} = Q_3 - Q_1$$

Examples

The math test marks of a class are as follows:

52, 45, 25, 75, 63, 86, 72, 85, 55, 65, 70, 82, 90, 48, 68, 86, 65, 64, 78, 75, 32, 42.

Find the interquartile range.

Step 1

Arrange the marks in ascending order as follows:

25	32	42	45	48	52	55	63	64	65	68	70	72	75	78	82	85	86	86	90
					a				65					a					
					Q_1				Q_2					Q_3					

Step 2

$$\text{Interquartile range} = Q_3 - Q_1 = 78 - 52 = 26 \text{ Marks}$$

This tells that the middle 50 per cent of the test score are dispersed over a range of 26 marks.

Limitations of Range

The range is a very crude measurement of the spread of data because it is extremely sensitive to 'outliers'. Thus, the range can sometimes be misleading when there are extremely high or low values. A single individual data value can greatly affect the value of the range. For example, consider the set of data 8, 11, 5, 9, 7, 6, 10.

The maximum value is 10, the minimum is 5 and the range is 5 (10 − 5 = 5). Now consider the same set of data, only with the value 100 included. The range now becomes 100 − 5 = 95. The inclusion of a single extra data point greatly affected the value of the range. The single value of 100 makes the range large, but most values are around 10.

Range is an informative tool used as a supplement to other measures such as the SD or semi-interquartile range, but it should rarely be used as the only measure of dispersion.

Mean Absolute Deviation from the Mean

'Mean absolute deviation from the mean' is average of absolute differences (differences expressed without plus or minus sign) between each value in a set of values and the average of all values of that set. For example, the average (arithmetic mean or mean) of the set of values 1, 2, 3, 4 and 5 is (15 ÷ 5) or 3. The difference between this average (3) and the values in the set is 2, 1, 0, −1 and −2; the absolute difference being 2, 1, 0, 1 and 2. The average of these numbers (6 ÷ 5) is 1.2 which is the mean deviation. Also called mean absolute deviation, it is used as a measure of dispersion where the number of values or quantities is small; otherwise, SD is used.

In Figure 12.3, we noted that the scores in section B cluster around the mean, while in section A the scores are spread away from the mean. Let us take the deviation of each observation from the mean and add all such deviations. If the sum is 'large', the dispersion is 'large'. If, however, the sum is 'small' the dispersion is small. If you simply compute this, then the sum of the deviations taken from the mean will always be 0 for any set of data.

Ungrouped or raw data

$$\text{Mean Deviation (MD)} = \sum|X - \bar{X}|/N$$

Where

$\Sigma\,|X - \bar{X}|$ = Deviations from the means (± signs ignored) and

N = Number of observations.

Example

Section A:	35	45	50	65	70	90	100
Section B:	60	65	65	65	66	64	70

SECTION A

X	$\lvert X - \bar{X}\rvert$
35	30
45	20
50	15
65	0
70	5
90	25
100	35
$\Sigma \boldsymbol{X} = \mathbf{455}$	130

SECTION B

X	$\lvert X - \bar{X}\rvert$
60	5
65	0
65	0
65	0
66	1
64	1
70	5
$\Sigma \boldsymbol{X} = \mathbf{455}$	12

$$\bar{X} = 455/7 = 65 \quad \bar{X} = 455/7 = 65$$

$$\text{MD} = 130/7 = 18.57 \;\; \text{MD} = 12/7 = 1.71$$

Mean Deviation of discrete data:

$$\text{Mean Deviation (MD)} = \sum f\,|X - \bar{X}|/N$$

Example

Find the mean deviation from the mean of the following data:

Size of Item X	4	6	8	10	12	14	16
Frequency (f)	2	4	5	3	2	1	4

Solution

Size of Item	Frequency (F)	$X - \overline{X}$	$\lvert X - \overline{X}\rvert$	$F\lvert X - \overline{X}\rvert$
4	2	–5.7	5.7	11.4
6	4	–3.7	3.7	14.8
8	5	–1.7	1.7	8.5
10	3	0.3	0.3	0.9
12	2	2.3	2.3	4.6
14	1	4.3	4.3	4.3
16	4	6.3	6.3	25.2
	21			69.7

$$\text{Mean Deviation (MD)} = \sum f \,|\, X - \overline{X} \,|\, / N = 69.7 / 21 = 3.319$$

Mean Deviation of Continuous Series:

$$\text{Mean Deviation (MD)} = \sum f \,|\,(D)\,|\, / N,$$

where
D = mid-value (m) – average and
m = mid-point.

Example

Calculate the mean deviation from mean of the following distribution:

Marks	0–10	10–20	20–30	30–40	40–50
No. of Students	5	8	15	16	6

Solution

Marks	Class Marks (x)	f	$X - \bar{x}$	$\lvert x - \bar{x}\rvert$	$f\lvert x - \bar{x}\rvert$
0–10	5	5	–22	22	110
10–20	15	8	–12	12	96
20–30	25	15	–2	2	30
30–40	35	16	8	8	128
40–50	45	6	18	18	108
Total		50			472

$$\text{Mean Deviation from the Mean (MD)} = \sum[f \,|\, x - \overline{X} \,|] / N$$

$$\text{MD} = 472 / 50 \text{ marks} = 9.44 \text{ Marks}$$

Coefficient of Mean Deviation

$$\text{Coefficient of Mean Deviation} = \frac{\text{Mean Deviation}}{\text{Mean}}$$

If the result were desired in percentage, the coefficient of mean deviation would be:

$$\text{Coefficient of Mean Deviation} = \frac{\text{Mean Deviation}}{\text{Mean}} \times 100$$

Mean Deviation is rarely being used as a measure of dispersion. It is rather useful in business and economics for studying business cycles.

Standard Deviation

Karl Pearson introduced the concept of SD in 1893. It is the most important measure of dispersion and is widely used in many statistical formulae and hence in research studies. It provides accurate result. Square of SD is called 'variance'. The SD is denoted by the Greek letter s (sigma).

> SD is the positive square root of the arithmetic mean of the square of the deviations of the given observation from their arithmetic mean.

The SD is a numerical value used to indicate how widely individuals in a group vary. If individual observations vary greatly from the group mean, the SD is big and vice versa.

Individual Series

Deviations taken from actual mean

$$\sigma = \sqrt{\frac{\sum (X - \overline{X})^2}{N}}$$

where

σ = SD;
Σ = Sum of
X = Value in the data set;
$\overline{X}$ = Mean of the values in the data set and
N = Number of values in the data set.

Deviations taken from assumed mean

$$\sigma = \sqrt{\frac{\sum d^2}{N} - \left(\frac{\sum d}{N}\right)^2},$$

where d = deviations from the assumed mean = $(X - A)$.

Example

SD for two sections

SECTION A		
X	$(X - \overline{X})$	$(X - \overline{X})^2$
35	–30	900
45	–20	400
50	–15	225

SECTION B		
X	$(X - \overline{X})$	$(X - \overline{X})^2$
60	–5	25
65	0	0
65	0	0

65	0	0
70	5	25
90	25	625
100	35	1,225
Σ*X* = 455		3,400

$\bar{X} = 455/7 = 65$

$$\sigma = \sqrt{\frac{\Sigma(X-\bar{X})^2}{N}} = \sqrt{\frac{3,400}{7}} = 22.03$$

65	0	0
66	1	1
64	1	1
70	5	25
Σ*X* = 455	12	52

$\bar{X} = 455/7 = 65$

$$\sigma = \sqrt{\frac{\Sigma(X-\bar{X})^2}{N}} = \sqrt{\frac{52}{7}} = 2.73$$

Discrete Series

Actual mean method

The following formula entails the actual mean method:

$$\sigma = \sqrt{\frac{\Sigma fd^2}{\Sigma f}}$$

1. Find deviations for various items from the means, that is, $x - x = \text{d}$.
2. Square the deviations ($= d^2$) and multiply by the respective frequencies (f) to obtain fd^2
3. Sum the product (Σfd^2), and then apply the following formula to obtain the assumed mean method:

$$\sigma = \sqrt{\frac{\Sigma fd^2}{\Sigma f} - \left(\frac{\Sigma fd}{\Sigma f}\right)^2},$$

where $d = (X - A)$, and $N = \Sigma f$.

Stepwise Deviation

$$\sigma = \sqrt{\frac{\Sigma f\dot{d}^2}{N} - \left(\frac{\Sigma f\dot{d}}{N}\right)^2} \times \text{C},$$

where

$\dot{d}$ = deviation from assumed mean and

C = CI.

Continuous Series

$$\sigma = \sqrt{\frac{\Sigma f\dot{d}^2}{N} - \left(\frac{\Sigma f\dot{d}}{N}\right)^2} \times \text{C},$$

where

$$\dot{d} = \frac{m - A}{C} \text{ and } C = \text{ CI}$$

It is important to distinguish between the SD of a population and the SD of a sample. They have different notations, and they are computed differently. The SD of a population is denoted by σ and the SD of a sample by *s*. These methods have been explained in Chapter 4.

Variance

The 'variance' is a numerical value used to indicate how widely individuals in a group vary. If individual observations vary greatly from the group mean, the variance is big and vice versa. It is represented by (σ^2).

$$\text{Variance } (\sigma^2) = \sum (X_1 - \bar{X})^2 / N$$

$$\text{Standard Deviation}(\sigma) = \sqrt{\sigma^2}$$

And finally, the variance is equal to the square of the SD.

Coefficient of Variation

Basically, the SD is an absolute measure of dispersion. It is shown in terms of units in which the original figures are collected and stated. For instance, the SD of heights of students cannot be compared with the SD of weights of students, as both are expressed in different units, that is, heights in centimetre and weights in kilogrammes. It is, therefore, necessary to convert the SD into a relative measure of dispersion for the purpose of comparison. The relative measure is known as the 'coefficient of variation' (CV).

The CV is obtained by dividing the SD by the mean and multiply it by 100.

$$\text{CV} = \sigma / \bar{X} \times 100$$

If we want to compare the variability of two or more series, we can use CV. The series or groups of data for which the CV is greater indicate that the group is more variable, less stable, less uniform, less consistent or less homogeneous. If the CV is less, it indicates that the group is less variable, more stable, more uniform, more consistent or more homogeneous.

Example

Suppose there are two schools, School A and School B, with teachers' average monthly salary (in US$) and SDs as follows:

School A	Average (US$) ($\bar{X}$)	Standard Deviation (σ)	No. of Teachers
A	87.5	5.5	100
B	81.5	4.5	100

Q1. Which school, A or B, pays out a larger amount as monthly salary?
Q2. Which school, A or B, has greater variability in individual salaries?

Solution

Question 1

- Total salary paid by School A = 87.5 × 100 = 8,750
- Total salary paid by School B = 81.5 × 100 = 8,150

Therefore, School A pays larger amount as monthly salary.

Question 2

- $\text{CV(School A)} = 5.5/87.5 \times 100 = 6.28$
- $\text{CV(School B)} = 4.5/81.5 \times 100 = 5.52$

The total monthly salary paid by School A is more than School B. However, School B has greater variability in individual salaries, since CV of School A is greater than CV of School B.

Example

Public expenditure in primary education per pupil in five years in two districts is given as follows.

Year	District A (in US$)	District B (in US$)
2011	20	10
2012	22	20
2013	19	18
2014	23	12
2015	16	15
	Σ100	Σ75

Which district has more stable per pupil public expenditure in primary education?

District A

$$\overline{X}_A = 20$$

$$CV_A = 12.25\%$$

District B

$$\overline{X}_B = 15$$

$$CV_B = 24.60\%$$

District A had more stable public expenditure than District B, because CV is less in District A.

Skewness

Measures of skewness, like measures of central tendency and dispersion, study the characteristics of a frequency distribution. Averages describe the central value of the distribution and measures of dispersion inform us about the concentration of the items around a central value. These measures, however, do not disclose whether the dispersal of value on either side of an average is symmetrical or not. If observations are ordered in a symmetrical manner around a measure of central tendency, we obtain a symmetrical distribution; otherwise, it may be arranged in an asymmetrical order which gives asymmetrical distribution. Thus, 'skewness' is a measure that describes the degree and direction of departure from symmetry.

A symmetrical distribution, when presented on the graph paper, offers a 'symmetrical curve', where the value of mean, median and mode are exactly the same or equal. On the other hand, in an asymmetrical distribution, the values of mean, median and mode are different. When two or more symmetrical distributions are compared, the difference in them is studied with 'Kurtosis'. On the other hand, when

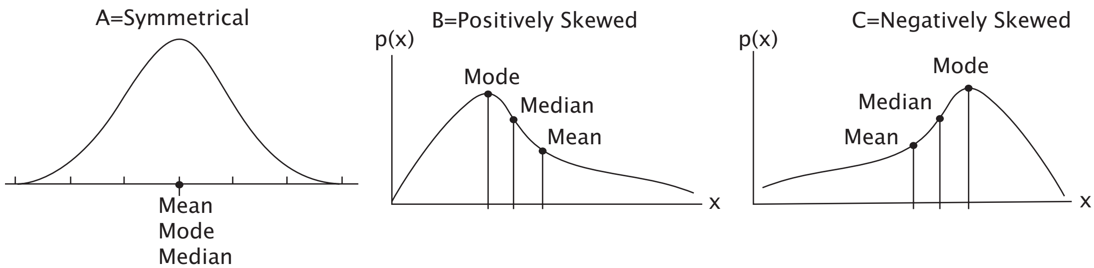


FIGURE 12.4 Shapes of Distribution

two or more symmetrical distributions are compared, they will give different degrees of skewness. These measures are mutually exclusive, that is, the presence of skewness implies absence of kurtosis and vice versa (School of Open Learning).

Skewness refers to deviation from symmetry or more precisely, the lack of symmetry. Curves B and C in Figure 12.4 are not symmetric and hence called skewed distribution. A distribution, or data set, is symmetric if it looks the same to the left and right of the centre point (curve A in Figure 12.4). Also, mean, median and mode coincide and quartiles are equispaced. Depending on the positions of mean and mode curve is specified as positively skewed curve, symmetric curve or negatively skewed curve.

The measure of asymmetry is called 'measures of skewness'. These are classified as absolute measures and relative measures. Absolute measures are known as measures of skewness and relative measures are known as coefficient of skewness.

- Measure of skewness is positive for positively skewed curve.
- Measure of skewness is 0 for symmetric curve.
- Measure of skewness is negative for negatively skewed curve.

That is:

- If sk > 0, the curve is positively skewed curve.
- If sk 3 = 0, the curve is symmetric curve.
- If sk 3 < 0, the curve is negatively skewed curve.

The important measures of skewness are:

a. Karl Pearson's Method or Pearsonian coefficient of skewness
b. Bowley's method

Before we discuss the computation methods, it is important to remind that for all the three methods listed before, we have to know the methods of calculation of mean, median, mode, quartiles for all types of data discussed before. In the following examples, we just present formulas only for discrete data.

Karl Pearson's Method or Pearsonian Coefficient of Skewness

According to Karl Pearson, the absolute measure of skewness = mean − mode. This measure is not suitable for making valid comparison of the skewness in two or more distributions because the unit of measurement may be different in different series. To avoid this difficulty, researchers use relative measure of skewness called Karl Pearson's coefficient of skewness given by:

$$\text{Karl Pearson's coefficient of skewness } (sk) = \frac{\text{Mean} - \text{Mode}}{\text{Standard Deviation}}$$

If the mode is ill defined in a distribution, the coefficient can be computed by the following formula:

$$\text{Coefficient of Skewness } (sk) = \frac{3(\text{Mean} - \text{Median})}{\text{Standard Deviation}}$$

Example

The following gives the age structure of nine university students. The coefficient of skewness can be computed as follows:

Age in Years (X)	Deviation from Assumed Mean A = 25(d)	d^2
25	0	0
15	–10	100
23	–2	4
40	15	225
27	2	4
25	0	0
23	–2	4
25	0	0
20	–5	25
N = 9	Σ = **–2**	Σd^2 = **362**

Mean ($\bar{X}$) = 24.78 Standard Deviation (σ) = 6.3 Mode = 25

$$\text{Karl Pearson's coefficient of skewness } (sk) = \frac{24.78 - 25}{6.3} = -0.03$$

Bowley's Coefficient of Skewness

In Karl Pearson's method of measuring skewness, the whole of the series is needed. Bowley has suggested a formula based on relative position of quartiles. In a symmetrical distribution, the quartiles are equidistant from the value of the median, that is,

$$\text{Median is equal to } Q_2 \text{ or median} - Q_1 = Q_3 - \text{median}$$

But in a skewed distribution, the quartiles will not be equidistant from the median. Hence, Bowley has suggested the following formula:

$$\text{Bowley's coefficient of skewness } (sk) = \frac{Q_3 + Q_1 - 2\text{ Median}}{Q_3 - Q_1}$$

Example

Find Bowley's coefficient of skewness for the following series.

2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22

Q_1 = Size of 3rd item = 6

Q_3 = Size of 9th item = 18

Median = Size of 6th item = 12

$$\text{Bowley's coefficient of skewness } (sk) = \frac{(18+6)-(2\times 12)}{18-6} = 0$$

Since $sk = 0$, the given distribution is symmetrical.

The measure of skewness based on moments is denoted by β_1 and is given by:

$$\beta_1 = \mu_3^2 / \mu_2^3$$

If μ_3 is negative, then β_1 is negative.

Moments

'Moments' can be defined as the arithmetic mean of various powers of deviations taken from the mean of a distribution. These moments are known as central moments.

The first four moments about arithmetic mean or central moments are defined in Table 12.4.

TABLE 12.4 Measures for the First Four Moments

	Individual Series	*Discrete Series*
First Moments about the Mean (μ_1)	$\frac{\Sigma(X-\bar{X})}{N} = 0$	$\frac{\Sigma f(X-\bar{X})}{\Sigma f} = 0$
First Moments about the Mean (μ_2)	$\frac{\Sigma(X-\bar{X})^2}{N} = \sigma^2$	$\frac{\Sigma f(X-\bar{X})^2}{\Sigma f}$
First Moments about the Mean (μ_3)	$\frac{\Sigma(X-\bar{X})^3}{N}$	$\frac{\Sigma f(X-\bar{X})^3}{\Sigma f}$
First Moments about the Mean (μ_4)	$\frac{\Sigma(X-\bar{X})^4}{N}$	$\frac{\Sigma f(X-\bar{X})^4}{\Sigma f}$

Source: Authors.

Example

Computation of first four moments

X	0	1	2	3	4	5	6	7	8
f	5	10	15	20	25	20	15	10	5

Solution

x	f	fx	$d = x - \bar{x}$ $(x - 4)$	fd	fd^2	fd^3	fd^4
0	5	0	–4	–20	80	–320	1,280
1	10	10	–3	–30	90	–270	810
2	15	30	–2	–30	60	–120	240
3	20	60	–1	–20	20	–20	20
4	25	100	0	0	0	0	0
5	20	100	1	20	20	20	20
6	15	90	2	30	60	120	240
7	10	70	3	30	90	270	810
8	5	40	4	20	80	320	1,280
	$\Sigma f = 125$	$\Sigma fx = 500$	$\Sigma d = 0$	$\Sigma fd = 0$	$\Sigma fd^2 = 500$	$\Sigma fd^3 = 0$	$\Sigma fd^4 = 4{,}700$

$$\overline{X} = 4$$

$$\mu_1 = \sum fd / 125 = 0/125 = 0 \qquad \mu_2 = \sum fd^2 / 125 = 0$$

$$\mu_3 = \sum fd^3 / 125 = 0/125 = 0 \qquad \mu_4 = \sum fd^4 / 125 = 4{,}700/125 = 37.6$$

Kurtosis

Kurtosis is a measure which determines whether the data are peaked or flat relative to a normal distribution, that is, data sets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly and have heavy tails. Data sets with low kurtosis tend to have a flat top near the mean rather than a sharp peak. A uniform distribution would be the extreme case.

As far as the measurement of shape is concerned, there are two characteristics—skewness which refers to asymmetry of a series and kurtosis which measures the peakedness of a normal curve. All the frequency curves consist of different degrees of flatness or peakedness. This characteristic of frequency curve is called as kurtosis. Measure of kurtosis indicates the shape of top of a frequency curve. Measure of kurtosis denotes the extent to which a distribution is more peaked or more flat topped than the normal curve, which is symmetrical and bell-shaped and is designated as 'mesokurtic' or 'normal curve'. If a curve is relatively narrower and peaked at the top, it is designated as 'leptokurtic'. If the frequency curve is more flat than normal curve, it is designated as 'platykurtic'. This is shown in Figure 12.5.

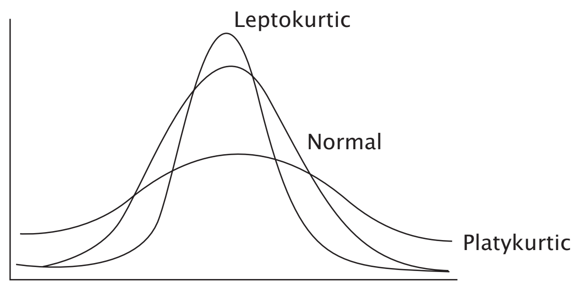


FIGURE 12.5 Shapes of Curves

Source: Authors.

Measure of Kurtosis

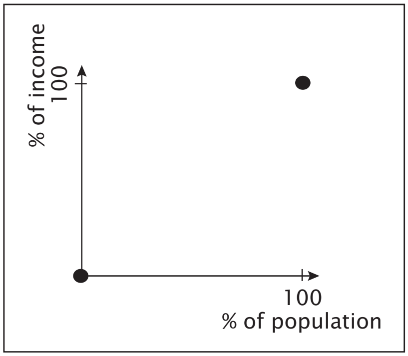
The measure of kurtosis of a frequency distribution is denoted as follows:

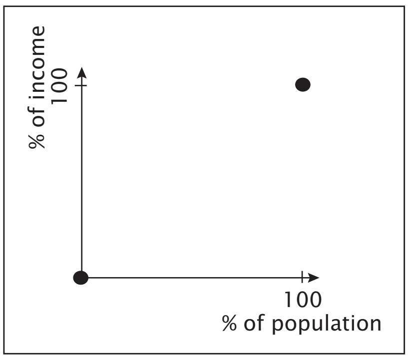
$$\beta_2 = \mu_4 / \mu_2^2$$

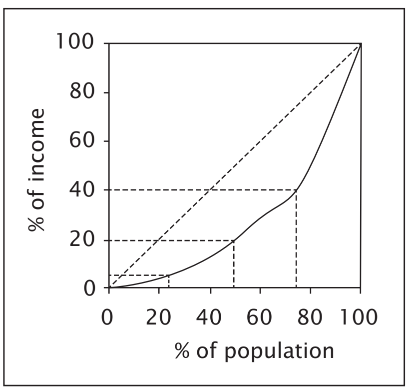
If $\beta_2 = 3$, the distribution is said to be normal and the curve is mesokurtic.
If $\beta_2 > 3$, the distribution is said to be more peaked and the curve is leptokurtic.
If $\beta_2 < 3$, the distribution is said to be flat topped and the curve is platykurtic.

Lorenz Curve: How to Measure Income Inequality?

Lorenz curve is a statistical tool which graphically shows the dispersion. It simply describes income distribution using a two-dimensional graph. To do this, consider lining people (or households depending on the context of the research study) in an economy up in order of income from smallest to largest. The horizontal axis of the Lorenz curve in this case would then show the cumulative percentage of these lined up people that are being considered.







For instance, 20 per cent on the horizontal axis will represent the bottom (poorest 20%) of income earners, 50 per cent will represent the bottom half of income earners and so on.

On vertical axis of the Lorenz curve, we show the per cent of total income in the economy.

The Lorenz curve starts with the points (0, 0) and (100, 100) which should be the ends of the curve. The obvious reason is that the bottom 0 per cent of the population (which has no people) has, by definition, 0 per cent of the economy's income, and 100 per cent of the population has 100 per cent of the income.

The rest of the curve is then is drawn by plotting all the percentages of the population between 0 and 100 per cent and plotting the corresponding percentages of income. This is shown in the adjacent figure. In this figure, the point (25, 5) represents the hypothetical fact that the bottom 25 per cent of people have 5 per cent of the total income. The point (50, 20) shows that the bottom 50 per cent of people have 20 per cent of the income, and point (75, 40) shows that the bottom 75 per cent of people have 40 of the income.

A Lorenz curve always bows downwards as in the example earlier. The reason is that it is mathematically impossible for the bottom 20 per cent of earners to make more than 20 per cent of the income, for the bottom 50 per cent of people to make more than 50 per cent of the income and so on.

The dotted line in the figure is the 45-degree line that represents perfect income equality in an economy. Perfect income equality means everyone making the same amount of money. That means the bottom 5 per cent has 5 per cent of the income, the bottom 10 per cent has 10 per cent of the income and so on.

Thus, Lorenz curves that are bowed further away from this diagonal correspond to economies with more income inequality.

Importance of Measures of Dispersion

We have discussed earlier the most widely used methods of dispersion in research. These measures occupy a significant place in any research endeavour. There are several reasons for which these methods are deployed in research. In research, measures of dispersion are required for many reasons. Among them, the following four reasons are highly significant. They:

- Determine the reliability of an average
- Serve as a basis for the control of the variability
- Facilitate comparisons between two or more series with regard to their variability
- Facilitate the use of other statistical measures

Reliability of an Average

Measures of variation indicate as to how far an average is representative of the mass. When dispersion is small, the average is a typical value in the sense that it closely represents the individual value and it is reliable in the sense that it is a good estimate of the average in the corresponding universe. On the other hand, when dispersion is large, the average is not so typical, and unless the sample is very large, the average may be quite unreliable.

Control of the Reliability

Another purpose of measuring dispersion is to highlight the nature and cause of variation in order to control the variation itself. In an education setting matter of students' participation rates (intake rates, enrolment ratios, attendance rates and so on), student flow rates (repetition, dropout, promotion, graduation, survival and TRs) and public expenditure are the basic guides commonly used in preparing the diagnosis of the system. Similarly, in matter of health, variations in body temperature, pulse beat and blood pressure are the basic guides to diagnosis. In an education setting, a special problem requiring the measurement of variability is the measurement of 'inequality' in demand for and supply of educational provision and services by various socio-economic traits of population.

Comparisons of Variability among Data Series

Measures of dispersion facilitate a comparison between two or more series with regard to their variability. The study of variation should be looked upon as a means of determining uniformity or consistency. A high degree of variation would signify little uniformity or consistency, whereas a low degree of variation would mean great uniformity or consistency.

Facilitating Other Statistical Measures in Data Analysis

Several powerful analytical statistical tools widely used in research such as correlation analysis, the testing of hypothesis, the analysis of fluctuations, techniques of production control, cost control and so on are grounded on measures of variation of one kind or another.

To sum up, while measures of central tendency are used to estimate 'normal' values of a data set, measures of dispersion are important for describing the spread of the data or its variation around a central value. Two distinct samples may have the same mean or median, but completely different levels of variability or vice versa. An appropriate description of a set of data should include both of these characteristics. There are several methods that researchers can use in order to measure the dispersion of a data set, each with its own set of advantages and disadvantages.

Index Numbers

Researchers construct index numbers to study the change in the effects of such factors which cannot be measured directly. According to Bowley, for example, changes in business activity in a country are not capable of direct measurement but it is possible to study relative changes in business activity by studying the variations in the values of some such factors which affect business activity and which are capable of direct measurement.

> 'An index number is a statistical measure designed to show changes in a variable or a group or related variables with respect to time, geographic location or other characteristics such as income, profession, etc'. (Andrew and Kotz 2011).
>
> 'An index number measures the relative change in price, quantity, value or some other item of interest from one time period to another'.

Most of you are familiar with indexes such as the consumer price index (CPI), which are released periodically and on a regular basis by the government, business publications and in most daily newspapers.

Why convert data to indexes? The simple answer is that an index is a convenient way to express a change in a diverse group of items. For example, the CPI comprises many items, such as items of food, clothing, shelter, tuition fees, transport expenses, medical bills and so on. Prices of all items are expressed, say, in US dollars per kilogramme, box, metre and many other different units. Only by converting the prices of these many diverse goods and services to one index number can the government and others concerned with inflation be kept informed of the overall movement of consumer prices.

Converting data to indexes also facilitates our understanding of the trend in a series composed of exceptionally large numbers. For example, the total public expenditure on education in Country A for 2015 was US$346.7 million and US$331.1 million for 2014. The increase of US$15.6 million appears significant. Yet if the 2015 expenditure were expressed as an index based on 2014 expenditure, the increase would be approximately 4.7 per cent.

$$\frac{\text{Public Expenditure on Education in 2015}}{\text{Public Expenditure on Education in 2014}} = \frac{346.7}{331.1} \times 100 = 104.7$$

In education research, index numbers are commonly used for:

- Analysing and framing suitable policies
- Revealing trends and tendencies
- Deflating the quantitative variables

In this section, we will discuss the several methods used in constructing the index numbers.

Construction of Index Numbers

There are several methods used for the construction of index numbers. Basically, indexes are divided into two main categories, namely,

- Simple indexes
- Weighted indexes

Simple Indexes

There might arise many situations during your research when you would be wishing or required to combine several items and develop an index to compare the cost of this aggregation of items in two different time periods. For example, you might be interested in an index for items that relate to the university student index. The items in the index might include the cost of books, tuition, stationery, transport, hostel, meals and entertainment. There are two most frequently used methods which help us combine the items to determine the index.

Simple Average of the Price Indexes

Example

Consider that you are the head of students' kitchen of a university. You have been providing six food items shown in Table 12.5 for breakfast to the boarding students of the university. The table also shows the prices of six food items for the years 2000 and 2015. Using these data, you wish to compute an index for this group of food items for 2015, assuming year 2000 as the base year (2000 = 100).

You could begin by computing a simple average of the price indexes for each item, using 2000 as the base year and 2015 as the given year. The simple index for bread is 257.1, found by using the following formula. You compute the simple index for the other items in Table 5.24 similarly. The largest price increase is for bread, 157.1 per cent (257.1 − 100 = 157.1), and milk was a close second, with 125 per

TABLE 12.5 Computation of Indexes for Food Prices 2015 (Base Year: 2000)

Item	*Unit (per)*	*2000 Price (US$)*	*2015 Price (US$)*	*Simple Index*
Bread (Britannia)	Pack	0.77	1.98	257.1
Eggs	Dozen	1.85	1.84	99.5
Milk (Fresh Full Cream)	Litre	0.88	1.98	225.0
Apples (Golden)	Kg	1.46	1.75	119.9
Jam (Carrefour)	Bottle	1.58	1.70	107.6
Coffee (Nestle)	400 g	4.40	3.99	90.7
TOTAL		**10.94**	**13.24**	

Source: Authors.

cent. The price of eggs dropped by half a per cent in the period, found by 100.0 – 99.5 = 0.5. Then it would be natural to average the simple indexes. The formula is:

$$P = P_1 / P_0(100) = 1.98 / 0.77 \times 100 = 257.1$$

Compute the simple index for other items in Table 12.5 similarly. The largest price increase is for bread, 157.1 per cent (257.1 – 100 = 157.1), and milk was a close second with 125 per cent. The price of eggs dropped by half a per cent in the period, found by 99.5 – 100.0 = –0.5. Then it would be natural to average the simple indexes. The formula is:

$$\text{Simple average of the price relatives} = P = \sum P_1 / n,$$

where

P_i refers to the simple index for each of the items and n the number of item. In our example, the index is 150.0, computed by

$$P = \sum P_1 / n = 257.1 + 99.5 + 225.0 + 119.9 + 107.6 + 90.7 / 6 = 150.0$$

This indicates that the mean of the group of indexes increased 50 per cent from 2000 to 2015.

A positive feature of the simple average of price indexes is that we would obtain the same value for the index regardless of the units of measurement. In the earlier index, if apples were priced in tonnes, instead of kilogrammes, the impact of apples on the combined index would not change.

A negative feature of this index is that it fails to consider the relative importance of the items included in the index. For example, milk and eggs receive the same weight, even though a typical family might spend far more over the year on milk than on eggs.

Simple Aggregate Index

This method expresses the aggregate of price of all commodities in the current year as a percentage of the aggregate price in the base year.

$$\text{Simple aggregate index} = P = \sum P_1 / \sum P_0 \times 100$$

Example

The index of the food items contained in Table 12.5 can be computed as follows:

$$P = 13.24 / 10.94 \times 100 = 121.0$$

This means that the aggregate group of prices had increased 21 per cent in the 15-year period.

Since the units of measurement can influence the value of a simple aggregate index, it is often not used on a regular basis. In our earlier example, the value of the index would differ significantly if we were to report the price of apples in tonnes rather than kilogrammes. It is also important to note the effect of coffee on the total index. For both the current year and the base year, the value of coffee is about 40 per cent of the total index, so a change in the price of coffee will drive the index much more than any other item. So we need a way to appropriately 'weight' the items according to their relative importance. For this to be addressed, we have to construct the weighted indexes.

Weighted Indexes

These are those indexes in which rational weights are assigned to various chains in an explicit fashion. Two methods of computing a weighted price index are the Laspeyres method and the Paasche method.

Laspeyres' Price Index

In this method, Laspeyres uses base period weights for computing a weighted index. Applying his method, a weighted price index is computed by

$$\text{Laspeyres'price index} = P = \sum p_1q_0 / \sum p_0q_0 \times 100,$$

where

P = the price index;

p_1 = the current price;

p_0 = the price in the base period and

q_0 = the quantity used in the base period.

In Table 12.6, the prices for the six food items taken from Table 12.5 have been repeated. Included in this is the number of units of each consumed item by the university dining hall in 2000 and 2015.

Applying his equation:

$$P = \sum p_1q_0 / \sum p_oq_0$$

$$\text{Laspeyres' index for 2015} = \text{US\$555.94} / \text{US\$336.16} \times 100 = 165.4.$$

Based on this analysis, we conclude that the price of this group of items has increased 65.4 per cent in the 15-year period. The advantage of this method over the simple aggregate index is that the weight of

TABLE 12.6 Computation of Laspeyres' Indexes of Food Price (Base Year: 2000)

		2000		*2015*			
Item	*Unit*	*Price (US$)*	*Quantity*	*Price (US$)*	*Quantity*	P_0Q_0	P_1Q_0
Bread (Britannia)	Pack	0.77	50	1.98	55	38.50	99.00
Eggs	Dozen	1.85	26	1.84	20	48.10	77.48
Milk (Fresh Full Cream)	Litre	0.88	102	1.98	130	89.76	201.96
Apples (Golden)	Kg	1.46	30	1.75	40	43.80	52.50
Jam (Carrefour)	Bottle	1.58	40	1.70	41	63.20	68.00
Coffee (Nestle)	400 g	4.40	12	3.99	12	52.80	57.00
						336.16	**555.94**

Source: Authors.

each of the items is taken into account. In the simple aggregate index, coffee had about 40 per cent of the weight in determining the index. In the Laspeyres' index, the item with the most weight is milk, because the product of the price and the units sold is the largest.

The major disadvantage of the Laspeyres' index is that it assumes that the base period quantities are still realistic in the given period, that is, the quantities used for the six items are about the same in 2000 as 2015. In this case, notice that the quantity of eggs purchased declined by 23 per cent, the quantity of milk increased by nearly 28 per cent and the number of apples increased by 33 per cent.

Paasche's Price Index

This method is an alternate to Laspeyres' method. This method uses the weights of the current period, that is, we use the sum of the products of the 2000 prices and 2015 quantities. If there has been a change in the quantities consumed since the base period, such a change is taken into account in this index. This is shown in Table 12.7.

$$P = \sum p_1 q_1 / \sum p_o q_1$$

TABLE 12.7 Computation of Paasche's Indexes of Food Price (Base Year: 2000)

		2000		*2015*			
Item	*Unit*	*Price (US$)*	*Quantity*	*Price (US$)*	*Quantity*	P_0Q_1	$P_!Q_1$
Bread (Britannia)	Pack	0.77	50	1.98	55	42.35	108.90
Eggs	Dozen	1.85	26	1.84	20	37.00	59.60
Milk (Fresh Full Cream)	Litre	0.88	102	1.98	130	114.40	257.40
Apples (Golden)	Kg	1.46	30	1.75	40	58.40	70.00
Jam (Carrefour)	Bottle	1.58	40	1.70	41	64.78	69.70
Coffee (Nestle)	400 g	4.40	12	3.99	12	52.80	57.00
						369.73	**622.6**

Source: Authors.

$$\text{Paasche's index for } 2015 = \frac{\text{US\$622.6}}{\text{US\$369.73}} \times 100 = 168.4$$

The index shows that in 2015 the package of food costed 68.4 per cent more than it did in 2000. The index is more reflective of the situation in 2015. An important limitation of this index is it tends to overweight goods whose prices have gone down.

Fisher's Ideal Index

This index (Fisher 1922) attempts to offset the shortcomings of both the indexes discussed before. It is the geometric mean of the Laspeyres' and Paasche's indexes. It balances the effects of the two indexes.

However, it is rarely used in practice because it has the same basic set of problems as the Paasche's index. It requires that a new set of quantities be determined for each year.

$$\text{Fisher's ideal index} = \sqrt{(\text{Laspeyres' index})(\text{Paasche's index})}$$

Example

$$\text{Fisher's ideal index} = \sqrt{(165.4)(168.4)} = 166.9$$

Value Index

A value index measures changes in both the price and quantities involved. A value index needs the original base year prices, the original base year quantities, the present year prices and the present year quantities for its construction.

$$V = \sum p_1 q_1 / \sum p_0 q_0 \times 100$$

For example, the prices and quantities sold at a department store for various household items for December 2010 and December 2015 are given in Table 12.8.

TABLE 12.8 Computation of Real Income for 2010 and the Present Year

	2010			2015		
Item	*Price (US$)*	*Quantity Sold ('000)*	*$p_0\ q_0$ (US$ '000)*	*Price (US$)*	*Quantity Sold ('000)*	*$P_1\ q_1$ (US$ '000)*
Grinder (each)	10	1,000	10,000	12	900	10,800
Washing Machine (each)	300	100	30,000	400	120	48,000
Microwave (each)	100	500	50,000	120	500	60,000

Source: Authors.

What is the index of value for December 2015 using December 2010 as the base period?

Total sales in December 2015 were US$118,800,000, and the comparable figure for 2010 is US$90,000,000. Thus, the index of value for December 2015 using 2000 = 100 is 132.0. The value of sales of household items in 2015 was 132.0 per cent of the 2000 sales. To put it another way, the value of sales of household items increased by 32.0 per cent from December 2010 to December 2015.

$$\text{Value Index (2015)} = \sum p_1 q_1 / \sum p_0 q_0 \times 100 = 118{,}800 / 90{,}000 \times 100 = 132.0$$

Consumer Price Index

The CPI is without a doubt the most popular inflation index all over the world. There are several different version of the CPI but they are all built upon the idea of tracking prices for a basket of goods and comparing them to a baseline year. We can construct both simple and weighted price indexes as discussed before.

The items covered by CPI differ from one country to another. However, in majority of countries it covers items of different categories including:

1. Food and beverages
2. Housing
3. Apparel
4. Transportation
5. Medical care
6. Recreation and entertainment
7. Education and communication
8. Other goods and services

Further to measuring changes in prices of consumer goods and services, the CPI is also used to:

- Compute real disposable personal income
- Deflate sales and other series
- Determine the purchasing power of currency
- Establish cost-of-living increases

Computation of Real Income

The meanings of deflated and real income are the same. A popular term for deflated income is 'income expressed in constant dollars'.

$$\text{Real Income} = \frac{\text{Money Income}}{\text{CPI}} \times 100$$

Example

Consider that the CPI is presently 122.3 with 2005 = 100. Now assume that you earned annually US$25,000 in the base period of 2005. Now you have a current income of US$30,575. What was your real income currently? Table 12.9 shows the computation.

The table shows that your 'money income' has increased by 22.3 per cent since the base period of 2005. At the same time, the prices you paid for the goods and services have also increased by 22.3 per cent.

TABLE 12.9 Computation of Real Income for 2005 and the Present Year

Year	*Money Income (US$)*	CPI (2005 = 100)	Real Income *(US$)*	Computation of Real Income (US$)
2010	25,000	100.0	25,000	$\frac{25,000}{100} \times 100$
Present Year	30,575	122.3	25,000	$\frac{30,575}{122.3} \times 100$

Source: Authors.

It means your standard of living has remained the same since 2005 (the base year) to the present time. In other words, price increases have exactly offset an increase in your income, so your present purchasing power (real income) is still US$25,000.

Summary

This chapter has provided an introduction to the presentation of simple descriptive statistics and index number construction methods commonly used in social science research. The treatment has been very general and undertaken at a very basic level. As much of what has been presented constitutes little more than common sense, data analysts should use their own common sense when facing particular issues regarding the analysis of their surveys. In other words, there is not a single hard and fast rule for data analysis but you need to look at your data and decide on the method of data analysis. Data organisation alone cannot help you in drawing conclusions but data analysis helps you in this regard. After analysing data, you get an organised and well-examined form of data that can help you know whether your hypothesis should be accepted or rejected. There are some basic tips you need to follow to analyse data in research papers and dissertations.

More sophisticated methods can also be used to analyse survey data, some of which are discussed in later chapters. All things considered, the data analysis for any given survey will have to be tailored to the main topics and objectives of the survey, and researchers will have to consult specialised books and journals to obtain guidance on issues specific to those topics.

These more sophisticated methods of data analysis are discussed in Chapter 13. This chapter broadly explains the second part of the statistical analysis, that is, the inferential statistical methods to interpret the results and trends, draw inferences from our data and test the research hypothesis.

Self-test Exercises

Exercise 12.1:

12.1.1. There are four girls in Grade 1 of a certain primary school. Their heights in inches are 54, 57, 53 and 52. The teacher finds out that according to national statistics, the average height of a 9-year-old girl is 55 inches or 4 feet and 7 inches. What is the mean or average height of these four girls?

12.1.2. You have four 10 km segments to your automobile trip. You drive your car:

Trip 1 100 km/hr for the first 10 km
Trip 2 110 km/hr for the second 10 km
Trip 3 90 km/hr for the third 10 km
Trip 4 120 km/hr for the fourth 10 km

What is your average speed?

12.1.3. Suppose that if you earn US$25,000 a year today and the inflation rate continues at 3 per cent per annum, you will need to make US$33,598 in 10 years to have the same buying power. Confirm that this statement is accurate by finding the geometric mean rate of increase.

12.1.4. Calculate the median of a statistical distribution given by the following table:

Class	Frequency (F)
60–63	5
63–65	18
66–69	42
69–72	27
72–75	8
TOTAL	100

12.1.5. Calculate the geometric mean of the following series of monthly income of a batch of families 180, 250, 490, 1,400 and 1,050.

12.1.6. The following series relates to the marks secured by students in an examination. Find the quartiles and interpret them.

Marks	No. of Students
0–10	11
10–20	18
20–30	25
30–40	28
40–50	30
50–60	33
60–70	22
70–80	15
80–90	12
90–100	10

Exercise 12.2:

12.2.1. In two primary schools A and B located in the same city, the average scores of Grade 6 students and standard deviations are as follows:

School	Average	Standard Deviation	No. of students
A	34.5	5.0	476
B	28.5	4.5	524

Which school A or B has greater variability?

Exercise 12.3:

12.3.1. If the prices of different milk products in a shop changed by +21; +1.8; +1.4; +1.6 and –0.9 per cent, then the aggregate price index can be ….

A. 123 per cent
B. 91 per cent
C. 107 per cent

12.3.2. The RPI in 2007 was 115.2 and by 2008, it was 138.5 using 2006 as the base year.

This represents a rise in the price of goods by ....

A. 15.0 per cent
B. 20.2 per cent
C. 23.3 per cent
D. 11.8 per cent

Exercise 12.4

12.4.1. In the following table, profit earned is given for number of companies belonging to two areas A and B. Draw in the same diagram their Lorenz curves and interpret them.

Profit Earned ('000 US$)	Number of Companies	
	Area A	Area B
5	7	13
26	12	25
65	14	43
89	28	57
110	33	45
155	25	28
180	18	13
200	8	6

References

Aliaga, M. and B. Gunderson. 2002. *Interactive Statistics*. Thousand Oaks, CA: SAGE.

Andrew, I.D and Kotz, Samuel. 2011. *Arthur L Bowley: A Pioneer in Modern Statistics and Economics*. Singapore: World Scientific Publishing.

Croxton, E.F. and D.J. Cowden. 1946. *Applied General Statistics*. Boston, NY: Prentice Hall.

Dikhanov, Yuri. 2005. *Trends in Global Income Distribution, 1970–2000, and Scenarios for 2015* (Occasional Paper 2005/8). United Nations Human Development Report.

Fisher, I. 1922. *The Making of Index Numbers*. Boston, NY: Houghton Mifflin Company.

Franklin, D. 2007. *Calculating the Quartile, (Or Why Are My Quartile Answers Different?)*. Available at http://www.nesug.org/proceedings/nesug07/po/po08.pdf (accessed on 17 November 2017).

School of Open Learning. Undated. *Measures of Skewness and Kurtosis*. New Delhi: University of Delhi. Available at https://sol.du.ac.in/mod/book/view.php?id=1317&chapterid=1067 (accessed on 12 February 2018).

Weiss, N.A. 1999. *Introductory Statistics*. Boston, NY: Addison Wesley.

13
CHAPTER

Data Analysis: Inferential Methods and Analysis of Time Series

Introduction

Descriptive and inferential statistics are the two most important branches of the science of statistics. We have seen in Chapter 12 that descriptive statistics tell us basic information about the population or data set under study. Descriptive statistics deals with the information about our immediate group of data. For example, we could compute the arithmetic mean and standard deviation (SD) of the exam scores for the 100 students and this could provide valuable information about this group of 100 students. Any group of data like this, which contains all the data you are interested in, is called a 'population'. A population can be small or large, as long as it comprises all the data you are interested in. For example, if you were only interested in the exam scores of 100 students, then 100 students would represent your population. Descriptive statistics are used to study and describe the properties of populations. The arithmetic mean or SD are called 'parameters' as they represent the whole population (i.e., everybody you are interested in). Here the goal is to describe and numerical measures are used to tell about features of a set of data.

We know that often the researchers do not have access to the whole population they are interested in investigating, but only a limited number of data instead. For example, they might be interested in the exam scores of all university students in a given country X. It is neither feasible nor practical to measure all exam scores of all students in the whole of country X. Thus, they have to measure a smaller sample of students (e.g., 100 students), who can be considered as representatives of the larger population of all students in country X. Most quantitative in social sciences operates using inferential statistics for the simple reason, as it is typically too costly or time-consuming to study an entire population of people. We can conduct research that otherwise would not be possible by seeing a statistically valid sample and inferential statistics.

Inferential statistics are generated by more complex mathematical calculations. They allow the researcher to infer trends about a larger population based on a study of a sample taken from it. In social sciences, inferential statistics are used extensively to examine the relationships between variables within a sample and then make generalisations or predictions about how those variables will relate within a larger population.

Researchers use inferential statistics to draw inferences about a sample of participants taken from the population under study. For example, you might attempt to conclude something about a population of students (e.g., all 4th graders in a school district) by studying a sample of these students. Depending on inferential statistics, you can infer that the results from the sample of 4th graders are equally true of the population of 4th graders. Inferential statistics are also used to make inferences about the differences between two or more groups of observations.

This chapter explains simple techniques such as correlation, linear regression analysis, ANOVA and analysis of covariance (ANCOVA) that social scientists use to examine the relationships between variables and thereby to create inferential statistics. If you are intending to use many of the multivariate methods such as factor analysis, multidimensional scaling, cluster analysis, discriminant function analysis, logistic regression analyses, stepwise regression analysis, structural equation modelling and survival analysis and so on. You are advised to refer to textbooks on advanced statistical methods. In particular, the chapter answers the following questions:

- What are inferential statistics?
- Why learn about inferential statistics?
- Why use inferential statistics?
- When are inferential statistics utilised?
- Which types of inferential statistics are most commonly used and when?
- What is important for you to know about inferential statistics?

What are Inferential Statistics?

As seen in Chapter 12, descriptive statistics is concerned with telling about certain features of a data set. Although this is helpful in learning things such as the spread and centre of the data you are studying, nothing in the area of descriptive statistics can be used to make any sort of generalisation. In descriptive statistics, measurements such as the mean and SD are expressed as exact numbers. Though you may use descriptive statistics all you would like in examining a statistical sample, this branch of statistics does not allow us to say anything about the population.

With inferential statistics, our aim is trying to reach conclusions that extend beyond the immediate data alone. For instance, we use inferential statistics to conclude from the sample data what the population might think. Or we use inferential statistics to construct judgements of the probability that an observed difference between groups is a dependable and might have happened by chance in this study. Thus, we use inferential statistics to draw inferences from our data to more general conditions; we use descriptive statistics simply to describe what is going on in our data.

> 'Inferential statistics provide ways of testing the reliability of the findings of a study and "inferring" characteristics from a small group of participants or people (your sample) onto much larger groups of people (the population)' (Psychology Glossary).

> 'Inferential statistics is the branch of statistics, which derive conclusions about the concerned population from the data set obtained from a sample subjected to random, observational, and sampling variations' (Croxton and Crowden 1946).

The focus of inferential statistics is on how to generalise the statistics obtained from a sample as accurately as possible to represent the population. An important factor that we have to keep in mind or a matter of concern is the nature of the sample. If the sample is biased, then the results are also biased, and the parameters based on these will not represent the whole population correctly. Therefore, sampling is the key study of inferential statistics. Statistical assumptions, statistical decision theory and estimation theory, hypothesis testing, design of experiments, ANOVA and analysis of regression are prominent topics of study in the theory of inferential statistics. A good example of the use of inferential statistics is to predict the results of an election prior to the voting by means of polling.

In summary:

- Inferential statistics infer from the sample to the population.
- They determine probability of characteristics of population based on the characteristics of the sample.
- They help assess strength of relationship between our independent (causal) variables and our dependent (effect variables).

Whereas descriptive statistics attempts to describe:

- How widely dispersed are a given data?
- Are there a lot of different values? Or are many of the values the same?
- What value is in the middle of a given data?
- Where does a particular data value stand with respect with the other values in the data set?

Why Learn about Inferential Statistics?

Whenever you plan to conduct research using inferential statistics, it is important and necessary that you conduct a test of significance in order to know and decide whether you can generalise your results to a larger population. Common tests of significance are the chi-square test and *t*-test. These two tests tell us the probability that the results of our analysis of the sample are representative of the population that the sample represents. According to Carol Albrecht, a Utah State University extension specialist, following are the reasons to learn about inferential statistics:

- Before you use any intervention, you should do some research and determine if there is evidence that it works (i.e., does the head start programme increase educational performance for low-income children?).
- Before you work with any group, you want to base your judgements on research, not on stereotypes (i.e., you may want to know what proportion of Latino boys join gangs?).
- Before you make recommendations, you want to understand the probabilities of success (i.e., what is the probability that a child will have success in school if they participate in your tutorial programme?).
- Before you continue on with a programme/intervention, you want to reassure yourself that this programme is worth your time and effort.
- As you apply for grants, you want to ensure the grantees that you can implement an evidence-based programme.
- When making policy recommendations or participating in political advocacy, you want to provide empirical support that your intervention actually works.

Why Use Inferential Statistics?

Both descriptive and inferential statistics have their benefits and shortcomings. In most cases, researchers use an inferential statistic to test some hypothesis. Do groups differ on some outcome variable? Is the difference more than would be expected by chance? Can one factor predict another? You do not need to understand the underlying calculus, but you do need to know which inferential statistic to use and how to interpret it.

With inferential statistics, you do not need the data of the entire population to make your conclusions; this level of statistics only needs accurate samples in order for you to draw your conclusions. You can build an educated guess on what the parameters of the entire population are, no matter how large it may be. In other words, inferential statistics allows you to generalise your findings to the larger population. It can determine not just what 'can' happen, but 'what tends to happen' in programmes like yours. It helps assess strength of the relationship between your independent (causal) variables and your dependent (effect variable. Finally, inferential statistics assess the relative impact of various programme inputs on your programme outcomes/objectives.

Unfortunately, this does prevent you from having accurate data. Although inferential statistics does give you a good guess of what the data may look like, it does not compare to the accuracy that you will get with something more concrete, as with descriptive statistics.

Thus, both forms of statistics are great and you should use both types of statistical methods in unison to get accurate parameters of a small population and then take those parameters further and obtain great approximations of what a much larger population's statistics are.

When to Use Inferential Statistics?

For using inferential statistical methods for analysing your survey data, you should ensure that the following three conditions are met:

- You should have a complete list of the members of population.
- You should draw a random sample from this population.
- You should use a pre-established formula and determine that your sample size is large enough and represents the population.

Now the question that can come to your mind: 'Can I use inferential statistics even if my data do not meet the above three prerequisites?' With inferential statistics, you can determine the strength of relationship within your sample. In other words, you can assess the strength of the impact of your independent variables (programme inputs) on your outcomes (programme outputs). If you find it difficult to obtain a population list and/or draw a random sample, then you do the best you can with what you have. In this case, you can use both descriptive and inferential statistics.

The following types of inferential statistics are relatively common and easy to interpret.

1. One sample test of difference/one sample hypothesis test
2. Confidence interval
3. Contingency tables and chi-square statistics
4. *t*-test or ANOVA
5. Pearson correlation

6. Bivariate regression analysis
7. Multivariate regression analysis

We have already discussed methods 1 to 3 in Chapter 5 on hypothesis testing. In this chapter, we discuss methods 4 to 7 in detail to enable you better understand these methods of inferential statistics.

Univariate Statistics—One Sample Hypothesis Test

You use this method when your aim is to (a) compare responses of respondents of your study/programme on a pre- and post-test and (b) determine if implemented programme had an impact on one particular outcome. Its interpretation is quite simple. For instance, if the probability is 0.05 or less than that, you will make a mistake in asserting there is a difference between the pre- and post-test scores in the population, then you can assert that the programme did make a difference on this outcome. In other words, your programme is working.

Univariate Statistics—Confidence Interval

This is used to determine a value/score in a population based on the score of the participants in your sample. Its interpretation is relatively simple. For instance, a 95 per cent confidence interval indicates you are 95 per cent confident that you can predict/infer the value/score of a population within a specified range based on the value/score of your sample.

Bivariate Statistics—Contingency Tables and Chi-square Statistics

You use this method to (a) analyse two categorical variables and (b) to know if they are related and the strength of the relationship. For instance, you will use this method when your aim is to see the gender and score on outcome measurement. Suppose you want to know: Does knowing someone's gender help you predict their outcome score/value? In this case, if the probability associated with the chi-square statistics is 0.05 or less, then you can assert that the independent variable can be used to predict scores on the dependent or outcome variable.

You can also use the contingency table to compare the actual scores across the independent variable on the dependent variable or outcome measurement (i.e., compare the number/per cent of males who agreed that the programme had positive impact on their live to the per cent of females who agreed).

*Bivariate Statistics—*t*-test or ANOVA*

You can use this method when you have (a) a categorical and continuous variable, and when you want to (b) compare mean scores of two or more groups (i.e., you want to compare mean guided reading practice [GRP] of students you have tutored across race). For its interpretation, you will use either *t*-test or *f*-statistics to determine if the groups have significantly different means. If the probability associated with the *f*-statistics is 0.05 or less, then you can assert that there is a difference in the means.

Bivariate Statistics—Pearson Correlation

We will explain further this method of analysing statistics in detail. You can use Pearson correlation method when you have a continuous independent variable and a continuous dependent variable. Its interpretation is relatively simple. For instance, when the probability associated with *t*-statistics is 0.05 or less, then you can assume there is a relationship between the dependent and independent variable. Suppose you want to know if the number of hours student spend in your programme is positively related to their scores on school exams, you will compute the Pearson correlation coefficient.

Bivariate Statistics—Regression Analysis

Regression analysis is used when you have a continuous independent variable and a continuous dependent (outcome) variable. For instance, you may want to know if the number of hours participants spend in your programme is positively related to their scores on school exams. The interpretation will be as follows. When the probability associated with *f*-statistics is 0.05 or less, then you can assume there is a relationship between the dependent and the independent variable. It is important to note that the Pearson correlation and bivariate regression are very similar.

Multivariate Statistics—Elaborated Chi-square Statistics

This method is used when you have more than one independent categorical variables and one dependent categorical variable. For interpretation, you divide one of the independent variables into two groups and then do chi-square for each group (i.e., divide gender into boy and girl students, then do a chi-square of boys and one for girls. So for girls you can do a chi-square test of outcome measurement by race and then do the same for boys).

Multivariate Statistics—Multivariate Regression

You can use multivariate regression when you have more than one independent (casual) variable and one dependent (effect/outcome) variable. Here your aim is not only to know if your intervention has an impact on the outcome, but you also want to know 'which' aspects of your intervention has an impact and/or the relative impact of different aspects of your intervention. For interpretation, if the probability associated with *f*-statistics or *t*-statistics is 0.05 or less, then you can assert that independent variable has an impact on the outcome, independent of the other variables. The value of *t*-statistics can be compared across the independent variables to determine the relative value of each independent variable.

In short, the important things that you should know about inferential statistics are as follows. You should be able to:

- Read and understand all the computed values as shown in computer printouts
- Construct tables and graphs from the computer printouts
- Interpret and explain these tables and graphs to an audience
- Make wise and intelligent decisions based on valid and accurate data

Descriptive and Inferential Statistics—The Difference

When you conduct a research study, you are quite often confused between descriptive and inferential statistics, making it hard for yourself to distinguish the best option to use in your research. If you look closely, the difference between descriptive and inferential statistics is already pretty obvious in their names. 'Descriptive' describes data, while 'inferential' infers or allows you to arrive at a conclusion based on the collected information.

Suppose you are tasked to research about teenage pregnancy in a high school, say High School X. By using both descriptive and inferential statistics, you will be researching the number of teenage pregnancy cases in the school for a specific number of years. The difference is that with descriptive statistics, you are merely summarising the collected data and, if possible, detecting a pattern in the changes. For example, it can be said that for the past five years, the majority of teenage pregnancies in High School X happened to those enrolled in the third-year level. There is no need to predict that on the sixth year, the third-year level would still be the ones that would have the larger case of teenage pregnancies. Conclusions as well as predictions are only done for inferential statistics.

The difference between descriptive and inferential statistics can be summarised as follows:

- Descriptive statistics merely 'describes' research and does not allow for conclusions or predictions.
- Inferential statistics makes it possible for the researcher to arrive at a conclusion and predict changes that may occur regarding the area of concern.
- Descriptive statistics usually operate within a specific area that contains 'all' the target population.
- Inferential statistics usually takes a sample of a population especially if the population is too big to conduct research.
- Descriptive statistics is focused on summarising the data collected from a sample. The technique produces measures of central tendency and dispersion which represent how the values of the variables are concentrated and dispersed.
- Inferential statistics generalises the statistics obtained from a sample to the general population to which the sample belongs. The measures of the population are termed as parameters.
- Descriptive statistics make only summarisation of the properties of the sample from which data were acquired, but in inferential statistics, the measure from the sample is used to infer properties of the population.
- In inferential statistics, the parameters were obtained from a sample, but not the whole population; therefore, always some uncertainty exists compared to the real values.
- Inferential statistics starts with a hypothesis (a statement of, or a conjecture about, the relationship between two or more variables that you intend to study) and investigates whether the data are consistent with that hypothesis.

Although descriptive and inferential statistics both are used for purpose of analysis of the data, still both of them are different in various ways.

Limitations of Inferential Statistics

There are two key limitations to the use of inferential statistics. The first, and most important limitation, which is inherent in all inferential statistics, is that you are providing data only about a part of the population, that is, of the population that you have not fully measured. Therefore, you cannot ever be completely sure that the values/statistics you have calculated are correct. You should remember that inferential statistics are based on the concept of using the values measured in a sample to estimate/infer the values that would be measured in a population; there will always be a degree of uncertainty in doing this. The second limitation is associated with the first limitation. Some, but not all, inferential tests require the researcher to make educated guesses (based on theory) to run the inferential tests. Again, there will be some uncertainty in this process, which will have repercussions on the certainty of the results of some inferential statistics.

Data Analysis—Inferential Statistics

In Chapter 12, we examined numerical methods used to describe various characteristics of univariate data (descriptive statistics), that is, the data involving only one variable. You may recall that in univariate data, only one variable is associated with each unit of observation. However, we may have data in which more than one variable can be associated with each unit of observation. For instance, for providing information about the marks obtained by the students of a class in two subjects, say economics and statistics, you can associate two variables, one representing the marks in economics and the other marks in statistics to each unit of observation, namely, a student in the class. When we have two variables for which values are being observed for each unit of observation, we say that we have 'bivariate data'. In general, the study of those data which involves more than two variables is termed 'multivariate data'.

The most commonly used statistical methods to analyse bivariate and multivariate data (inferential statistics) include:

- Correlation
- Linear regression
- ANOVA

In the following paragraphs, explain these statistical methods. For more advanced methods, you are advised to refer to any standard textbook on statistical methods.

Correlation

Correlation measures the relationship between two variables. It is the most commonly used statistical technique to identify and determine the relationship between two continuous variables. By comparing numbers from two different data sets together, correlations glance at how movement in the value of numbers in one data set is related to movement in the value of numbers in the other data set. Thus, the main purpose of doing correlations is to allow researchers to make a prediction about one variable based on what is known about another variable.

> 'Correlation is a statistical technique that is used to measure and describe the "strength" and "direction" of the relationship between two variables'.

For example, evidence suggests that there is a correlation between income and education. We know that people with higher income have more years of education (you can also phrase this as people with more years of education have higher income). When it is observed that there is a correlation between these two variables, it enables you to make a prediction. If you know a group's income, for example, then you can calculate their years of education.

The value of the correlation coefficient lies between –1 to 1, where 1 shows that two variables move in an exactly same direction, that is, if one variable increases, the other variable also increases and if one variable decreases other variable also decreases.

Correlation requires two scores from the 'same' individuals. These scores are normally labelled as *X* and *Y*. You can also score pairs and can list them in a table or present in a scatter plot. Usually the two variables are observed, not manipulated.

Before we discuss further the properties of the coefficient of correlation, it is worth mentioning here what a scatter plot is.

> 'A scatterplot is a statistical graphic that displays the "strength," "direction" and "shape" of the relationship between two variables'.

A scatter plot displays the *X* variable on the horizontal (*X*) axis and the *Y* variable on the vertical (*Y*) axis. Each individual unit is identified by a single point (dot) on the graph which is located so that the coordinates of the point (the *X* and *Y* values) match the individual's *X* and *Y* scores.

Figure 13.1 shows various examples of correlation coefficients.

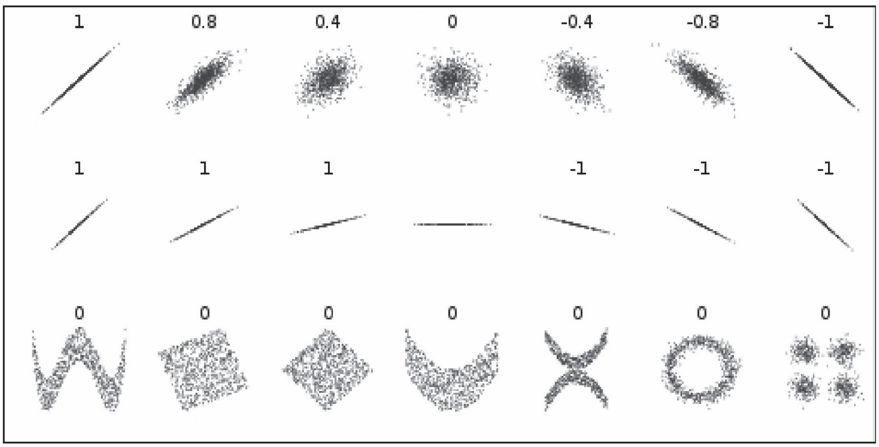


FIGURE 13.1 Examples of Correlation Coefficient

Source: Wikipedia.

The form or shape of a relationship refers to whether the relationship is straight or curved.

- Linear: A straight relationship is termed as linear, because it approximates a straight line. The relationship is said to be linear if the amount of change in one variable tends to bear a constant ratio to the amount of change in the other variable. For example, consider the following data:

Variable *X*:	10	20	30	40	50
Variable *Y*:	40	80	120	160	200

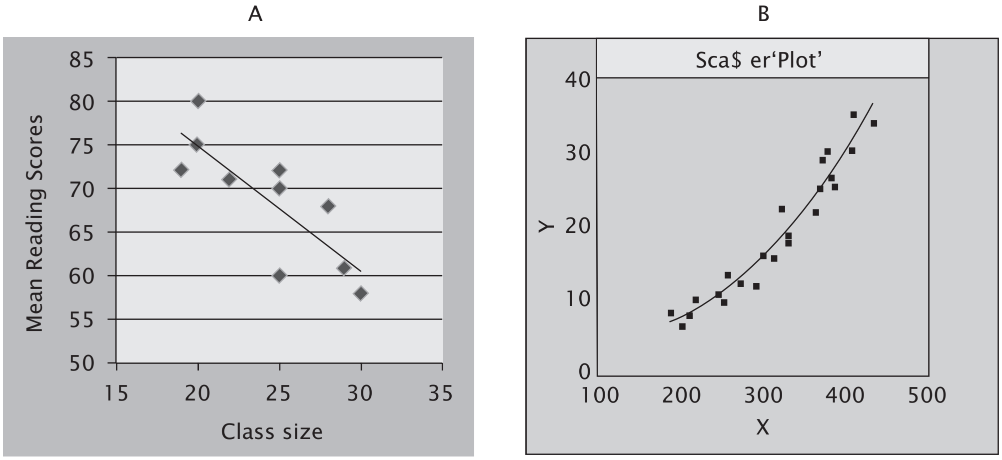


FIGURE 13.2 Scatter Plots

Source: Authors.

In this example, you can observe that the ratio of changes between the two variables is same and hence the correlation between *X* and *Y* is linear. It may be remarked that if the values of two variables, which are linearly correlated, are plotted on a graph paper all the plotted points would be on a straight line (linear line).

- Curvilinear: A curved relationship is called curvilinear, because it approximates a curved line. For curvilinear or non-linear relationship, the amount of change in one variable does not bear a constant ratio to the amount of change in the other variable. For example, if the amount of fertiliser used is doubled, the production of wheat may not necessarily be doubled. Then, you can say that the relation between the amount of fertiliser and the production of wheat is non-linear.

Figures 13.2(A) and 13.2(B) show the linear relationship and curvilinear relationship, respectively. The curvilinear relationship in Figure 13.2(B) is positive.

The correlated variables can move in the same direction or they can move in opposite directions. This leads to the concept of correlation type.

Types of Correlation

Broadly speaking, correlations can be classified into seven types. They are shown in Figure 13.3.

Positive correlation: When two variables move in the same direction, then the correlation between these two variables is said to be positive (Figure 13.3A).

- Negative correlation: Negative correlation occurs when an increase in the value of one variable leads to a decrease in the value of the other (Figure 13.3B).
- Strong correlation: A correlation is said to be stronger when the points in the scatter plot are closely located to one another along a straight line (Figure 13.3C).

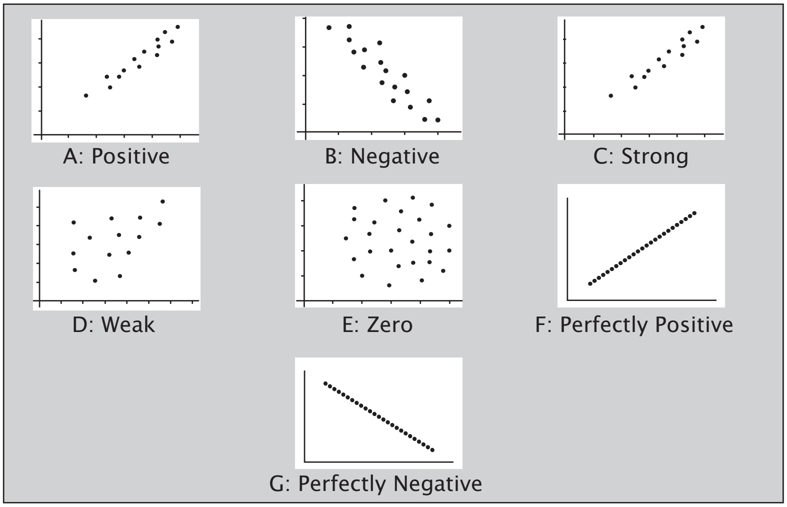


FIGURE 13.3 Types of Correlations

Source: Authors.

- Weak correlation: A correlation is said to be weaker when the points are located farther apart along a straight line (Figure 13.3D).
- Zero correlation: When the two variables are independent and the change in one variable has no effect on the other variable, then the correlation between these two variables is known as zero correlation. In zero correlations, there is either (a) no meaningful relationship between the two variables (the association is a random one) or (b) the relationship is other than linear (i.e., a curvilinear relationship). As you will see in a later section, this implies a correlation coefficient equal to '0'.
- Perfectly positive correlation: When a change in one variable, say *X*, always induces a change in the other variable, say *Y*, in the same direction, then these two variables have a perfectly positive correlation. A perfectly positive correlation means that for 100 per cent of the time, the relationship that appears to exist between two variables is positive. In this case, all the dots on the scatter plot lie on a straight line.
- Perfectly negative correlation: When the relationship between two variables *X* and *Y* is such that a change in *X* always induces causes a change in *Y* in the opposite direction, then the two variable have a perfectly negative correlation. A perfectly negative correlation means that the relationship that appears to exist between two variables is negative 100 per cent of the time.

Correlation is also classified on the basis of number of variables involved in the study.

- Simple correlation: It involves the study of only two variables, that is, in simple correlation, we measure the degree of association between two variables only. For example, when we study the correlation between the price and supply of a commodity, it is a problem of simple correlation.
- Partial correlation: It involves the study of three variables, but considers only two variables to be influencing each other, the effect of third variable being kept constant. Thus, in partial

correlation, we measure the degree of relationship between the variable Y and the variable X_1 with the effect of X_2 variable removed. For example, if we consider three variables, namely, yield of wheat, amount of rainfall and amount of fertiliser and limit our correlation analysis to yield and rainfall, with the effect of fertiliser removed, it becomes a problem relating to partial correlation only.

- Multiple correlation: It involves the study of three or more variables simultaneously. Thus, in multiple correlation we measure the degree of relationship between the variable Y and all the variables $X_1, X_2, \ldots, X_n$ taken together. For example, we study the relationship between the yield of wheat per acre of land and both amount of rainfall and the amount of fertiliser used, it becomes the problem relating to multiple correlation.

Correlation and Causation

You have noticed that correlation measures the degree of association or relationship between two or more variables, but it fails to explain anything about cause and effect between two or more variables. Even a high degree of correlation between two variables does not necessarily imply causation or functional relationship between the variables though the existence of causation always implies correlation. The high degree of correlation between the variables may be due to the following reasons:

- Mutual dependence: At times, it becomes difficult to find out from two correlated variables which is the cause and which is the effect.
- Influence of third variable: A high degree of correlation between two variables may be due to effect or influence of a third variable or a number of other variables.
- Pure chance: It is just possible that a high degree of correlation between two variables may arise either because of pure random sampling variation or because of the bias of the investigator in selecting the sample. For example, we may observe a high degree of correlation between the incomes and heights of a group of persons. Such a correlation is called 'spurious' or 'nonsense' correlation.

Uses of Correlation

Correlations are used for prediction, validity, reliability and verification.

Prediction

An important use of correlation is prediction. To predict the values of the dependent or criterion variable (variable Y) based on the values for the predictor variable (variable X), there is another mathematical technique called 'regression analysis', which makes extensive use of correlation analysis. Figure 13.3(A) highlights a linear relationship between variable X and variable Y.

A regression analysis can identify the equation that best describes the linear relationship between class size and reading score in the graph (Figure 13.4). You can use the regression equation for estimating the mean reading scores based on class sizes. Unless there is a perfect correlation between two variables (i.e., $r = \pm 1.00$), the prediction based on regression analysis will be imperfect. The standard error of

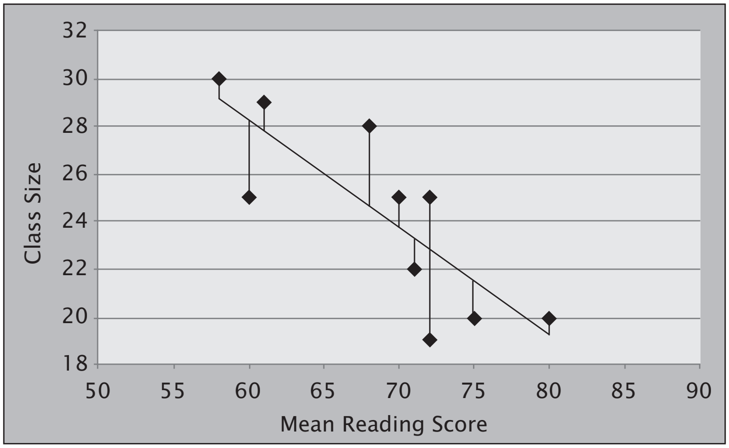


FIGURE 13.4 Linear Relationship Between Two Variables

Source: Authors.

estimate depicts how accurately the equation can predict values of a variable. In the example, the standard error of estimate in Figure 13.4 is 4.44, which is the average distance between the line in a graph of the regression equation and the actual data points for the mean reading score.

Another simple way to express the prediction error is that the smaller the numerical value of the correlation, the smaller would be the coefficient of determination, and the more error there will be when using the correlation for prediction.

Correlations can be used to help make predictions. If two variables have been known in the past to correlate, then we can assume they will continue to correlate in the future. We can use the value of one variable that is known now to predict the value that the other variable will take on in the future.

For example, we require high school students to take SAT (a standardised test widely used for college admissions; ViSta: The Visual Statistics System) exam because we know that in the past SAT scores correlated well with the grade point average (GPA) scores that the students get when they are in college. Thus, we predict high SAT scores will lead to high GPA scores and conversely.

Validity

Suppose we have designed a new test of intelligence. With this test, we can determine if it is really measuring intelligence by correlating the new test's scores with, for example, the scores that the same people get on standardised IQ tests, or their scores on problem-solving ability tests, or their performance on learning tasks and so on.

This is a process for validating the new test of intelligence. The process is based on correlation.

Reliability

We can use correlations to determine the reliability of some measurement process. For example, we could administer our new IQ test on two different occasions to the same group of people and see what the correlation is. If the correlation is high, the test is reliable. If it is low, it is not.

Theory Verification

There are several psychological theories which make specific predictions about the relationship between two variables. For example, it is predicted that parents and children's intelligences are positively related. We can test this prediction by administering intelligence tests (IQ tests) to the parents and their children and determining the correlation between the two scores.

Statistical Significance of the Correlation

A statistically significant correlation is shown by a probability value of less than 0.05. This means that the probability of obtaining such a correlation coefficient by chance is less than five times out of 100, so the result indicates the presence of a relationship. In Table 13.1, there is a statistically significant positive relationship between English and mathematic scores ($p < 0.001$), such that the probability of this correlation occurring by chance is less than one time out of 1,000.

Effect Size of the Correlation—The Coefficient of Determination

For correlations, the effect size is called the 'coefficient of determination' and is defined as r^2. The value of coefficient of determination can be anywhere from 0 to 1.00. The coefficient of determination shows that the proportion of variation in the scores can be predicted from the relationship between the two variables. In Table 13.1, the coefficient of determination (r^2) is 0.66 ($0.81^2 = 0.66$), which means that 66 per cent of the variation in the mean reading scores among the different classes can be foreseen from the relationship between class size and reading scores (conversely, 35% of the variation in mean reading scores cannot be explained).

TABLE 13.1 Class Size and the Mean Reading Scores

Variable 1 Score in English Language (x)	*Variable 2 Score in Mathematics (y)*	$X = x - \bar{x}$ $X = x - 67$	X^2	$Y = y - \bar{y}$ $Y = y - 68$	Y^2	XY
64	66	–3	9	–2	4	6
65	67	–2	4	–1	1	2
66	65	–1	1	–3	9	3
67	68	0	0	0	0	0
68	70	1	1	2	4	2
69	68	2	4	0	0	0
70	72	3	9	4	16	12
$\sum x = 469$	$\sum y = 476$	0	28	0	34	25

Source: Authors.

A correlation simply indicates the presence or absence of a relationship but not the nature of the relationship. For example, it cannot be concluded that smaller class sizes cause higher reading scores, even if the correlation is 1.0. Correlation is not causation. There is always the possibility that a third variable might exert influence on the results. For example, perhaps the students in the small classes were higher in verbal ability than the students in the large classes or they were from higher income families or had higher quality teachers.

Calculation of Correlation

When there exists some relationship between two variables, we have to measure the degree of relationship. This measure is called the measure of correlation or correlation coefficient and it is shown by r.

The co-variation (COV) between the variables X and Y is defined as:

$$\text{COV}(x, y) = \sum(x - \bar{x})(y - \bar{y}) / n$$

Where $\bar{x}$, $\bar{y}$ are respectively means of x and y, and n is the number of pairs of observations.

Karl Pearson's Coefficient of Correlation

This is the most widely used method for measuring the magnitude of linear relationship between the two variables. It is known as Pearsonian coefficient of correlation. It is denoted by r. The formula for calculating r is:

$$r = \text{COV}(x, y) / \sigma_x \times \sigma_y$$

Where

σ_x and σ_y represent of x and y, respectively.

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{\left[n\sum x^2 - (\sum(x^2)\right]\left[n\sum y^2 - (\sum y)^2\right]}}$$

We have noticed above that

$$r = \frac{\sum xy}{n\sigma_x \times \sigma_y}$$

$$r = \frac{\sum XY}{\sqrt{\sum X^2 \times \sum Y^2}}$$

Where

$$X = x - \bar{x} \text{ and } Y = y - \bar{y}$$

The formula is easy to calculate, and it is not necessary to calculate the SDs of x and y series, respectively.

Example

A small study is conducted involving seven student of Grade 12 to investigate the association between mean reading scores (out of 100 marks) in mathematics and English language in a school. This is shown in Table 13.1.

$$\bar{x} = 469/7 = 67; \bar{y} = 476/7 = 68$$

$$r = \frac{\sum XY}{\sqrt{\sum X^2 \times \sum Y^2}} = \frac{25}{\sqrt{28 \times 34}} = \frac{25}{\sqrt{952}} = \frac{25}{30.85} = +0.81$$

In Table 13.1, each point represents a (X, Y) pair (in this case the score in English language and mathematics of a student). Since $r = +0.81$, the two variables, X and Y (English and math), are highly positively correlated. It means pupils with higher marks in math have higher marks in English language.

The coefficient of correlation can also be calculated as follows for data in Table 13.1. The results are shown in table 13.2.

TABLE 13.2 Calculation of Coefficient of Correlation

Variable 1 Score in English Language (x)	*Variable 2 Score in Mathematics (y)*	X^2	Y^2	*XY*
64	66	4,096	4,356	4,224
65	67	4,225	4,489	4,355
66	65	4,356	4,225	4,290
67	68	4,489	4,624	4,556
68	70	4,624	4,900	4,760
69	68	4,761	4,624	4,692
70	72	4,900	5,184	5,040
$\Sigma x = 469$	$\sum y = 476$	$\sum 31{,}451$	$\sum 32{,}402$	$\sum 31{,}917$

Source: Authors.

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{\left[n\sum x^2 - (\sum (x)^2\right]\left[n\sum y^2 - (\sum y)^2\right]}}$$

$$r = \frac{7 * 31{,}917 - (469)(476)}{\sqrt{\left[(7 \times 31{,}451) - (469)^2 \times (7 \times 32{,}402) - (476)^2\right]}}$$

$$r = \frac{223{,}419 - 223{,}244}{\sqrt{(220{,}157 - 219{,}961) \times (226{,}814 - 226{,}576)}}$$

$$r = \frac{175}{\sqrt{(196)(238)}} = \frac{175}{216} = +0.81$$

Coefficient of Determination $(r)^2 = (0.81)^2 = 0.66$.

Both equations yield the same result. So would be the case if we use other formulas. You can calculate the value r by using SPSS and EXCEL.

Spearman's Rank Correlation Coefficient

Sometimes, we are given a series of items where no numerical measure can be made, but where best and worst or most favoured and least favoured can be identified. Rankings are often applied in these situations to put the series into an order. For example, the characteristics such as beauty, intelligence, leadership ability, honesty and so on cannot be measured numerically. The individuals in the group can be arranged in order thereby for each individual a number indicating its rank in the group.

If we have a group of individuals ranked according to two different qualities, it is natural to ask the following question:

'Is there an association between the rankings?'

To answer this question, we use Spearman rank correlation when we have two ranked variables, and we want to see whether the two variables covary; whether, as one variable increases, the other variable tends to increase or decrease. We also use Spearman rank correlation if we have one measurement variable and one ranked variable. In this case, we convert the measurement variable to ranks and use Spearman rank correlation on the two sets of ranks.

This is a method that measures the correlation between two variables recorded from the same sample. It works on a scale from –1 to +1: '–1' means perfect negative correlation. 'Zero' means no correlation, and '+1' means perfect positive correlation.

$$\rho = 1 - \frac{6\sum D^2}{n^3 - n}$$

Where

ρ = rank correlation coefficient

ΣD^2 = sum of squares (SS) of differences between the pairs of ranks

n = number of pairs of observations

Example

The following data show the scores of 10 students in a physics test and a chemistry test.

Physics	15	36	33	22	35	27	36	34	23	24
Chemistry	21	37	35	25	33	30	39	36	27	20

When arranged in accordance to rank, we get the following:

Physics	Rank Physics	Chemistry	Rank Chemistry	D	D^2
15	1	21	2	1	1
22	2	25	3	1	1
23	3	27	4	1	1
24	4	20	1	3	9
27	5	30	5	0	0
33	6	35	7	1	1

34	7	36	8	1	1
35	8	33	6	2	4
36	9.5	39	10	0.5	0.25
36	9.5	37	9	0.5	0.25
					$\sum D^2 = 18.5$

$$\rho = 1 - \frac{6\sum D^2}{n^3 - n} = 1 = \frac{6\times 18.5}{10^3 - 10} = 1 - \frac{111}{990} = +0.888$$

This suggests that there is a strong positive correlation between the data. It is important to note that if two data values are the same, for instance, if rank 9 and rank 10 are both 36, they would both be ranked 9.5.

As we noted, sample correlation coefficients range from –1 to +1. In practice, meaningful correlations (i.e., correlations that are clinically or practically important) can be as small as 0.4 (or –0.4) for positive (or negative) associations. There are also statistical tests to determine whether an observed correlation is statistically significant or not (i.e., statistically significantly different from zero). Procedures to test whether an observed sample correlation is suggestive of a statistically significant correlation are described in detail in Chapter 5.

In conclusion, a correlation only reveals a relationship; it cannot offer a categorical basis for showing why a relationship exists. A correlative result does not uncover which variable has a significant or prominent influence over the other. For instance, a strong correlation between wealth and education does not elucidate whether having wealth leads to more education, or whether having an education leads to more wealth. No causes for either can be entailed, but until more investigation is completed, causation cannot be ascertained. Moreover, there is every possibility that a third unidentified variable might be the underlying cause affecting both these phenomena. For instance, residing in Shanghai or Delhi could as well lead to both wealth and education.

Therefore, while interpreting correlations you should be highly vigilant, especially where new disciplines are concerned. This has led to signify that 'Correlation does not necessarily mean Causation (Hazarika 2013)'. A strong correlation between data does not mean that one set of data is automatically producing the effect that is cropping up in the other set of data. For example, there may be a high correlation between the absorption rate of a prescribed medication by patients belonging to a particular age group in Hospital A, and the blood pressures of patients in the same age group in Hospital B. So does this mean that the procedures followed on patients in Hospital A have an effect on patients in Hospital B? Certainly not! You should always be cautious about what inference you can draw from correlational statistical analyses. You should make sure and discover that the relationship is logical. In addition, it is a good idea to keep in mind that other factors may be involved in a cause-effect relationship.

Regression Analysis

'Regression analysis' attempts to establish the 'nature of relationship' between variables, that is, to study the functional relationship between the variables and thereby provide a mechanism for prediction or forecasting.

Regression analysis is a statistical tool widely used for exploring relationships between variables. Usually, the researcher attempts to ascertain the causal effect of one variable upon another, for example, the effect of a price increase upon demand or the effect of changes in the supply of money upon the inflation rate. To explore such issues, the researcher gathers data on the variables of interest and employs regression to estimate the quantitative effect of the causal variables upon the variable that they influence. The researcher also typically evaluates the 'statistical significance' of the estimated relationships, that is, the degree of confidence that the true relationship is close to the estimated relationship.

A mathematical equation that allows us to predict value of one variable from known values of one or more other variables is called a 'regression equation'. The variable whose value is to be predicted is called the 'dependent variable' or 'explained variable'. The variables which are used to predict the values of a dependent variable are called 'independent variables' or 'explanatory variables'. The regression analysis confined to the study of only two variables, a dependent variable and an independent variable, is called 'simple regression analysis'. When the relationship between the dependent and the independent variable is linear, the technique for prediction is called 'simple regression analysis'.

Meaning of Regression

> 'Regression analysis is a statistical approach to forecasting change in a dependent variable on the basis of change in one or more independent variables' (Sykes 1993).

Suppose that we wish to identify and quantify the factors that determine teachers' salaries in an education setting. If we reflect a little bit, we will realise that there are several factors that are associated with variations in salaries across teachers—age, experience, educational attainment, training, motivation and innate ability come to mind, perhaps along with factors such as race and gender that can be of particular concern to educational researchers. For the time being, let us restrict our attention to a single factor—call it education. Regression analysis with a single explanatory variable is termed 'simple regression'.

Simple Regression

In any regression analysis, a researcher formulates some hypothesis about the relationship between the variables of interest, here in our example, level of education attained and teachers' salaries. Common knowledge and experience suggest that better educated people tend to make more money and the causal relation runs from education to earnings rather than the other way around. Thus, the tentative hypothesis could be 'higher levels of education cause higher levels of earnings, other things remaining the same'.

To investigate this hypothesis, you have to collect data on teachers' education and salaries for various teachers. Let 'E' denote education in years of schooling for each individual teacher, and let 'S' denote individual teacher's salary in dollars per month. We can plot this information for all of the individual teachers in the sample using a two-dimensional diagram conventionally termed a 'scatter' diagram. Each point in the diagram shows an individual teacher in the sample. This is shown in Figure 13.5.

Figure 13.5 indeed suggests that higher values of 'E' tend to yield higher values of 'S', but the relationship is not perfect—it seems that years of schooling (E) does not suffice for an entirely accurate

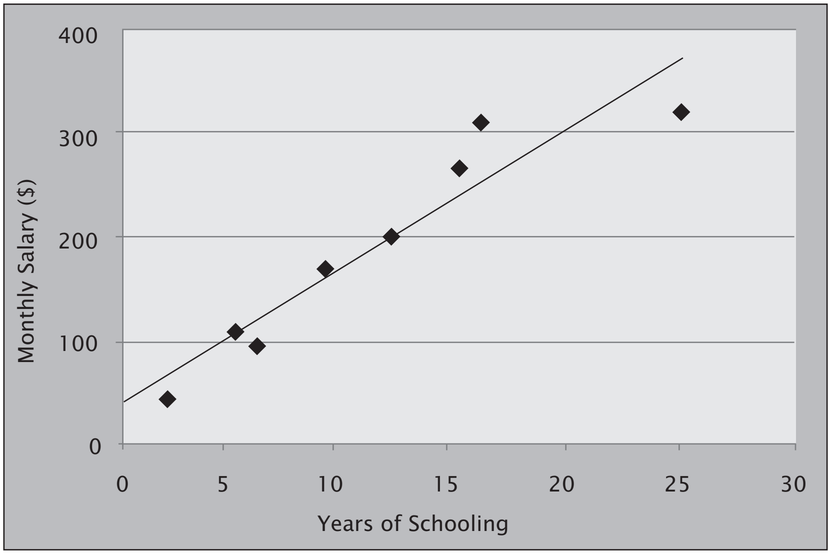


FIGURE 13.5 Relationship Between Teachers' Education and Salary (US$)

Source: Authors.

prediction about salaries (S). We can then deduce either that the effect of education upon salaries differs across teachers, or that factors other than education influence their salaries. Regression analysis ordinarily embraces the latter explanation. Thus, we now hypothesise that salary for each individual teacher are determined by education and by an aggregation of factors not included in the analysis. These omitted factors are termed as 'noise'.

In order to refine the research hypothesis further, it is natural to suppose that people in the labour force with no education nevertheless make some earnings, and that education increases earnings above this baseline. We might also suppose that education affects salaries in a 'linear fashion', that is, each additional year of schooling adds the same amount to income. This linearity assumption is common in regression studies but is by no means essential to the application of the technique and can be relaxed where the researcher has reason to suppose a priori that the relationship in question is non-linear.

Then the hypothesised relationship between education and salary may be written as:

$$S = \alpha + \beta E + \varepsilon$$

where
S = salary;
α = a constant amount (what one earns with 0 education);
β = the effect in dollars of an additional year of schooling on salary, (hypothesised to be positive) and
ε = the 'noise' term reflecting other factors that influence salary.

The variable S is termed the 'dependent' or 'endogenous' variable; 'E' is termed the 'independent', 'explanatory' or 'exogenous' variable; α is the 'constant term' and β the 'coefficient' of the variable E.

Remember what is observable and what is not. The data set contains observations for S and E. The noise component ε consists of factors that are unobservable or at least unobserved. The parameters α and β are also unobservable. The task of regression analysis is to produce an 'estimate' of these two parameters, based upon the information contained in the data set and upon some assumptions about the characteristics of ε.

We have seen above that in simple linear regression, a single independent variable is used to predict the value of a dependent variable. Let us see how we can calculate the simple regression coefficient.

Example

A researcher has collected the following data on pupil-teacher ratio (independent variable X and the mean score (out of 10 marks) in Grade 1 pupils in five primary schools as shown further.

Suppose the researcher wants to know the approximate mean score (Y value) for the PTR (X value) = 63.

PTR (X)	Mean Score (Y)
70	4.0
60	4.5
55	6.0
50	6.5
45	8.0

Regression Equation (Y)= $a + bX$

~~Where~~

$$\text{Slope } (b) = \frac{\left[N\sum XY - (\sum X)(\sum Y)\right]}{\left[N\sum X^2 - (\sum X)^2\right]}$$

$$\text{Intercept } (a) = \frac{\left[\sum Y - b(\sum X)\right]}{N}$$

x and y are the variables

b = the slope of the regression line

a = the intercept point of the regression line and the y-axis

N = number of values or elements

X = first score

Y = second score

ΣXY = sum of the product of first and second scores

ΣX = sum of first scores

ΣY = sum of second scores

$\Sigma X2$ = sum of square first scores

We can also calculate values of the slope (b) and intercept (a) by

X Value	Y Value	XY	X^2
70	4.0	70 × 4.0 = 280	4,900
60	4.5	60 × 4.5 = 270	3,600
55	6.0	55 × 6.0 = 330	3,025
50	6.5	50 × 6.5 = 325	2,500
45	8.0	45 × 8.0 = 360	2,025
$\sum X = 280$	$\sum Y = 29$	$\sum XY = 1,565$	$\sum X^2 = 16,050$

Where

$$\sum X = 280, \sum Y = 29, \sum XY = 1{,}565 \text{ and } \sum X^2 = 16{,}050$$

By substituting in the above slope formula:

$$\text{Slope } (b) = \frac{\left[5(1{,}565) - (280)(29)\right]}{\left[5(16{,}050) - (280)^2\right]}$$

$$= \frac{7{,}825 - 8{,}120}{80{,}250 - 78{,}400} = \frac{-295}{1{,}850} = -0.1595$$

Now by substituting in the above intercept formula:

$$\text{Intercept } (a) = \frac{\left[\sum Y - b(\sum X)\right]}{N}$$

$$\text{Intercept } (a) = \frac{\left[29 - (-0.1595)(280)\right]}{5} = \frac{73.66}{5} = 14.73$$

By substituting these values in regression equation formula:

$$\text{Regression Equation } (Y) = 14.73 - 0.1595X$$

Suppose if we want to know the approximate Y value for the variable $X = 63$. Then we can substitute the value in the above equation.

$$= 14.73 - 0.1595(63) = 4.7 \text{ Marks.}$$

Example

This example explains almost all the topics that we have presented before.

A random sample of eight drivers insured with a company and having similar insurance policies was selected. The following table lists their driving experiences (in years) and monthly auto insurance premiums.

Driving Experience (Years) (X)	Monthly Insurance Premium (US$) ($Y$)
5	64
2	87
12	50
9	71
15	44
6	56
25	42
16	60

We will see how we can compute and answer the following problems.

Problem 1: Does the monthly insurance premium depend on years of driving experience or does the years of driving experience depend on monthly insurance premium? Do we expect a positive or negative relationship between these two variables?

Problem 2: Compute SDs of X (SS_{xx}) and Y (SS_{yy}) and SS_{xy}

Problem 3: Find the least squares regression line by choosing appropriate dependent and independent variables based on your answer of Problem 1.

Problem 4: Interpret the meaning of the values of a and b calculated in part in Problem 3.

Problem 5: Plot the scatter diagram and the regression line.

Problem 6: Calculate r and r^2 and explain their meanings.

Problem 7: Predict the monthly insurance premium for a driver with 10 years of driving experience.

Problem 8: Compute the SD of errors.

Problem 9: Construct a 90 per cent confidence interval for B.

Problem 10: Test at the 5 per cent significance level whether B is negative.

Problem 11: Using $\alpha = 0.05$, test whether ρ is different from 0.

Solution

Answer 1: Based on empirical evidence, we expect the insurance premium to depend on driving experience. Therefore, the insurance premium is a dependent variable and driving experience is an independent variable in the regression analysis or equation. The insurance companies consider a new driver a high risk, and he or she has to pay a higher premium for insurance. On average, the insurance premium should decrease with an increase in the years of driving experience. Therefore, we expect a negative relationship between these two variables, that is, both the population correlation coefficient ρ and the population regression slope B are expected to be negative.

Answer 2: Table shows the calculation of Σx, Σy, Σxy, Σx^2 and Σy^2.

Driving Experience (Years) (X)	*Monthly Insurance Premium (US$) (Y)*	*xy*	x^2	y^2
5	64	320	25	4,096
2	87	174	4	7,569
12	50	600	144	2,500
9	71	639	81	5,041
15	44	660	225	1,936
6	56	336	36	3,136
25	42	1,050	625	1,764
16	60	960	960	3,600
$\sum X = 90$	$\sum Y = 474$	$\sum XY = 4,739$	$\sum X^2 = 1,396$	$\sum Y^2 = 29,642$

The values of $\bar{X}$ and $\bar{Y}$ are

$$\bar{X} = \sum X/\text{N} = 90/8 = 11.23 \text{ and } \bar{Y} = \sum Y/\text{N} = 474/8 = 59.25$$

The values of SDs of X (SS_{xx}) and Y (SS_{yy}) and SS_{xy} are computed as follows:

$$SS_{xy} = \sum xy - (\sum X)(\sum Y) / N = 4{,}739 - (90)(474) / 8 = -593.5000$$

$$SS_{xx} = \sum X^2 - (\sum X)^2 / N = 1{,}396 - (90)^2 / 8 = 383.5000$$

$$SS_{yy} = \sum Y^2 = (\sum Y)^2 / 8 = 29{,}642 - (474)^2 / 8 = 1{,}557.5000$$

Answer 3: To find the regression line, we can also calculate α and β as follows:

$$b = SS_{xy} / SS_{xx} = -593.5000 / 383.5000 = -1.5476$$

$$\alpha = \bar{\text{Y}} - \text{b}\bar{\text{X}} = 59.25 - (-1.5476)\ (11.25) = 76.6605$$

Thus, our estimated regression line $Y = 76.6605 - 1.5476\ X$

Answer 4: The value of $\alpha = 76.6605$ gives the value of Y for $X = 0$; that is, it gives the monthly insurance premium for a driver with no driving experience. However, as mentioned earlier, we should not attach much importance to this statement because the sample contains all drivers with only two or more years of experience. The value of b gives the change in Y due to a change of one unit in X. Thus, $b = -1.5476$ indicates that, on average, for every extra year of driving experience, the monthly insurance premium decreases by US$1.55. Note that when b is negative, Y decreases as X increases.

Answer 5: Figure 13.6 shows the scatter diagram and the regression line for the data on eight drivers. The regression line slopes downward from left to right. This result is in line and consistent with the negative relationship between driving experience and insurance premium.

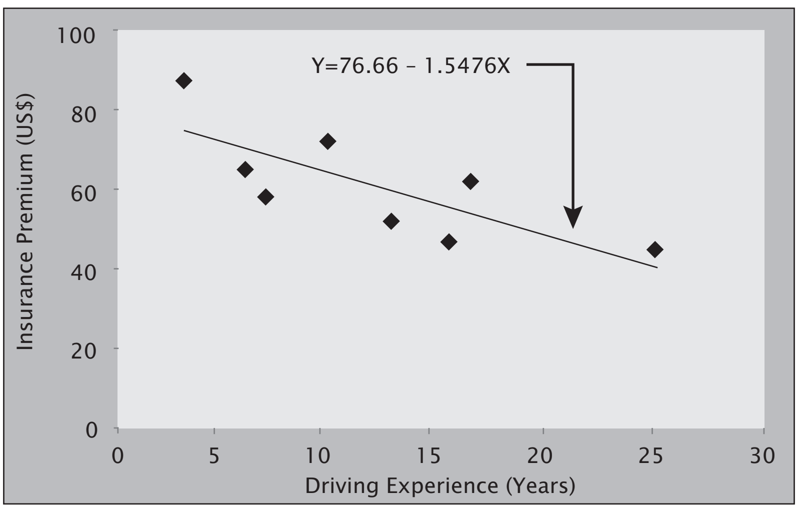


FIGURE 13.6 Scatter Diagram and the Regression Line

Source: Authors.

Answer 6: The values of r and r^2 are computed as follows:

$$r = \frac{SS_{xy}}{\sqrt{SS_{xy}SS_{yy}}} = \frac{-593.5000}{\sqrt{(383.5000)(1,557.5000)}} = -0.77$$

$$r^2 = \frac{b\ SS_{xy}}{SS_{yy}} = \frac{(-1.5476)(-593.5000)}{1,557.5000} = 0.5929$$

The value of $r = -0.77$ indicates that the driving experience and the monthly insurance premium are negatively related. The simple linear relationship is strong but not very strong. The value of $r^2 = 0.5929$ suggests that 59 per cent of the total variation in insurance premiums is explained by years of driving experience and 41 per cent is not. The low value of r^2 states that there could be several other important variables that affect and determine the amount of auto insurance premiums. For example, the premium may depend on the driving record of a driver and the model type and age of the car.

Answer 7: Using the estimated regression line, we find the predicted value of Y for $X = 10$ which is:

$$Y = 76.6605 - 1.5476X = 76.6605 - 1.5476(10) = \text{US\$} = 61.18$$

Answer 8: The SD of errors is

$$S_e = \sqrt{\frac{SS_{yy} - bSS_{xy}}{N-2}} = \sqrt{\frac{1,557.5000 - (-1.5476)(-593.5000)}{8-2}} = 10.3199$$

Answer 9: To construct a 90 per cent confidence interval for β, first we calculate the SD of β:

$$sB = \frac{s_e}{\sqrt{SS_{xx}}} = \frac{10.3199}{\sqrt{383.5000}} = 0.5270$$

For a 90 per cent confidence level, the area in each tail of the t-distribution is

$$\alpha/2 = \frac{1-0.90}{2} = 0.05$$

The df are

$$df = n - 2 = 8 - 2 = 6$$

From the t-distribution table, the t-value for 0.05 area in the right tail of the t-distribution and 6 df is 1.943. The 90 per cent confidence interval for B is

$$B \pm ts_b = -1.5476 \pm 1.943\ (0.5270)$$
$$= -1.5476 \pm 1.0240 = -2.57 \text{ to } -0.52$$

Thus, we can state with 90 per cent confidence that B lies in the interval −2.57 to −0.52, that is, on average, the monthly insurance premium of a driver decreases by an amount between US$0.52 and US$2.57 for every extra year of driving experience.

Answers 10 and 11: We perform the following five steps to test the hypothesis about B.

Step 1: State the null and alternative hypotheses.

The null and alternative hypotheses are written as follows:

$$H_0 : \text{B} = \text{(B is not negative)}$$
$$H_1 : \text{B} < 0 \text{ (B is negative)}$$

Note that the null hypothesis can also be written as $H_0 : B \geq 0$.

Step 2: Select the distribution to use.

Because '$\sigma\varepsilon$' is not known, we use the *t*-distribution to test the hypothesis.

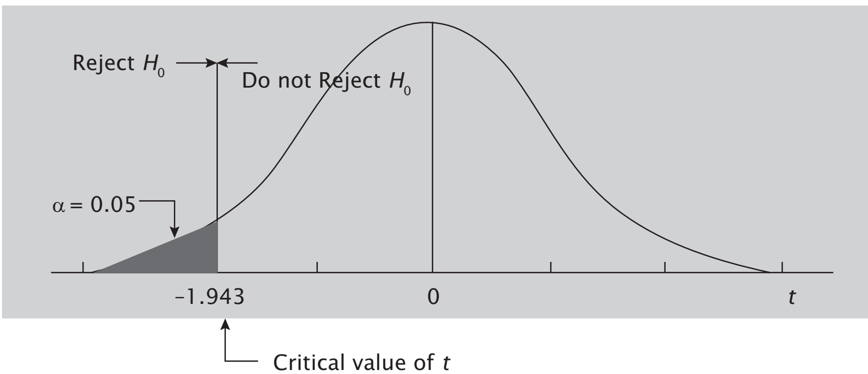
Step 3: Determine the rejection and non-rejection regions.

The significance level is 0.05. The '<' sign in the alternative hypothesis indicates that it is a left-tailed test.

Area in the left tail of the *t*-distribution = α = 0.05.

df = n − 2 = 8 − 2 = 6

From the *t*-distribution table, the critical value of *t* for 0.05 area in the left tail of the *t*-distribution and 6 *df* is = −1.943.



Step 4: Calculate the value of the test statistic.

The value of the test statistic for *B* is calculated as follows:

$$t = \frac{b - B}{s_b} = \frac{-1.5476 - \overset{\text{From } H_0}{\downarrow}0}{0.5270} = -2.937$$

Step 5: Make a decision.

The value of the test statistic $t = -2.937$ falls in the rejection region. Hence, we reject the null hypothesis and conclude that *B* is negative, that is, the monthly insurance premium decreases with an increase in years of driving experience.

Using the *p*-value to Make a Decision

From *t*-distribution table, we can find the range for the *p*-value and make a decision by comparing that *p*-value with the significance level. In our example before, df = 6 and the observed value of *t* is = −2.937.

From the t-distribution table in the row of $df = 6$, 2.937 is between 2.447 and 3.143. For these values, the corresponding areas in the right tail of the t-distribution are 0.025 and 0.01. However, our test is left-tailed and the observed value of t is negative. Thus, $t = -2.937$ lies between −2.447 and −3.143. The corresponding areas in the left tail of the t-distribution are 0.025 and 0.01. Therefore, the range of the p-value is

$$0.01 < p\text{-value} < 0.025$$

Thus, we can suggest that for any α equal to or greater than 0.025 (the upper limit of the p-value range), we will reject the null hypothesis. For our example, $\alpha = 0.05$, which is greater than the upper limit of the p-value of 0.025. Therefore, we reject the null hypothesis.

There are following five steps to test the hypothesis about the linear correlation coefficient ρ.

Step 1: State the Null and Alternative Hypotheses

The null and alternative hypotheses are:

$H_0 : \rho = 0$ (the linear correlation coefficient is 0)

$H_1 : \rho \neq 0$ (the linear correlation coefficient is different from 0)

Step 2: Select the distribution to use

Assuming that variables x and y are normally distributed, we will use the t-distribution to perform this test about the linear correlation coefficient.

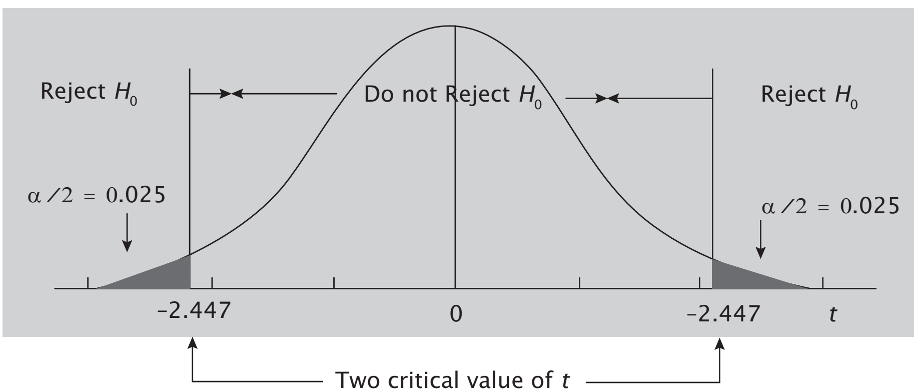
Step 3: Determine the rejection and non-rejection regions

The significance level is 5 per cent. From the alternative hypothesis, we know that the test is two-tailed. Hence,

Area in each tail of the t-distribution = 0.05 / 2 = 0.025

$Df = n - 2 = 8 - 2 = 6$

From the t-distribution table, the critical values of t are −2.447 and 2.447.



Step 4: Calculate the value of the test statistic

The value of the test statistic t for r is calculated as follows:

$$t = r\sqrt{\frac{n-2}{1-r^2}} = (-0.77)\sqrt{\frac{8-2}{1-(-0.77)^2}} = -2.956$$

Step 5: Make a decision

The value of the test statistic $t = -2.956$ falls in the rejection region. Hence, we reject the null hypothesis and conclude that the linear correlation coefficient between driving experience and insurance premium is different from 0.

We can also find the range for the p-value from the t-distribution table and make a decision by comparing that p-value with the significance level. For this example, $df = 6$ and the observed value of t is -2.956. From the t-distribution table in the row of $df = 6$, $t = 2.956$ is between 2.447 and 3.143. The corresponding areas in the right tail of the t-distribution curve are 0.025 and 0.01. Since the test is two-tailed, the range of the p-value is

$$2(0.01) < p - \text{value} < 2(0.025)$$

or

$$0.02 < p - \text{value} < 0.05$$

In our example before, $\alpha = 0.05$. This value is equal to the upper limit of the p-value. Consequently, we reject the null hypothesis.

Briefly speaking, it is clear from the above examples that the major outputs you need to be concerned about for simple linear regression are the R-squared, the intercept and the coefficients of the dependent variable. The R-squared number in this example is 59 per cent—this shows how our model predicts or forecasts the monthly premium of insurance. Next, we have an intercept of 76.66, which tells us the monthly insurance premium for a driver with no driving experience. And at last, the driving experience's correlation coefficient of –1.5476 tells us that if the driving experience increases by one year, the monthly premium of insurance will likely to decrease to US$75.11.

Now the question is: How would you use this simple model in your education research? Well if your research leads you to believe that a change in PTR by certain numbers affect the pupils' score (see example), you can plug that number into the model and generate a forecast for pupils' mean score. This can help you to assess and develop a more objective plan and budget for the upcoming year.

So far we have discussed just a simple regression. There are regression models that you can build that use several independent variables called 'multiple linear regressions' (MLRs). But MLRs are more complicated and have several issues. In the following section, we will simply introduce the concept. However, there are several statistical packages such as SPSS or Excel by which you can use this analysis within no times.

Multiple Linear Regression

The main limitation that you have with correlation and simple linear regression as you have just learned above how to do it is that it only works when you have 'two' variables. The problem is that most things are not as simple as that. We have to use models that allow the researcher to see the effect of more than two variables. For example, suppose you have been asked the following question:

'Why does a person receive the compensation for the work he or she does?'

What would you say? Your possible answers might say something like the following. It may be due to:

- How long he or she has worked for the organisation
- How much experience he or she has doing the specific kind of work
- How old is he or she

- His or her performance rating
- What part of the country he or she lives in

If you were going to use standard correlation to study why people receive the compensation they do, you would be limited to only looking at one of these things at a time. For example, you can use correlation to observe the relationship between a person's current compensation and his or her time with the company (as we did in linear regression). You can also use another correlation to study the relationship between a person's current compensation and the number of years of schooling they have completed. However, you cannot do both to find out how a person's current compensation is related to both his or her education and the length of service (years of service) he or she worked for the company. It is important to remember that Pearson's correlation is a 'bivariate' tool. It means that it is designed to find relationships between only two variables. And yet we know that life is so complicated that it takes way more than two variables to even begin to explain/predict why things are the way they are. For solving this, you need a new tool—multiple regression.

Multiple regression is a highly advanced statistical tool and it is very powerful when we are trying to develop a 'model' for predicting a wide variety of outcomes. It allows us to examine how multiple independent variables are related to a dependent variable. Once we have identified how these multiple variables relate to our dependent variable, we can take information about all of the independent variables and use it to predict accurately about why things are the way they are. This latter process is called 'multiple regression'. The goal of MLR is to establish the relationship between the explanatory and response variables.

The model for MLR, given n observations, is:

$$Y = a + b_1X_1 + b_2X_2$$

Where

Y' = A predicted value of Y (which is your dependent variable).

b_1 = The change in Y for each 1 increment change in X.

b_2 = The change in Y for each one increment change in X_2.

X = an X score (X is your independent variable) for which you are trying to predict a value of Y.

You can use the following equations to compute values of a, b_1 and b_2.

For computing a,

$$a = \overline{Y} - b_1\overline{X}_1 - b_2\overline{X}_2$$

Where

$\overline{Y}$ = The mean of Y (dependent variable).

$b_1\overline{X}_1$ = The value of b_1 multiplied by the mean of first independent varaible.

$b_2\overline{X}_2$ = The value of b_2 multiplied by the mean of second variable.

For computing b_1,

$$b_1 = \left[\frac{r_{y,x_1} - r_{y,x_2}r_{x_1,x_2}}{1-(r_{x_1,x_2})^2}\right]\left[\frac{SD_y}{SD_{x_1}}\right]$$

$$b_2 = \left[\frac{r_{y,x_2} - r_{y,x_1}r_{x_1,x_2}}{1-(r_{x_1,x_2})^2}\right]\left[\frac{SD_y}{SD_{x_2}}\right]$$

Where

$r_{y,\,x_1}$ = Correlation between variable X_1 and variable Y.
$r_{y,\,x_2}$ = Correlation between variable X_2 and variable Y.
$r_{x_1,\,x_2}$ = Correlation between variable X_1 and variable X_2.
$(r_{X_1,\,X_2})^2$ = The coefficient of determination.
SD_y = SD of Y variable.
SD_{x_1} = SD for the first variable X_1.
SD_{x_2} = SD for the second variable X_2.

In MLR, the researcher takes a group of random variables and tries to establish a mathematical relationship between them. The MLR model generates a relationship in the form of a straight line (linear) that best approximates all the individual data points.

Let us consider a simple case. Suppose you have collected data on three psychological variables, four academic variables (standardised test scores) and the type of educational programme the student is attending for 600 high school students.

You are interested in assessing:

'How the set of psychological variables is related to the academic variables and the type of programme the student is attending'.

- Psychological variables
  1. Locus of control
  2. Self-concept
  3. Motivation
- Academic variables
  1. Standardised test scores in reading
  2. Standardised test score in writing
  3. Standardised score in science
- Categorical variable (programme) giving the type of programme the student is in
  1. General education
  2. Teacher training
  3. Vocational education

This is a complex problem. You need not to worry. There are several statistical packages and software such as SPSS or Excel which are highly useful in computing and solving such complex problems. You have to simply create your data sheets (files) for the systematic entry of data in these programmes.

Example

The Chinese Ministry of Higher Education collected data on enrolment in Nanjing University, unemployment rate (UNEM) in China, number of Nanjing University graduates (GRAD) and per capita income (PCI). On the basis of these data, the Ministry seeks to predict the total enrolment in Nanjing University via the UNEM, GRAD and PCI.

The computation found the following:

	Coefficients		
Intercept	UNEM	GRAD	PCI
−9,153.2545	450.1245	0.4065	4.2749

From this output, we can determine that the intercept is −9,153.2545, the coefficient for UNEM rate is 450.1245, the coefficient for GRAD is 0.4065 and the coefficient for PCI is 4.2749. Therefore, the complete regression equation is

$$\text{ENNR} = -9{,}153.2545 + (450.1245 \times \text{UNEM}) + (0.4065 \times \text{GRAD}) + (4.2745 \times \text{PCI})$$

This equation tells us that the predicted total enrolment for the university will increase by 450.1245 students for 1 per cent increase in the UNEM, 0.4065 students for every one graduate and 4.2745 per cent for every one dollar of PCI.

While doing the above computation, you should always keep the following in mind:

- You want to calculate a single summary number that tells you how strong the relationship is between 'all' the independent variables and the dependent variable. What you want is similar to the correlation coefficient *r*. You should, however, remember that *r* is only used with two variables. The statistic that we are explaining to you here is called *R* (which is a capitol *r*) and different than *r* statistic. A capitol *R* is expressed just like any regular correlation coefficient except it indicates the strength of the combined relationships between all the independent variables in the dependent variable.
- When we talked earlier, we talked about regular old correlation where we explained the 'coefficient of determination' which is symbolised as r^2. We can treat *R* in the same manner to get an understanding of how much variation in the dependent variable is explained by the independent variables. To get this statistic, all you need to do is square your *R*-value which gives you R^2.
- Finally, you would be able to actually make predictions about a dependent variable taking into account all of the information provided by all the independent variables.

In short, you can use a computer-based MLR model to do much more than just calculate expected values. In this case, the summary (OBJECT) function is a highly useful tool. With this function, you can generate a wealth of important information about a linear model. It provides *t*-test, *F*-test, *R*-squared, residual and significance values. You can use all of these data to answer important questions related to our models.

Once you have constructed a multiple regression equation, you can check how good it is (in terms of predictive ability) by examining the coefficient of determination (R^2). R^2 always lies between 0 and 1. A value of R^2 close to 1 shows the better is the model and its prediction.

Analysis of Variance

ANOVA is relatively a sophisticated hypothesis-testing technique widely used in research studies. It is used to evaluate mean differences between two or more populations. Like all other inferential statistical methods, we use in ANOVA sample data as the basis for drawing overall conclusions about populations. The key merit of ANOVA is that we can use this method to compare two or more populations. In other words, ANOVA provides researchers with much greater flexibility in designing experiments and interpreting results.

Although ANOVA can be used in a wide variety of research situations, here we introduce ANOVA in its simplest form. In particular, the analysis deals with only single-factor designs, that is, the analysis:

- Examines studies that have only one independent variable
- Considers only independent-measures design, that is, studies that use separate sample for each treatment condition

Looking from this perspective, the null hypothesis would be stated as follows:

$$H_0 : \mu_1 = \mu_2 = \mu_3$$

Similarly, the research hypothesis would be stated as follows:

$$H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \text{ (all three means are different)}$$

Let us consider the following example that will provide you the procedures to calculate and understand better ANOVA.

Example

Suppose we have a sample of employees under normal times and we want to measure the level of stress. We do a sample of five employees under the normal working condition (green sample) shown in Figure 13.7(A). Then we do another sample (pink sample) of the same five employees after the announcement of lay-offs. Finally, we take third sample (blue sample) of these five employees with level of stress during lay-offs. With this information, we want to measure the impact of the announced lay-offs on the employees' level of stress. ANOVA, we have seen above, allows us to compare different samples at different points in times.

In this case, we have to make several calculations all around the variance also known as the SS. In order to find out the value of SS, it has to be calculated for all the three samples. This is called sum of squares within groups (SSWG).

The next calculation we make is to look at the variance between groups. This is known as sum of squares between groups (SSBG). We also take these three samples to make one large sample or one large study. This is called total sum of squares (TSS). It will then be followed by the computation of an *F*-ratio to see if it falls within the rejection region or not (blue area of the curve in Figure 13.7[B]). If the value of *F*-ratio falls in the grey region, then we reject the hypothesis and if it falls within the white area of the curve, then we fail to reject the null hypothesis.

Finally, we will make the following three calculations and combine with the *df* to know the value of *F*-ratio. This computed value has to be compared with critical value in *F*-table either for accepting or rejecting the null hypothesis.

TSS = SSBG + SSWG

(TSS)		(SSBG)		(SSWG)
257.5	=	203.3	+	54

Let us see how we can calculate the SSWG. For this calculation, we take the green sample. The same calculation should be carried out for the red and blue samples.

Observation (*X*)	$(X - \text{Mean})^2$
2	$(2-4)^2 = 4$
3	$(3-4)^2 = 1$
7	$(7-4)^2 = 9$
2	$(2-4)^2 = 4$
6	$(6-4)^2 = 4$
$\text{Mean} = \frac{\sum X}{N} = \frac{20}{5} = 4$	$\sum (X - \text{Mean})^2 = 22$

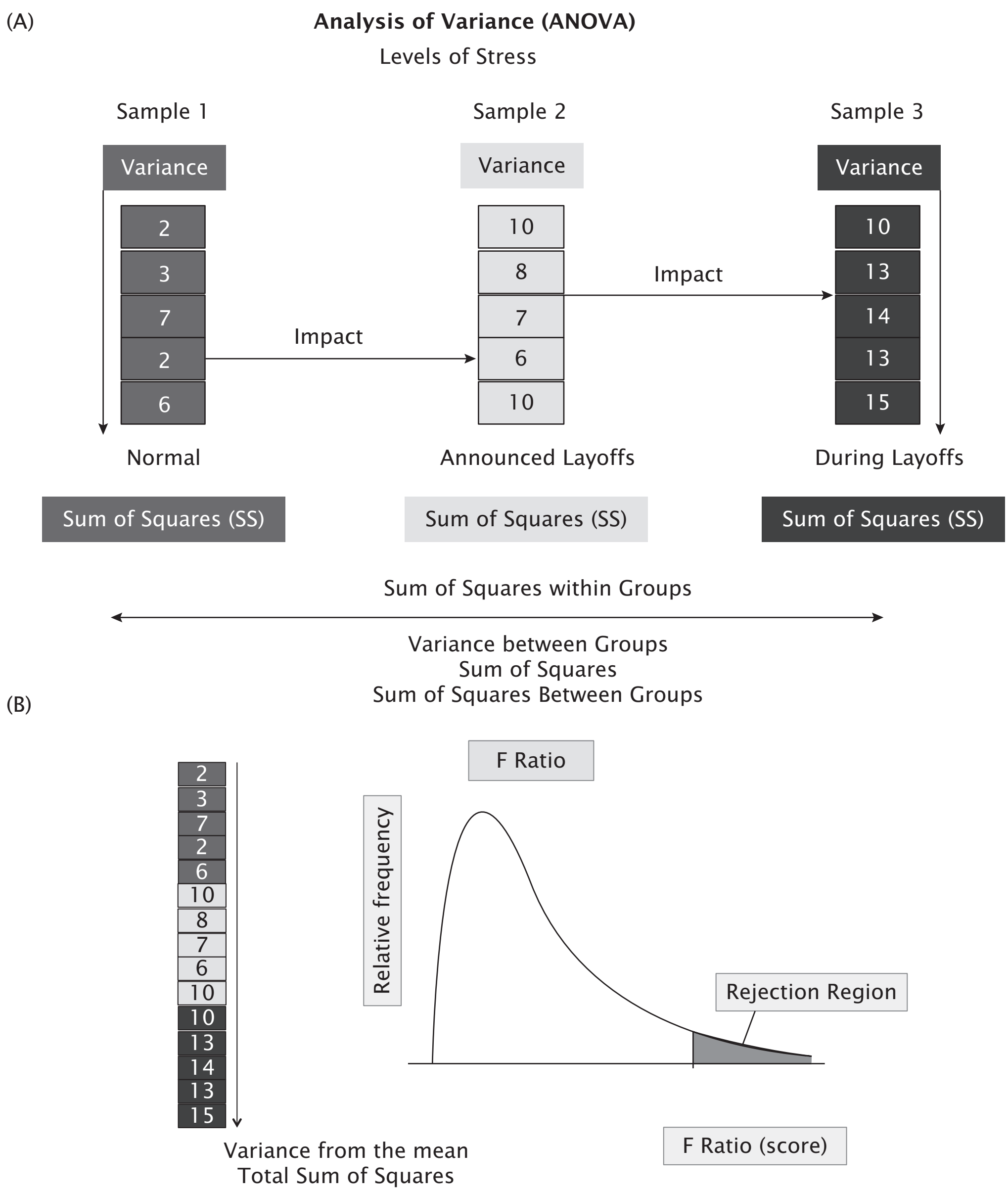


FIGURE 13.7 Calculation of Variance

Source: Authors.

We repeat this for red and blue samples. The mean values are 8 and 13, and the SSWG for the two samples are 18 and 14, respectively. The SSWG of three samples is:

$$\text{SSWG} = 22 + 18 + 14 = 54$$

Now, we have to calculate TSS. We take all the three samples together and calculate the variance from the mean and calculate its TSS. The calculations are shown as follows:

Observation (X)	$(X - \text{Mean})^2$
2	$(2 - 8.3)^2 = 39.7$
3	$(3 - 8.3)^2 = 28.1$
7	$(7 - 8.3)^2 = 1.7$
2	$(2 - 8.3)^2 = 39.7$
6	$(6 - 8.3)^2 = 5.3$
10	$(10 - 8.3)^2 = 2.9$
8	$(8 - 8.3)^2 = 0.1$
7	$(7 - 8.3)^2 = 1.7$
5	$(5 - 8.3)^2 = 10.9$
10	$(10 - 8.3)^2 = 2.9$
10	$(10 - 8.3)^2 = 2.9$
13	$(13 - 8.3)^2 = 22.1$
14	$(14 - 8.3)^2 = 32.5$
13	$(13 - 8.3)^2 = 22.1$
15	$(15 - 8.3)^2 = 44.9$
$\text{Mean} = \frac{\sum X}{N} = \frac{125}{15} = 8.3$	TSS = 257.5

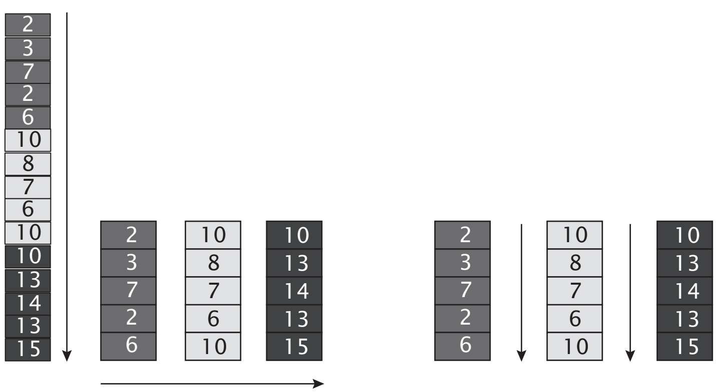
TSS = 257.5

Once we know the values of SSWG and TSS, we can calculate the value of SSWG as follows:

$$\text{SSWG} = \text{TSS} - \text{SSWG or } 257.5 - 54.0 = 203.5$$

Now we have all the calculated means

$$(\text{mean}_1 - \text{mean grp})^2 + (\text{mean}_2 - \text{mean grp})^2 + (\text{mean}_3 - \text{mean grp})^2 \times 5$$
$$(4 - 8.3)^2 + (8 - 8.3)^2 + (13 - 8.3)^2 \times (5)$$
$$(-4.3)^2 + (-0.3)^2 + (4.7)^2 \times (5)$$
$$(18.8 + 0.1 + 21.8)(5)$$
$$(40.7)(5) = 203$$



TSS = SSBG + SSWG

(TSS)		(SSBG)		(SSWG)
257.5	=	203.3	+	54

Now we have all the three calculations, we will do the final calculation.

$$\text{SSBG} / df = 203.3 / 2 = 101.667$$

Where

df = Number of groups −1 = 3 − 1 = 2

Similarly,

$$\text{SSWG} / df = 54 / 12 = 4.5$$

Here df = total number of observation −3 = 15 − 3 = 12

Thus,

$$F = 101.667 / 4 = 22.59$$

Now we have to look up this number in the *F*-table for our numerator which in this case is 2 and denominator which is 12, that is $F(2, 12) = 22.59$ at 5 per cent ($p < 0.05$) level of significance.

F-Distribution F (2, 12) = 22.59, $p < 0.05$											
		df numerator									
df denominator	$df_2 \backslash df_1$	1	2	3	4	5	6	7	8	9	10
	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6	5.96
	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.1	4.06
	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
	8	5.32	4.46	4.07	3.84	3.69	3.58	3.5	3.44	3.39	3.35
	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
	10	4.96	4.1	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
	11	4.84	3.98	3.59	3.36	3.2	3.09	3.01	2.95	2.9	2.85
	12	4.75	3.89	3.49	3.26	3.11	3	2.91	2.85	2.8	2.75
	13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67

Thus, our critical value from *F*-table is = 3.89, whereas our calculated *F* is = 22.59. These are shown in Figure 13.8.

From Figure 13.8, you can see that *F*-ratios of 3.89 or above are unusual occurrences. The area to the right of 3.89 represents the probability of an *F* that is large or larger and is equal to 22.59 and therefore the null hypothesis can be rejected. The conclusion that at least one of the population means is different from at least one of the others is justified.

In short, ANOVA is used to compare several means. It is important to keep in mind that a *t*-test is used to test differences between two means, that is, the mean of the experiment group versus control group. An ANOVA test, on the other hand, is indicated when there are three or more means or populations to be examined.

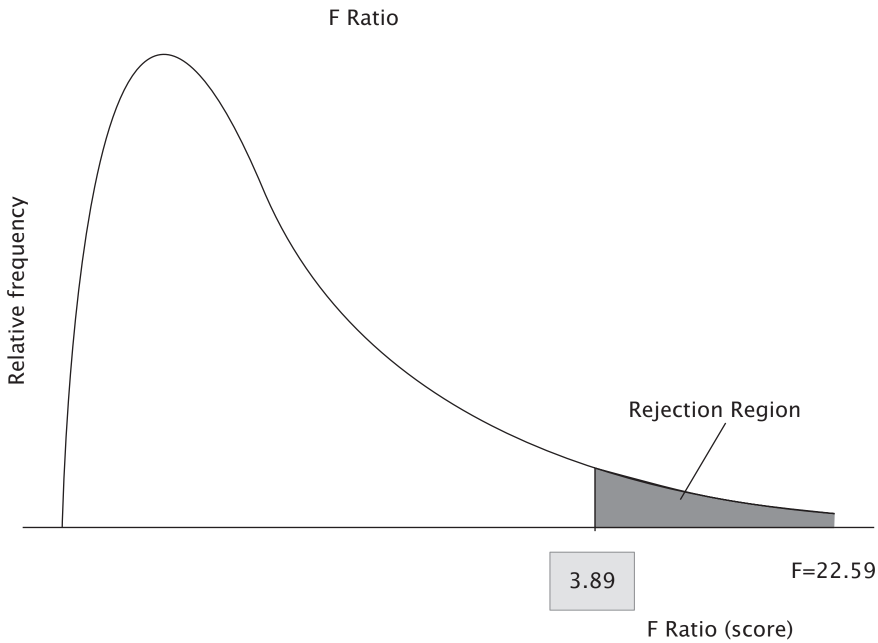


FIGURE 13.8 Hypothesis Testing

Source: Authors.

Analysis of Time Series

Time series modelling is a dynamic research area. The main aim of time series analysis is to carefully collect and rigorously study the past observations of a time series to develop an appropriate model which describes the inherent structure of the series. This model is then applied to generate future values for the series, that is, to make forecasts. Time series forecasting thus can be termed as the act of predicting the future by understanding the past (Raicharoen, Lursinsap and Sanguanbhoki 2003). Due to the indispensable importance of time series forecasting in numerous practical fields such as business, economics, finance, education and so on, proper care should be taken to fit an adequate model to the underlying time series.

In the domain of educational research, time series designs are an alternative to pretest and post-test methods that are able to identify and measure the impacts of multiple educational interventions, even for small student populations. For instance, we use a questionnaire comprising standard multiple-choice conceptual questions to collect data from students at regular intervals. The questions are amended by asking students to distribute 100 confidence points among the options in order to indicate the perceived likelihood of each answer option being the correct one. Tracking and monitoring the class-averaged ratings for each option produces a set of time series. Autoregressive integrated moving average (ARIMA, discussed further) analysis is then used to test and measure changes in each series. In particular, it is possible to discern which educational interventions produce significant changes in class performance.

The purpose of this section of the chapter is to provide an introduction to time series analysis and forecasting. It first reviews techniques used to identify patterns in time series data (such as smoothing and curve fitting techniques and autocorrelations) and then introduces a general class of models that can be used to represent time series data and generate predictions (autoregressive and moving average models). Finally, it reviews some simple but commonly used modelling and forecasting techniques based on linear regression.

Meaning of Time Series Analysis

Suppose we have been asked to provide annual forecasts (projections) of primary level enrolment in a Country X over the coming five-year period. Teachers' requirements, pedagogic material, students' space (number of classrooms) and so on will all be affected by the annual forecasts we provide. Consequently, poor forecasts may result in poor planning and increased costs for the education system. How should we go about providing the annual enrolment forecasts? Good judgement, intuition and an awareness of the state of the education system may give us a rough idea or 'feeling' of what is likely to happen in the future, but converting that feeling into a number that can be used for forecasting pupils enrolment for a given year is a difficult task.

An arrangement of data by successive time period is called time series. For example, the total monthly sales receipts in a departmental store, the annual yield of a crop in a country for a number of years, hourly temperature recorded at a locality for a period of years, the weekly prices of wheat in a Country X, the monthly consumption of electricity in a certain town, the monthly total of passengers carried by rail, the quarterly sales of a certain fertiliser, the annual rainfall at Town Y for a number of years, the enrolment of students in a college or university over a number of years and so on.

> A time series is a sequence of observations on a variable measured at successive points in time or over successive periods of time. The measurements may be taken every hour, day, week, month or year, or at any other regular interval.

For understanding how the time series has behaved in the past, it is important to know and study the pattern of the data. If such behaviour can be expected to continue in the future, you have to use the past pattern to guide us in selecting an appropriate forecasting method. A forecast is simply a prediction of what will happen in the future. As researchers, we must learn to accept that regardless of the technique used, we will not be able to develop perfect forecasts.

When the data are arranged on the basis of time of their occurrences, often they show fluctuations from time to time—from day to day, from week to week, from month to month and from year to year. These fluctuations are caused due to the combined effect of number of factors, commonly known as the 'component of a time series'. Therefore, it is necessary to isolate and measure the separate effects of these components in a given time series. The analysis of time series enables us to isolate and measure the separate effects of these components in a given time series.

The analysis of time series is extremely useful to an educational planner and researcher in planning future operations and in assessing the effect of an intervention in the system. It helps in evaluating current achievements which can be compared with the expected performance and analyse the cause of variations.

Components of a Time Series

A typical time series has four types of movements usually called components of a time series. They are:

- Secular trend or long-term movement (T)
- Seasonal movements or variations (S)
- Cyclical movements variations or fluctuations (C)
- Irregular, accidental or random movements (I)

Secular trend (T): The general tendency of a time series data to increase or decrease over a long period (20 to 30 years or more) is called 'secular trend' or simply 'trend'. The concept of trend does not refer to short-term fluctuations or variations but refers to steady movements showing continuous growth, for example, relating to population, production, sales, prices, wages, epidemics and so on. The main objective of studying trend is to have a general idea about pattern of behaviour of the phenomenon under study. This helps in forecasting and planning future operations.

Seasonal variations (S): They are short-term periodic variations which occur regularly every year. The major factors these variations are climate and weather conditions, social customs, traditions and religious festivals. For example, the demand of cold drinks rises in summer and falls in winter. The sales get increased before Christmas or other religious festivals.

Cyclical variations (C): They refer to long-term oscillatory movements in a time series with a period of oscillation greater than one year. Statistical data in a number of cases show periodic or cyclical fluctuations. There are peaks from prosperity through troughs from recession, depression and recovery, and back to prosperity, which are well-known four phases of a business cycle (depression, revival or recovery, prosperity or boom, contraction or recession) and very important example of cyclical movements. These changes happen repeatedly at regular intervals ranging from 7 to 10 years. The study of the cyclical fluctuations helps business executives in the formulation of policies aimed at establishing the level of business activity.

Irregular or random movements (I): These movements are irregular and unsystematic in nature and happen as a result of abnormal events such as floods, earthquakes, wars and strikes, and so on. For example, prices rise during civil strife or war; the production of industries goes down due to labour strikes and so on. These variations are complex in character and cannot be isolated easily. They can occur at any time, but without any pattern or regularity.

Methods of Measuring Trends

The analysis of time series comprises the description, measurement and isolation of the various components present in the series. This analysis helps the economists, businessmen, researchers, planners and so on.

There are two mathematical models—'additive' and 'multiplicative'—that are commonly used for the decomposition of a time series into its components.

In the additive model, the value of the time series (Y) follows the additive law, that is:

$$Y = T + C + S + I$$

This model assumes that different components are absolute quantities expressed in original units and can take positive and negative values. The four components are independent of one another.

Whereas multiplicative model is the product of four components—trend (T), Cyclical (C), Seasonal (S) and Irregular (I) movements.

$$Y = T \times C \times S \times I$$

In this model, only trend is expressed in terms of original values, while the seasonal and cyclical components are expressed as relatives or percentages.

The following four methods are commonly used for measuring trends.

Free Hand Curve or Graphic Method

This method makes use of graphs where the data points are plotted on X-axis of a graph showing the time units (year, months and so on) and the value of the time series variable along the Y-axis. A trend line or smooth curve is drawn through the graph in such a way that it shows the general tendency of the values. The trend values for different years (or months) are read from the trend line or curve. Table 13.3 presents data on primary enrolment in Country X during the period 2005–2009.

TABLE 13.3 Primary Level Enrolment in Country X (In Million)

Year	*Enrolment*	*Year*	*Enrolment*
2004	30	2010	40
2005	33	2011	47
2006	28	2012	46
2007	38	2013	55
2008	42	2014	52
2009	35	2015	57

Source: Authors.

These data are plotted in Figure 13.9.

As far as merits of this method are concerned, it is the simplest method of estimating trend. It is very flexible in the sense that it can be used to describe all types of trend—linear or non-linear. However, its demerit is its highly subjective nature and consequently different people are likely to draw different trend curves for the same set of data.

Semi-averages Method

The free hand curve method, as we have presented before, depends largely on personal judgement and thus yields subjective results. Another simple method for measuring secular trend is the method of semi-averages. In this method, the data are divided into two equal parts (in case the number of values is odd, either the middle value is ignored [left out] or the series is divided unequally). The averages for each part are calculated and placed against the centre of each part. The averages are plotted and joined by a line. The line is extended to cover the whole data. Trend values corresponding to different time periods can be read from this trend line. See example featuring the Table 13.4.

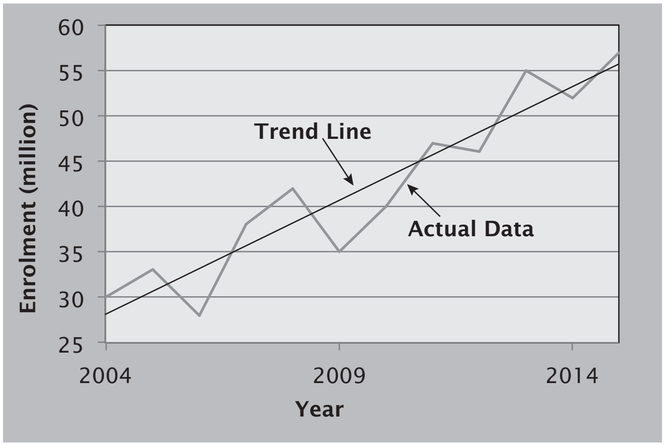


FIGURE 13.9 Fitting a Trend Line by the Free Hand Curve Method

Source: Authors.

TABLE 13.4 Primary School Teachers' Attrition

Year	*Attrition ('000)*	*Semi-total*	*Semi-average*
2006	38		
2007	45		
2008	41	225	225 / 5 = 45
2009	53		
2010	48		
2011	60		
2012	56		
2013	64	320	320 / 5 = 64
2014	72		
2015	68		

Source: Authors.

These data has been plotted in Figure 13.10
The merits of the semi-averages method are:

- It is a simple method.
- The trend line can be extended both ways to obtain future or past estimates.
- The method is objective in the sense that any two persons will get the same trend for the given time series.

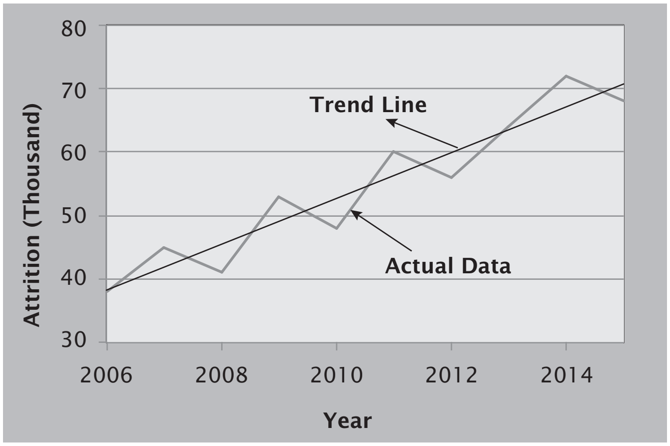


FIGURE 13.10 Fitting a Trend Line by the Method of Semi-averages

Source: Authors.

Demerits of the method are:

- This method assumes the presence of linear trend which may not exist.
- Since the mean is generally affected by extreme values, therefore, if extreme values are present in either half or both halves of the series, then the trend line does not give the true picture of the growth factor.

Method of Moving Averages

We have seen that the free hand curve method is subjective because it relies too much on individual judgement. Likewise, the method of semi-average is can be used and appropriate only when the trend is linear. Thus, there is another simple method which you can use to eliminate seasonal, cyclical and irregular movements. This method is referred to as moving averages. In this method, you can find the simple average successively taking a specific number of values at a time.

Suppose you want to find out 3-year moving average. First, you will calculate the average of the first three values, then drop the first value and include the fourth value. The process will be continued till all the values in the series are included. The averages so obtained are placed in the middle of the group for which the average is calculated. When you find the moving average taking an even number of values, the middle of the group will lie between two years. In order to make the average coincide with a particular year, we centre the averages by calculating further a 2-year moving average of the even order moving averages. The averages obtained are called moving averages (centre). The computation is shown in Table 13.5.

Merits of moving averages:

- The method is quite simple to apply.
- It is flexible in the sense that even if we add few more observations, the entire calculations do not change.

TABLE 13.5 Computation of 3-year Moving Average

Year (1)	*Value (2)*	*3-year moving total (3)*	*3-year moving average (4 = (3) ÷ 3*
2009	2	–	–
2010	4	2 + 4 + 6 = 12	(12 ÷ 3) = 4
2011	6	4 + 6 + 8 = 18	(18 ÷ 3) = 6
2012	8	6 + 8 + 13 = 27	(27 ÷ 3) = 9
2013	13	8 + 13 + 12 = 33	(33 ÷ 3) = 11
2014	12	13 + 12+ 14 = 39	(39 ÷ 3) = 13
2015	14	–	–

Source: Authors.

- The method is objective in the sense that it is independent of personal bias, that is, anybody working on a problem with this method will get the same results.
- In addition to the measurement of trend, the method of moving averages is also used for measurement of seasonal, cyclical and irregular variations.
- If the period of the moving average happens to coincide with the period of cyclical movement in the given series, the cyclic fluctuations are automatically eliminated.

Demerits of moving averages:

- Trend values cannot be calculated for all the years.
- The method cannot be used in forecasting or predicting future values.
- The method fails to reveal the trends in a non-linear trend.

Least Square Method

The principle of least square refers to 'the SS of the deviations of the observed values from the corresponding estimated values should be least' in this method a straight line $Y = a + bX_1$, second-degree parabola $Y = a + bX + cX^2$ and a third-degree parabola $Y = a + bX + cX^2 + dX^3$ are fitted to the observed time series by the method of least squares.

The straight line obtained by this method is the line of 'best fit' that approximates the given time series data.

Linear Trend Line

According to the method of least squares, a line of 'best fit' for a given time series data is the line such that the sum of the squares of the individual time series observations from the corresponding estimated trend values by the line is least. The linear trend is calculated from the following:

$$Y_c = a + bX_1$$

Where Y_c represents the estimated value of the trend, X represents the deviation to some convenient time period, 'a' and 'b' are constants. In this method, we have to determine the values of both constant so as to minimise:

$$S = \sum(Y - Y_c)^2$$

This can be done by solving the following two normal equations.

$$\sum Y = na + b \sum X$$

and

$$\sum XY = a \sum X + b \sum X^2$$

Where

$$a = \sum Y / n$$

$$b = \sum XY / \sum X^2 \text{ and}$$

n represents the number of years for which data are given.
For example, suppose that we have the following data on total expenditure on primary level of education in Country Z (Table 13.6) and we want to estimate the likely expenditure for the year 2020 and calculate the trend value.

TABLE 13.6 Public Expenditure on Primary Education in Country Z 2009–2015 (Million US$)

Year (t)	*Expenditure (million US$) Y*	*X = t – 2012*	*XY*	*X₂*	*Trend Values (Y_c)*
2009	60	–3	–180	9	61.42
2010	72	–2	–144	4	66.28
2011	75	–1	–75	1	71.14
2012	65	0	0	0	76.00
2013	80	1	80	1	80.86
2014	85	2	170	4	85.72
2015	95	3	285	9	90.58
$n = 7$	$\sum Y = 532$	$\sum X = 0$	$\sum XY = 136$	$\sum X^2 = 28$	

Source: Authors.

Since $\sum X = 0$, the values of a and b S

$$a = \sum Y / n = 532/7 = 76 \text{ and } b = \sum XY / \sum X^2 = 136/28 = 4.86$$

Hence, the equation of the straight line is: $Y_c = 76 + 4.86X$
For the year 2020, $X = 8$. Thus,

$$Y_{2020} = 76 + 4.86(8) = 114.88$$

Similarly, you can compute the estimated values of second- and third-degree trend lines by using the following equations:

Second-degree Parabola/Second-degree Trend Line

$$Y = a + bX + cX^2$$

$$\Sigma Y = na + b\Sigma X + c\Sigma X^2$$

$$\Sigma XY = a\Sigma X + b\Sigma X^2 + c\Sigma X^3$$

$$\Sigma X^2\ Y = a\Sigma X^2 + b\Sigma X^3 + c\Sigma X^4$$

By substituting the values, we can calculate *a*, *b* and *c*
Merits of method of least squares

- The method is free from subjectivity as it is based on a mathematical basis.
- It provides the line of 'best fit' because it is this line from where the sum of positive and negative deviations is zero and sum of the squares of deviations is least.
- It enables us to compute the trend values for the entire time period.
- The trained equation obtained by this method can be used to estimate or predict the values of the variable for any given time period *t* in future and the forecasted values are quite reliable.

Demerits of method of least squares

- The method requires several calculations and is quite tedious and time-consuming as compared to other methods.
- The addition of even a single new observation requires re-calculation.
- Predictions or forecasts are based on long-term variations and completely ignore the cyclical, seasonal and irregular fluctuations.

In conclusion, time series analysis serves two main goals: (a) identifying the nature of the phenomenon represented by the sequence of observations, and (b) forecasting (predicting future values of the time series variable). Both of these goals require that the pattern of observed time series data is identified and more or less formally described. Once the pattern is established, we can interpret and integrate it with other data (i.e., use it in our theory of the investigated phenomenon, e.g., seasonal trends in a variable). Regardless of the depth of our understanding and the validity of our interpretation (theory) of the phenomenon, we can extrapolate the identified pattern to predict future events.

Summary

The process of data analysis deals with scanning, examining and interpreting data available in tabulated form. Data analysis basically serves two purposes: (a) a clear understanding of the nature of data and (b) reach a conclusion. Data analysis actually provides answers to the research questions or research problems that the researcher has formulated. Without data analysis, you cannot draw or arrive at any conclusion. After analysing data, you get an organised and well-examined form of data that can help you know whether your hypothesis got accepted or rejected.

While analysing your data, an important thing that you have to keep in mind is to look at your data and decide on the method of data analysis. There are some basic tips you need to follow to analyse data in research papers and dissertations.

The first consideration is the organisation of your data before scanning, examining or interpreting it. Data organisation is necessary because you cannot analyse haphazard data. One easier way is to organise your data in tables or groups. However, this is easier to do if your data are quantitative. But qualitative data are difficult to tabulate. You can first arrange your data in groups or categories. And then under each category you can tabulate the data. For qualitative data, you have to follow different methods of data organisation. Well-organised data lend themselves easily to analysis.

In the next step, describe the data that are shown in both tabulated and graphical forms. This description will help you draw main conclusions. Study and explore the graphs and tables, and find out how you can write down the interpretation of your research study. You can correlate the variables and you can also explain the results. Try to make the interpretation simple, specific and to the point. You should avoid extremely lengthy explanations, as they are unnecessary in most cases; on the other hand, a specific interpretation of the data is easy to understand.

In the last stage, you should test the hypothesis and either reject or accept in the light of your interpretations. You have to ensure that your hypothesis proved right or it proved wrong. You can use any one of the statistical methods for confirmation of the hypothesis. Generally, you can use ANOVA, *t*-test, *z*-test or chi-square to test the hypothesis. There are also software that can help you in this regard. You can also seek help of a statistician to apply statistical methods to your research. Statistical application is important because it makes your research valid and generalisable.

Chapter 14 describes the several tabular and graphical presentations of your analysed data. It also explains some simple ways to organise the writing of your final dissertation or thesis.

Self-test Exercise

Exercise 13.1: For the following set of scores

X	*Y*
4	9
9	2
6	7
7	4
7	7
9	7

13.1.1. Compute the Pearson correlation.

13.1.2. Add 2 points to each *X* value and compute the Pearson correlation for the modified scores.

13.1.3. Sketch a scatter plot showing the six data points.

Exercise 13.2:

13.2.1. A researcher wishes to determine whether cooperative learning, computer assisted learning or self-paced learning provides the most academic achievement for students of educational research.

Determine which statistical test(s) should or could be used for each example. Possible tests include: correlation, *t*-test, ANOVA or chi-square.

13.2.2. The basic divisions of statistics are:
A. Population and sample
B. Sampling and scaling
C. Inferential and descriptive
D. Mean and median

13.2.3. Unlike *t*-tests, an ANOVA uses both differences between group means and differences within groups to determine whether or not the differences are significant.
A. True
B. False

Exercise 13.3: Using the data given in the following table

Disposable Income and Sale of Vehicle

Year	*New Motor Vehicle Sales (Units)*	*Real Personal Disposable Income per person (1992 US$)*
1981	1,190,882	16,368
1982	920,902	16,107
1983	1,081,088	15,765
1984	1,283,502	16,252
1985	1,530,410	16,710
1986	1,515,920	16,747
1987	1,533,637	16,856
1988	1,565,501	17,443
1989	1,483,875	17,875
1990	1,317,869	17,756
1991	1,287,790	17,164
1992	1,227,419	17,096
1993	1,192,934	17,016
1994	1,260,056	17,000
1995	1,166,535	17,047
1996	1,204,557	16,861
1997	1,424,380	16,994
1998	1,428,932	17,334
1999	1,542,041	17,735
2000	1,587,561	18,068
2001	1,597,949	18,185

13.3.1. Plot a scatter diagram.

13.3.2. Determine the regression equation.

13.3.3. Plot the regression line.

13.3.4. Compute the predicted vehicle sales for disposable income of $16,500 and of $17,900.

13.3.5. Compute the coefficient of determination and the coefficient of correlation.

Exercise 13.4: In Delhi, the monthly food expenses of a typical family were the following:

		2010		*2015*	
Product Bread	*Measurement Kg*	*Quantity*	*Price (₹)*	*Quantity*	*Price (₹)*
Bread	Kg	18	94	25	110
Milk	Litre	32	60	44	80
Apple	Kg	30	74	38	120

13.4.1: Calculate how the following changed during these years.
A. The expenses by product type separately and also in total.
B. The volume of consumption by product type separately and also in total.
C. The selling price by product type separately and also in total.

13.4.2: Highlight the relations between the various indices.

13.4.3: Calculate how much more the family spent on food in 2015 due to the price and/or the volume change.

References

Croxton, E.F. and D.J. Cowden. 1946. *Applied General Statistics*. Prentice Hall.

Hazarika, N. 2013, June. *MajesticSEO Beginners Guide to Correlation*. Available at https://blog.majestic.com/case-studies/majesticseo-beginners-guide-to-correlation-part-3/ (accessed on 12 February 2018).

Raicharoen, T., C. Lursinsap and P. Sanguanbhoki. 2003, 25–28 May. *Application of Critical Support Vector Machine to Time Series Prediction*. Circuits and Systems. ISCAS '03, Proceedings of the 2003 International Symposium on Volume 5.

Sykes, A.O. 2000. *An Introduction to Regression Analysis*. In *Chicago Lectures on Law and Economics* edited by Eric A. Posner. New York, NY: Foundation Press.

ViSta: The Visual Statistics System. *Correlation: The Relationship of Two Variables*. Available at http://forrest.psych.unc.edu/research/vista-frames/help/lecturenotes/lecture11/overview.html (accessed on 17 November 2017).

14
CHAPTER

Data Presentation and Report Writing

Introduction

We have learned in earlier chapters how data are collected and organised. As data are generally voluminous, they need to be condensed and put in a presentable form. In presenting the data and statistical computations, it is often desirable to use appropriate forms of graphic presentation. In addition to tabular forms, graphic presentation entails use of graphics, charts and other pictorial devices such as diagrams. These forms and devices condensed large masses of statistical data to a form that can be quickly understood at a glance. It is often difficult for the mind to grasp and retain the meaning of figures in a tabular form. Properly constructed graphs and charts ease the mind of burdensome and complex details by depicting facts concisely, logically and simply. They, by emphasising new and significant relationships, are also significantly meaningful in exploring new facts and in developing hypotheses.

The device of graphic presentation is especially useful when our prospective readers are non-technical people or general public. It is useful to even technical people for magnifying certain points about data for important points can be more effectively seized when presented in pictures than in tables. However, graphic forms are not substitutes for tables, but they are additional tools for the researcher to emphasise the research findings.

There are several ways in which you can present numerical data. In this chapter, we explain the various ways in which we can do this. It offers practical advice on how to incorporate numerical information into dissertations and doctoral thesis. The chapter outlines the role of text, tables, graphs and charts as formats for presenting numerical data. It focuses on issues that should be addressed when presenting numerical data for different audiences and suggests ways that will maximise the impact of such data and ensure that they are easy to read and interpret. Finally, the chapter looks at a number of different aspects of dissertation/thesis writing to assist researchers in the process of writing and presenting their thesis.

Making Data Meaningful

Both tabular and graphic presentation must be planned with utmost care and diligence. For instance, numbers are most effective in the main body of the text of an essay, report or dissertation when there are only two values to compare. If you are presenting and discussing three or more numbers, including them within the main body of text does not facilitate comprehension and/or comparison and it is often more pertinent to use a table incorporated within the text. All the salient features of data presentation by tables have been discussed in earlier chapters.

While presenting data in graphical forms, you must always keep in mind that graphs that you are going to use should be simple, clear and accurate, and also be appropriate to the data. In planning this work, the following must be considered:

- What is the purpose of the diagram?
- What facts are to be emphasised?
- What is the educational level of the audience?
- How much time is available for the preparation of the diagram?
- What kind of chart will portray the data most clearly and accurately?

Forms of Data Presentation

There are generally three forms in which we can present data:

- Textual or descriptive presentation
- Tabular presentation
- Graphical presentation

Textual or Descriptive Presentation

In this form of data presentation, all data are presented in the form of texts, phrases or paragraphs. It involves enumerating important characteristics, emphasising significant figures and identifying important features of data. Let us consider the following example.

Example

Suppose a researcher is asked to present the performance of a section in the statistics test. The following are the test scores of 40 students in the class.

34	42	20	50	17	9	34	43
50	18	35	43	50	23	23	35
37	38	38	39	39	38	38	39
24	29	25	26	28	27	44	44
49	48	46	45	45	46	45	46

A possible way of presenting this textual information in a meaningful way could be as follows.

In the statistics class of 40 students, 3 obtained the perfect score of 50. Sixteen students secured a score of 40 and above, while 3 got 19 and below. Generally, the students performed well in the test with 23 or 70 per cent students getting pass score of 38 marks and above.

The presentation of data in a textual form has both advantages and disadvantages. The advantage is that the data would be more interpreted, because there is a direct explanation of how the data happened to be like that. It may appear overwhelming and too long so its disadvantage is that it cannot interpret data in an easier way that majority could understand.

Further details of textual presentation are given in the section 'Writing the Research Report'.

Tabular Presentation

This type of presentation is already discussed in earlier chapters dealing with data collection and analysis.

Graphical Presentation

A graph is a visual representation of a relationship between, but not restricted to, two variables. Graphs are an appropriate and good means of describing, exploring or summarising numerical data, because the use of a visual image can simplify complex information and help to highlight patterns and trends in the data. They are especially an effective way of presenting voluminous data in a simple visual format and can also be used instead of a table to present smaller data sets. There are many different types of graphs to choose from but a key consideration is to ensure that the graph type selected is the most appropriate for the data. Having done this, you should then ensure that the design and presentation of the graph help the reader or audience interpret the data.

A summary of the types of data that can be presented in the most common types of graphs is provided further together with some general guidelines for designing readily understandable graphs. You can present several types of data and present them in the most common types of graphs. These graphs that are commonly used in research reports.

When and Why to Use Graphs

It is important for the researcher to have some basic knowledge of how to convey information graphically in presenting statistics. In deciding how best to present his or her data, the researcher should always consider the purpose of his or her graph or chart and the information that he or she intends to present and highlight. It should then be followed by the choice of selecting variables he or she wants to include and the manner such as frequencies, percentages or categories in which these variables should be presented.

When the researcher decides what kind of graph or chart best illustrates his or her data, he or she should consider what type of data he or she is working with. The main reasons for deciding when graphs should be used are for:

- Making comparison: How much? Which item is bigger or smaller?
- Investigating changes over time: How does a variable evolve?

- Understanding frequency distribution: How are the items distributed? What are the differences?
- Establishing correlation: Are two variables linked?
- Assessing the relative share of a whole: How does one item compare to the total?

Likewise, the reasons for using different types of graphs are obvious because:

- Graphs are quick to draw and are direct in presenting and sorting out critical information.
- They highlight the most important facts.
- They facilitate understanding of the data.
- They can convince readers.
- They can be easily remembered.

Guidelines for Formatting Graphs

Charts and graphs have found their way into news, presentations and comics, with users from art to design to statistics. The design principles for these data graphics will vary depending on what you are using it for. You should always keep your charts and graphs extremely simple and avoid using a lot of text.

Once you decide that a chart is the most appropriate way to present your data, then no matter what type of chart you use, you need to keep the following guidelines in mind:

- Check the data: This should be obvious. Data form the basis of charts and graphs. If your data are weak, your graph would be weak. It means your data should make sense. Just start with some simple and easy graphs to see if there are any outliers or weird spikes. Verify anything that does not make sense. You might be surprised by the number of data entry typos you find in the spreadsheets people send you.
- Explain encodings: You can use a colour scale to indicate magnitude or the size of a square to represent values. It can be a combination of both. Explain what these encodings are supposed to indicate, and do not assume the reader knows what everything means. Most likely, he or she does not.

  You can provide explanations in a variety of ways, but the most common are providing a legend, directly labelling shapes, or describing your graphic in a lead-in paragraph. Without your pointers, it would be a guessing game for the reader.
- Define your target audience: What does the audience know about the issue? Take into consideration who and what your graphs and charts are for, and design the graphs accordingly. You might design highly detailed graph for a poster that people can look at for hours. But if it is for a presentation, you should keep the words to a minimum. In other words, describe your story clearly and communicate the data accurately.
- Determine the message you want to communicate: What do the data show? Is there more than one message?
- Determine the nature of your message: Do you want to compare items, show time trends or analyse relationships in your data?
- Keep it simple and avoid flashy special effects: Present only essential information. Avoid using three-dimensional bars that confuse the reader.

- Title your graph clearly to convey the purpose: The title provides the reader with the overall message you are conveying.
- Specify the units of measurement on the *X*-axes and *Y*-axes: Without labels or any explanation, the grid lines are just decoration. You should show the axes of the graph by appropriate labels so that readers know what scale points are plotted on. Include some units while you are at it. If you just leave it with naked numbers, it could mean anything from a percentage, to a volume, to the number of students who drop out from the school. You should eliminate the need for any guesswork from the reader. Years, income, teachers, number of participants trained and type of school personnel are examples of labels for units of measurement.
- Label each part of the graph: In case of too much information, it is better that you show a legend to label each part of the graph (see the line graph). It would be meaningful if you use different shades, colours or variations in patterns to help the reader distinguish categories and understand your graph or chart.
- Include your sources: You should always include the sources of data—where the data come from—in your graph. You can put it directly in a graph, or if it is part of an article, the source can be specified in the copy. It is necessary for two reasons. First, your graphs becomes authentic and more reputable, and second, readers who are interested can look for deeper or fact check the contents of your graph.

Table 14.1 suggests a list of appropriate terms to describe your graph.

According to the United Nations Economic Commission for Europe (United Nations Economic Commission for Europe 2009), your graph:

- Should draw and grab the reader's attention
- Should present the information simply, clearly and accurately
- Should not mislead
- Should display the data in a concentrated way (e.g., one-line chart instead
- of many pie charts)
- Should facilitate data comparison and highlight trends and differences
- Should illustrate messages, themes or storylines in the accompanying text

TABLE 14.1 Checklist of Appropriate Terms to Describe the Graph

If your graph…	Use the Following Terms
Describes components	Share of per cent of the smallest, the majority of
Compares items	Ranking, larger than, smaller than, equal to
Establishes a time series	Change, rise, growth, increase, decrease, decline, fluctuation
Determines a frequency	Range, concentration, most of, distribution of *X* and *Y* by age
Analyses relationships	Increase with, decrease with, vary with, despite, correspond to, relate to

Source: Statistics Canada: Learning Resources, Using Graphs.

When Is It Not Appropriate to Use a Graph?

A chart is not always the most appropriate tool to present statistical information (refer to Figures 14.1–14.4). Sometimes a text and/or data table may provide a better explanation to your audience and save you considerable time and effort. You should reconsider using charts when your data:

- **Are very dispersed**

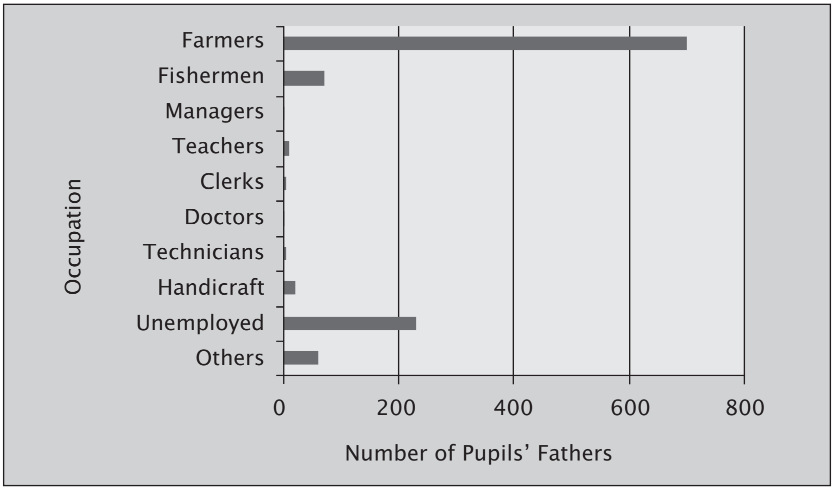


FIGURE 14.1 Division of Pupils' Fathers by Occupation in a Village

Source: Authors.

- **Have too few values**—there are too few data (only one, two or three data points).

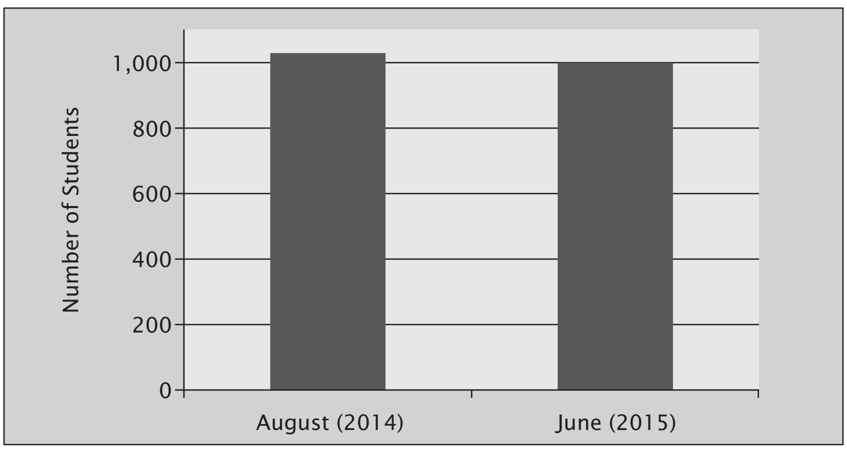


FIGURE 14.2 Enrolment in a Secondary School

Source: Authors.

- **Have too many values**

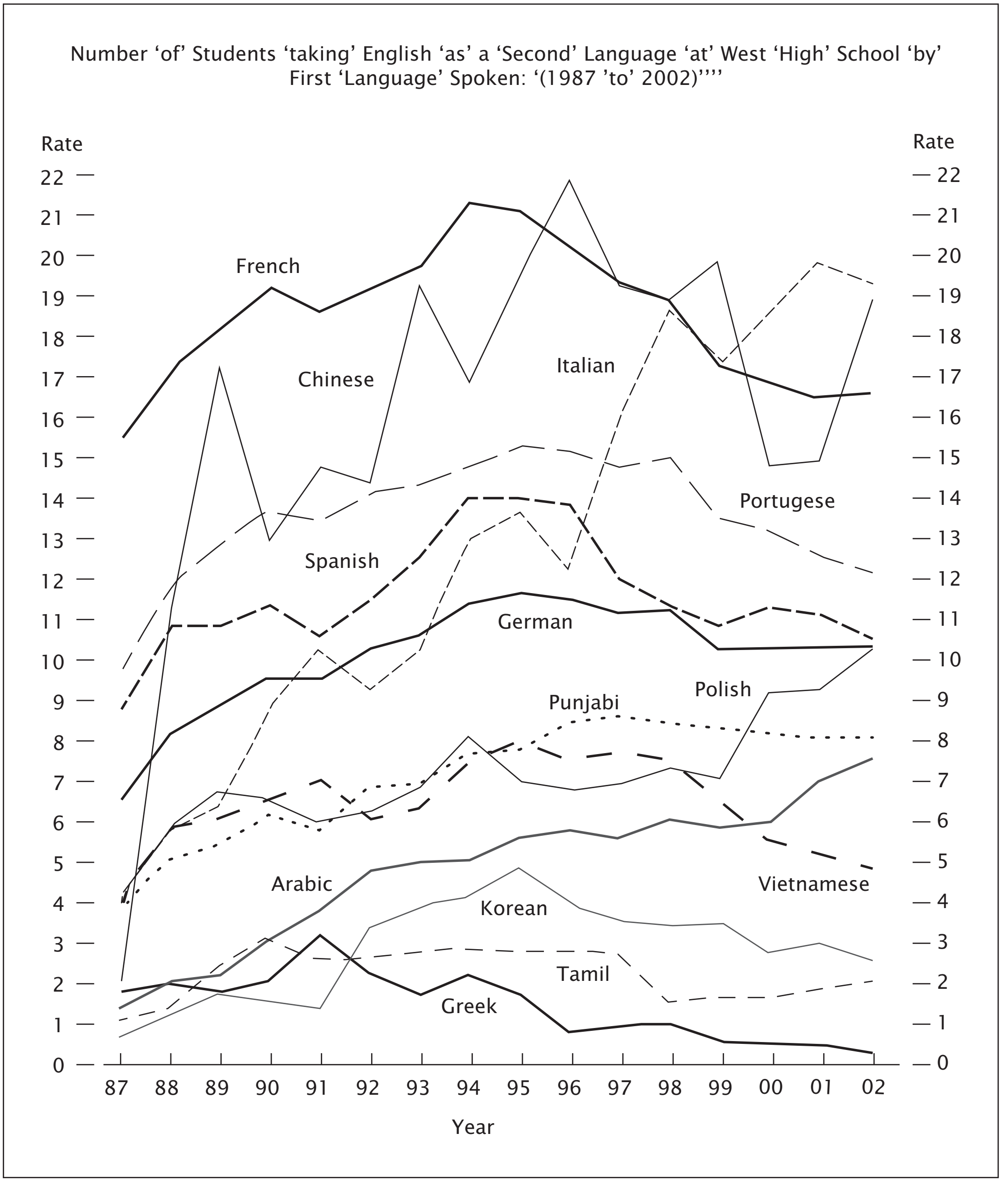


FIGURE 14.3 Example of a Bad Graph

Source: Statistics Canada, Learning Resources (Using Graphs)

- **Data show little or no variations**

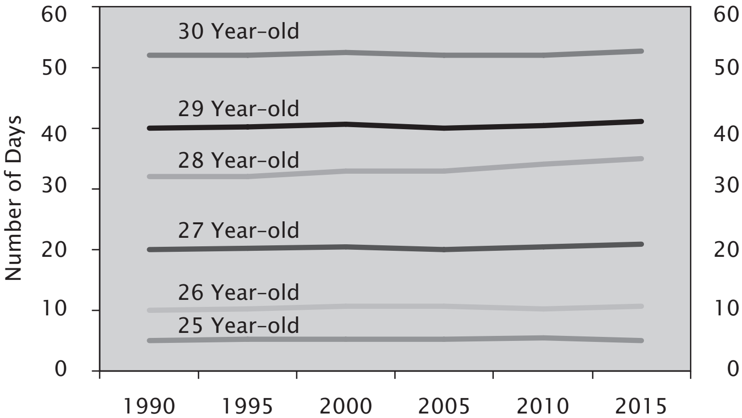


FIGURE 14.4 Number of Youth Illiterates Who Attended Literacy Class by Age in a Community (1990–2015)

Source: Authors.

Components of a Good Graph

Although there are many different types of graph, there are a number of elements that are common to the majority of them such as axes. Here we provide some general guidelines to help you design your graph and ensure that you apply these elements in a way that will help the reader or audience interpret the data you are presenting. The different components of a graph are shown in Figure 14.5.

In Figure 14.5, there are several components of a graph but basically the chart components comprise three main categories.

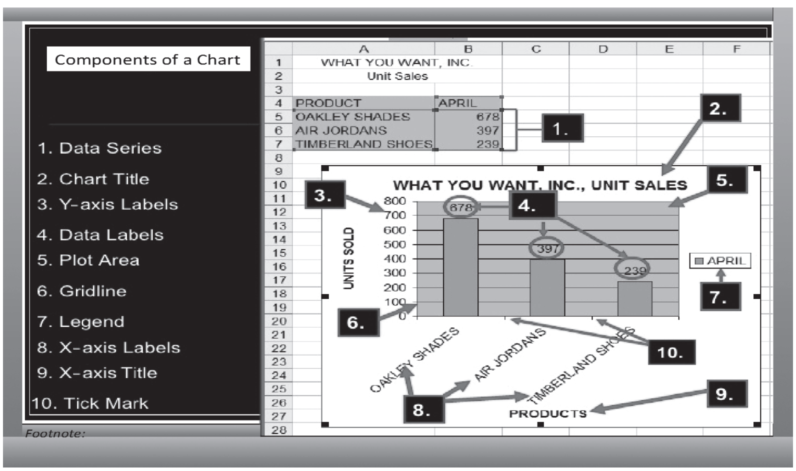


FIGURE 14.5 Components of a Graph

Source: Adapted from Hubert Hunter's Understand Charts and Graphs Used in Business.

- Data components: These components represent the data: bars, lines, areas or points.
- Support components: Those components that assist in understanding the data: title, legend, data labels, gridlines, footnotes and data source.
- Decorative features: Those components which are not related to the data.

A brief description of all these components is given further.

- Chart title: It describes the purpose and content. The title should be such that it gives a clear idea of what the chart is about. It has to be short and concise. The titles can be of two types:
  - Informative title: It provides all the information needed to understand the data. It should answer the three questions 'what', 'where' and 'when'.
  - Descriptive title: It is a caption that highlights the main pattern or trend displayed in the chart. It narrates in a few words the story that the chart illustrates.
- Y-axis: It is the left vertical side; it contains the numerical data.
- X-axis: It shows the bottom horizontal side; it contains the category information.
- Axis titles: It describes the X- and Y-axis data.
- Data markers: It represent values.
- Data series: A collection of related values from the worksheet; one row/column on the spreadsheet.
- Gridlines: This is a horizontal or vertical line that extends across the plot area of the graph to make it easier to read and understand the values.
- Plot area: The rectangular area bound by the category and value axes.
- Tick mark: It is used to add clarification to the data categories.
- Legend and data labels: They are the objects that explain the symbols, colours or pattern used to differentiate the data. You should not display the legend when only one series of values is represented in the chart. Whenever possible, you should use data labels rather than a legend. Data labels should be displayed on or next to the data components (bars, areas, lines) to facilitate their identification and understanding.
- Footnote: It is used to provide definitions or methodological information.
- Data source: It should be placed at the bottom of the chart.

Advantages of Using Graphs

Graphs and maps are useful tools to present the outcomes of our data analysis. There are several advantages of graphical presentation. The important ones are as follows:

- Portraying data graphically reveals patterns in the data that are difficult to detect otherwise.
- Visual depictions of data are almost universally understood without requiring knowledge of language.
- A graph is the well-designed presentation of interesting data—a matter of substance, of statistics and of design.
- A graph consists of complex ideas communicated with clarity, precision and efficiency.
- A graph gives the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space.

- A graph is almost always multivariate.
- A graph requires telling the truth about the data (Tufte 1983).
- The automated graph are very important and by comparing actual data with the data in graphical form, you can obtain an overview of the data, to explore their meaning, to discover regularities and irregularities, to suggest, test or discard scientific hypotheses.
- Animation of graphs may constitute an additional tool of research, mainly in regard with dynamic phenomena.
- Graphs may be of help in communicating the statistical information needed to make decisions by politicians, and business and administrative executives.

Distortion in Graphs

It should be noted that graphs used for the various purposes require different properties. For instance, while graphs for executives should be easily understandable, graphs for internal use by statisticians or scientists may be more sophisticated and may not require the same degree of properties. On the other hand, pleasant and attractive graphs, such as those obtained today by the automatic use of colour, may have important applications as illustrative and educational tools.

However, all graphs should be produced in a way which guarantees objectivity and honesty in the transmission of information, and which give a trustful representation of the data, apt to be interpreted quickly and in a more or less similar way by their readers. To reach this aim, they should stand up to acceptable scientific standards.

Use of graphs and maps is not free from limitations. A distorted graph can deceive the reader to think something other than what the data really say.

Graphs get distorted in several ways. Perhaps the most common way that graphs get distorted is when the distance along the vertical or horizontal axis is changed in relation to the other axis, that is, axes can be outstretched or contracted to create any desired result. For example, if you were to contract the horizontal axis (*X*-axis), it could make the slope of your line graph appear steeper than it actually is, giving the feeling that the results are more dramatic than they are. Similarly, if you stretch the horizontal axis while keeping the vertical axis (*Y*-axis) the same, the slope of the line graph would be more gradual, giving the impression that the results are less significant than they really are.

While creating and editing your graphs, you should make sure that the graphs do not get distorted. Oftentimes it may happen by accident when editing the range of numbers in an axis, for example. Therefore, you should pay attention to how the data look in the graphs and make sure that your results are being presented accurately and appropriately so as to not cheat the readers.

Tufte and Cleveland Principles of Graphical Integrity

In view of the distortions discussed before, Tufte (1983) and Cleveland (1994) suggested the following principles of graphical integrity:

- 'Show data variation, not design variation'. Tufte calls superfluous graphical elements 'chart junk'.
- Never use 3-D effects. Never.
- No need to entertain or distract the reader (Tufte 1983).

Cleveland (1994) suggests:[1]

- Make the data stand out. Avoid superfluity.
- Use visually prominent graphical elements to show the data, that is, do not obscure data by using ambiguous or overlapping symbols, inadequate spacing, poor labelling and so on.
- Strive for clarity.
- Proofread graphs.
- Visual clarity must be preserved under reduction and reproduction.
- Make captions comprehensive and informative.
- Draw attention to the important features of the data.
- Clear, detailed and thorough labelling should be used to defeat graphical distortion and ambiguity.
- The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities represented.
- The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.

In short, the purpose of graphical data presentation is to communicate information clearly and accurately.

Types of Graphs

Different types of graphs are used to depict different situations. It is for this reason we explain further to inform you a little bit about what the available graphs are. The selection of an appropriate graph quite often is determined by the kind of data we wish to show. For example, in this section, graphs would be used to show the quantified results obtained from the analysis of data discussed in earlier chapters on data analysis.

Graphs are pictures representatives for a single or more sets of information and how these visually relate to one another. There are varieties of charts and graphs of diverse complexity. You can represent almost any numerical data set with a graph type that is appropriate for its display. Graphs show your data in a meaningful way. It is one thing to see a data listed on a page and it is another to actually understand the details and trends of the data.

Bar Charts

A bar graph (Figure 14.6), also known as a Pareto diagram, is the simplest type of chart to draw and interpret. Bars are used to display and compare the number, frequency or other measure (e.g., mean) for different discrete categories or groups. In other words, bar graphs are best means to compare two different entities that are similar in their content. The bars can be either vertical (up and down) or horizontal (across). The data can be shown in both words or numbers. Example might include data on

[1] Also see https://www.amazon.com/Elements-Graphing-Data-William-Cleveland/dp/0963488414 (accessed on 17 November 2017).

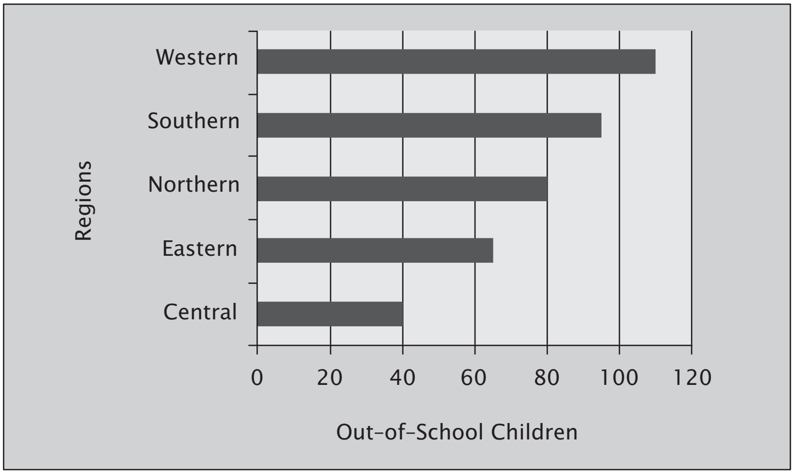


FIGURE 14.6 Bar Chart: Out-of-school Children in Country X in 2015 ('000)

Source: Authors.

number of teachers in primary, secondary and higher levels of education, number of out-of-school children by regions in a hypothetical country Ardania.

Stacked Bar Chart

A stacked bar chart is used to show and compare segments of totals. Caution should be exercised when using this type of chart. It can be difficult to analyse and compare if there are too many items in each stack or if many items are fairly close in size (Figure 14.7).

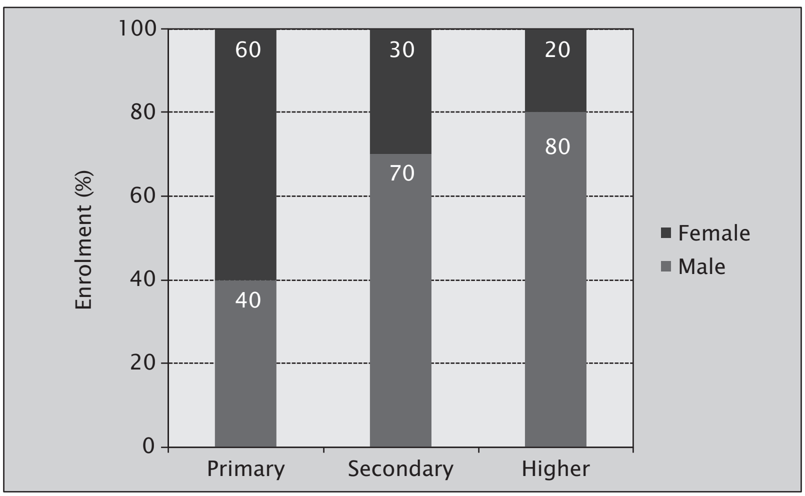


FIGURE 14.7 Stacked Bar Chart: Total Enrolment by Levels of Education and by Sex in 2015 (%)

Source: Authors.

Population Pyramid Chart

Population pyramids are used by researchers in general and demographers in particular to represent the age and sex distribution of a particular population at a specific point in time. They are easy to build. A population pyramid is a specifically formatted comparative histogram (Figure 14.8).

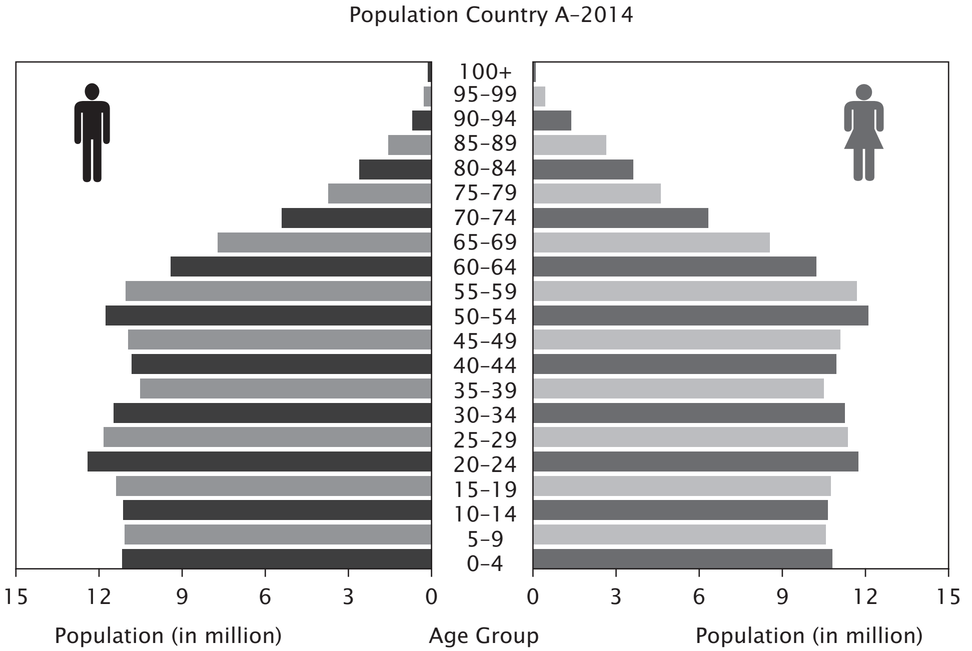


FIGURE 14.8 Population Pyramid: Population of Country A by Sex and Age Group in 2014

Source: Authors.

Line Chart

Line charts (Figure 14.9) display continuous data over time, set against a common scale, and are therefore ideal for showing trends in data at equal intervals or over time. Line chart is the most appropriate type of chart for time series. For example, if you want to see how steadily the youth unemployment in the East Asia, South Asia and Latin America and the Caribbean regions has been growing, the line chart can help you understand the overall trend without going into the individual values themselves. Line graphs show you how numbers have changed over time. They are used when you have data that are connected, and to show trends, for example, average monthly attendance in primary school in each month of the school year.

In a line chart, category data are evenly distributed along the horizontal axis, and all value data are distributed evenly along the vertical axis. As a general rule, you should use a line chart if your data have non-numeric x values. For numeric x values, it is usually better to use a scatter chart. You can modify and tune the chart parameters to better communicate your message, but you should be careful not to twist the data.

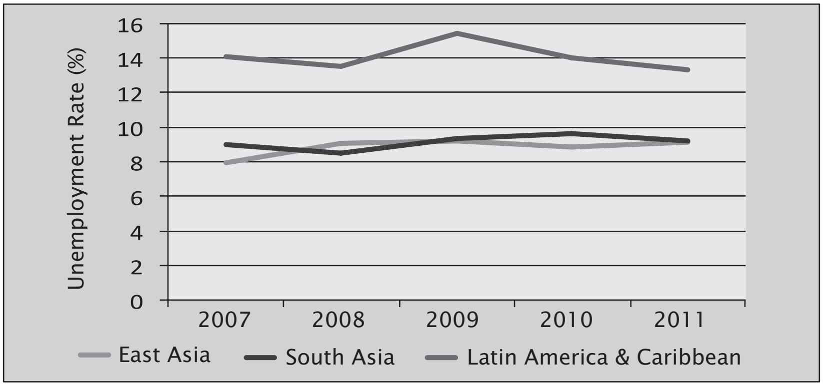


FIGURE 14.9 Line Chart: Youth Unemployment Rates, 2007–2011

Source: Databank, ILO.

Dual *Y* or Combination Chart

When the values in a chart vary widely from data series to data series, you can plot one or more data series on a secondary axis. A seconday axis can also be used as part of a combination chart when you have mixed types of data (e.g., pupils' survival rates and cost per pupil, price and volume sold and so on) in the same chart. This type of chart is known as combination chart. It is shown in Figure 14.10.

A dual *Y* chart is a chart that combines two or more charts types in a single chart, for example, a column charts series and a line charts series. In a dual *Y*-axis combination charts, each axis has its own unit and magnitude, and each data series conforms to one of these axis. For example, in Figure 14.10, if you have to compare survival rates with cost per pupil in primary level of education over a period, you can use the column series on primary-axis to show unit cost per pupil, and line series on secondary-axis to show survival rate. The survival rate figures on primary axis will reflect per cent values, whereas line series on secondary axis will reflect unit values.

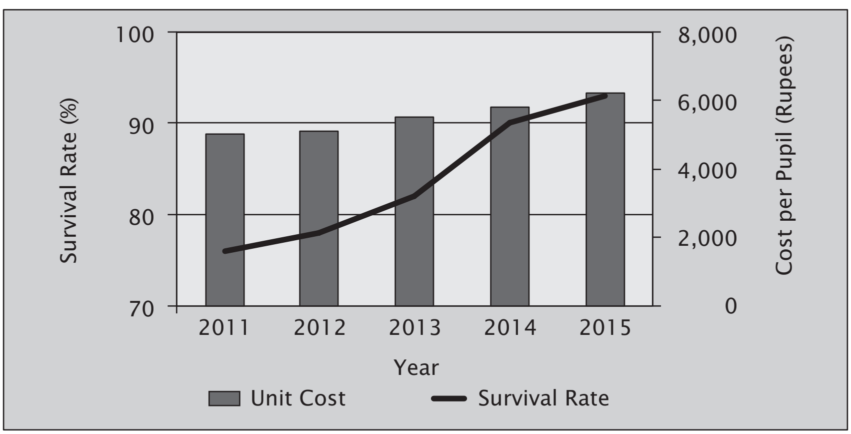


FIGURE 14.10 Dual *Y* or Combination Chart: Survival Rates and Unit Cost, 2007–2011

Source: Authors.

Histogram

Histogram is a graphical display of data using bars of different heights. It is similar to a bar chart, but in a histogram, a variable is grouped into several ranges. You are the one who decides what ranges to use (Figure 14.11).

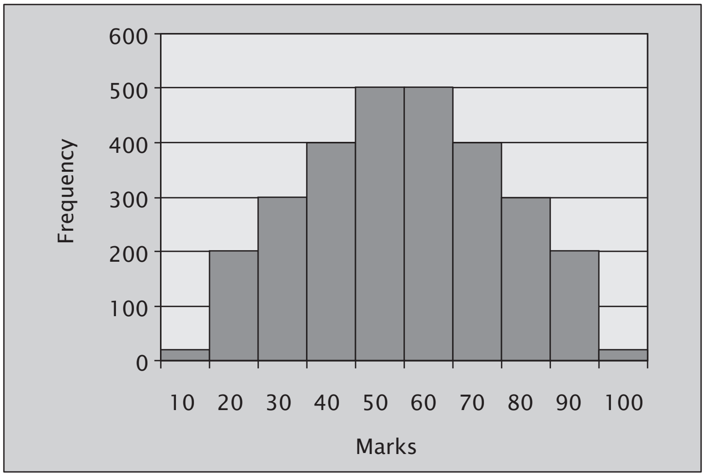


FIGURE 14.11 Histogram: Frequency Distribution of Pupils' Score

Frequency Curve

A frequency curve is a graphical device for understanding the shapes of distributions. It serves the same purpose as histograms, but is especially helpful for comparing sets of data. It is a smooth curve which corresponds to the limiting case of a histogram computed for a frequency distribution of a continuous distribution as the number of data points becomes very large (Figure 14.12).

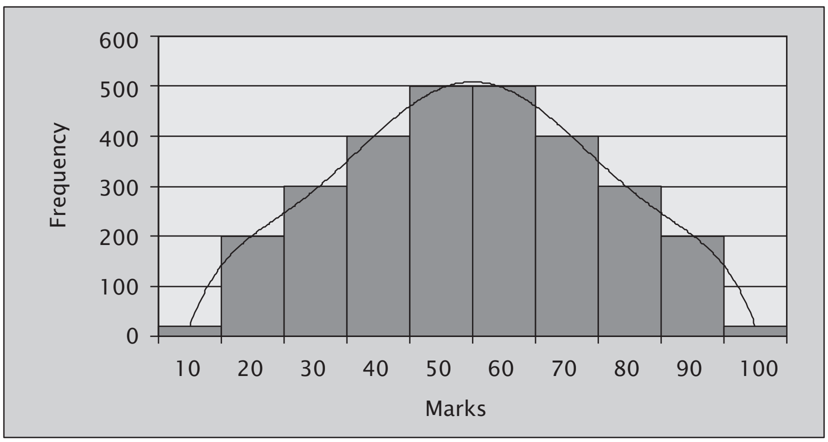


FIGURE 14.12 Frequency Curve: Frequency Distribution of Pupils' Score

Frequency Polygon (Ogive)

Frequency polygons are also used for displaying cumulative frequency distributions. A frequency polygon is a simple way of displaying quantitative data. It is usually used for showing ungrouped data (a histogram is more commonly used for grouped data). It is a plot of points whose *x*-coordinate represents values that are observed and the *y*-coordinate is a count of the number of times the value was observed. Figures 14.13 and 14.14 show frequency polygon and cumulative frequency polygon, respectively.

A cumulative frequency polygon is obtained from frequency distribution by joining the point in the successive interval time. Using this graph, we can also find the shape of frequency distributions which is known as a frequency polygon.

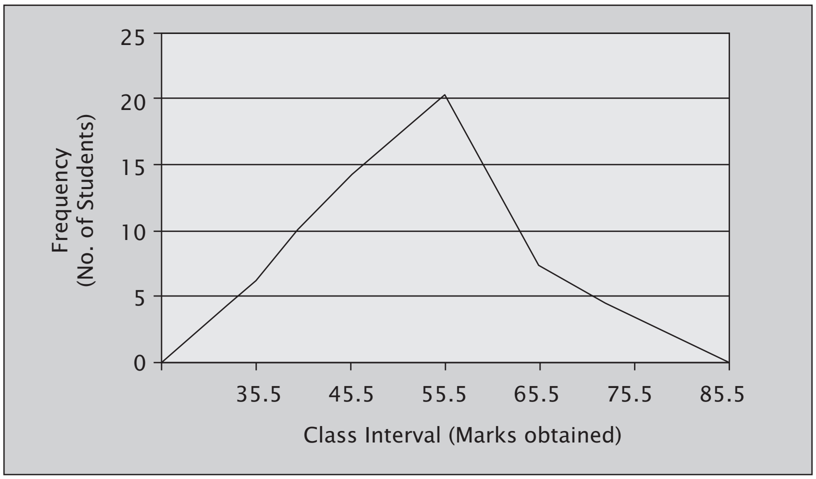


FIGURE 14.13 Frequency Polygon (Ogive): Math Test Score

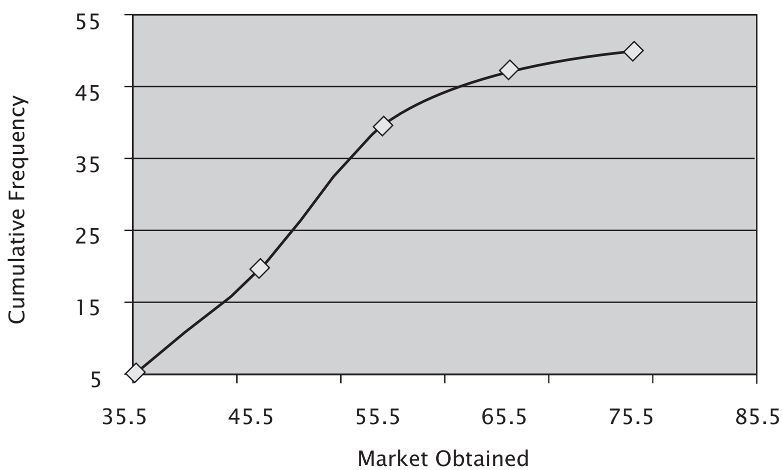


FIGURE 14.14 Cumulative Frequency Polygon: Math Test Score

Scatter Graphs

Researchers use scatter plots to show the relationship between pairs of quantitative measurements made for the same object or individual. For example, a scatter plot could be used to present information about the examination and coursework marks for each of the students in a class or mothers' year of schooling and number of children and so on. Figure 14.15 shows the paired measurement of number of children and mothers' schooling (years).

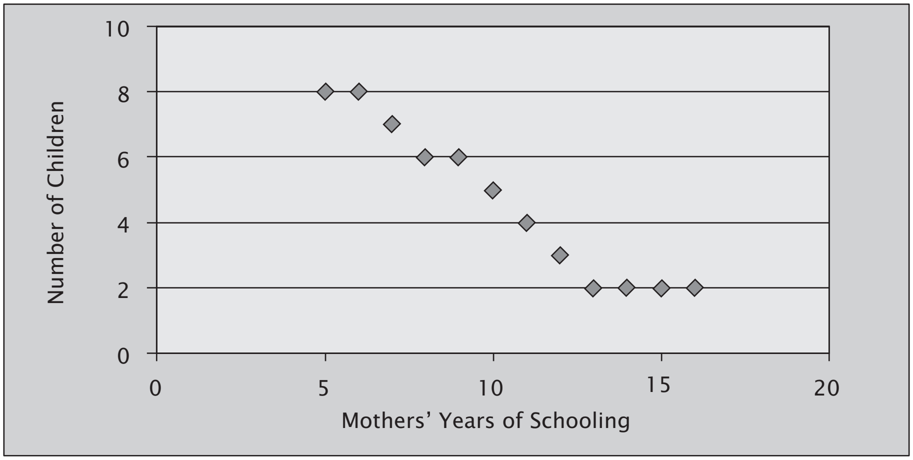


FIGURE 14.15 Scatter Plot Chart: Number of Children Decreases as Mothers Become More Educated

In a scatter plot, a dot represents each individual or object (mother in this case) and is located with reference to the *X*-axis and *Y*-axis, each of which represent one of the two measurements. By analysing the pattern of dots that make-up a scatter plot, it is possible to identify whether there is any systematic or causal relationship between the two measurements. For example, in this case it is clear from the downward trending pattern of dots that educated mothers would like to have less and less number of children, that is, number of children decreases with mothers' years of schooling. Regression lines can also be added to the graph and used to decide whether the relationship between the two sets of measurements can be explained or if it is due to chance.

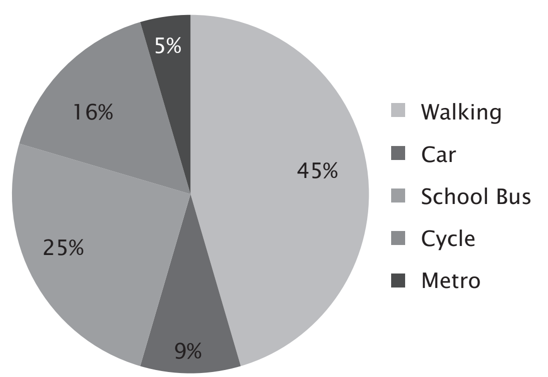


FIGURE 14.16 Pie or Circle Chart: Modes of Travelling to School

Pie or Circle Chart

Pie chart (or a circle chart) is a circular chart divided into sectors, illustrating numerical proportion (Figure 14.16). It is clear from the following pie chart that the arc length of each sector (and consequently its central angle and area) is proportional to the quantity it represents. While it is named for its resemblance to a pie which has been sliced, there are variations in the way it can be presented. Pie charts show us as to how a whole is divided into different parts. You might, for example, want to show how the public budget on education had been spent on different levels of education in a particular year.

The proportional segments give a clear picture of the relationship among the component parts. The percentages are placed inside each slice. A very small segment is shown by an arrow which is drawn to it and the legend is placed outside.

Pie charts are very widely used in education research, the business world and the mass media. However, pie charts have been criticised, and many experts recommend avoiding them, pointing out that research has shown it is difficult to compare different sections of a given pie chart, or to compare data across different pie charts. Pie charts can be replaced in most cases by other plots such as bar charts.

Pictograph

Nothing is as visually appealing as a picture. Pictograms represent huge volumes of data by a simple repetition of the same picture. It does require an explanation at one corner of the chart to let the viewers know what the value of a single picture is. These serve the purpose of grabbing attention, where in the purpose of transferring the information is not of much importance (Figure 14.17).

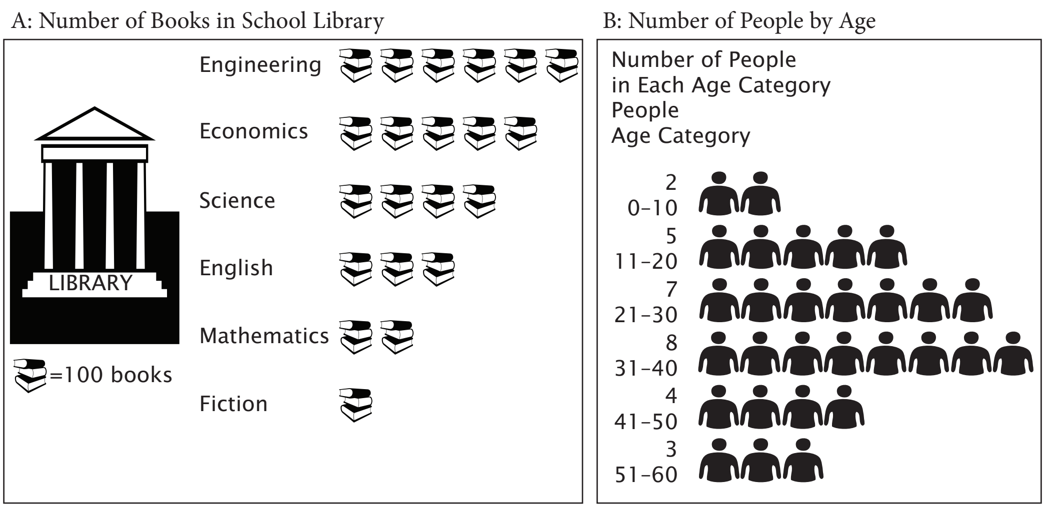


FIGURE 14.17 Pictograph

Flow Charts

A flow chart is a diagram that shows step-by-step progression through a procedure or system especially using connecting lines and a set of conventional symbols. Each of the flow chart's shapes represents specific actions and the legend codes are included to inform you. Flow charts consist of inputs and outputs, a start point and an end point that basic symbols represent.

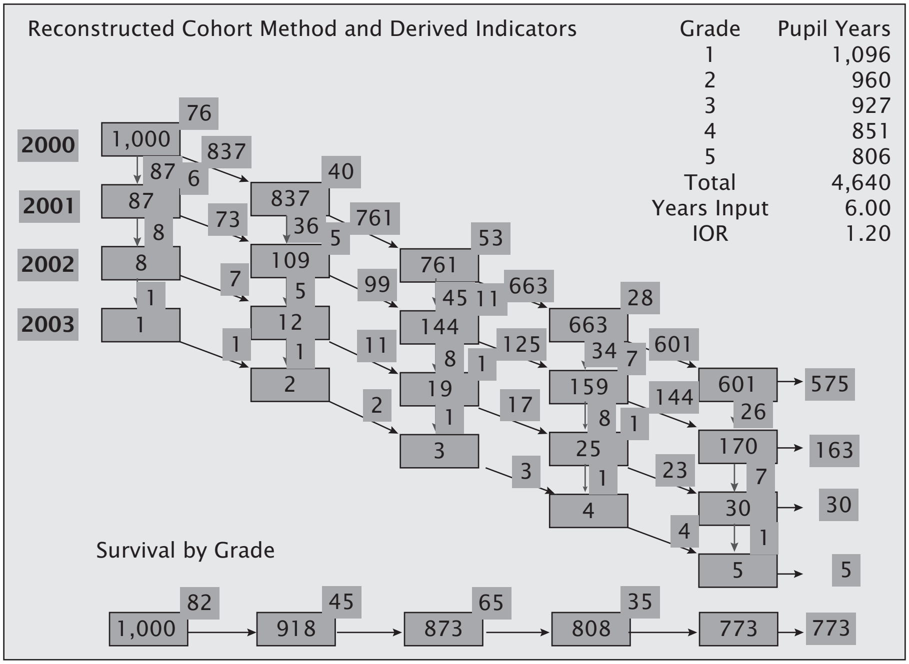


FIGURE 14.18 Students Flow Chart

Figure 14.18 provides an explicit, logical and integrated framework of students' flow during the primary level cycle of education. It also provides explicit assumptions about the future development of the propensities of children to enter school as well as to be promoted, to be repeating, dropping out, transferring or graduate once they are enrolled. It shows tools for reconstructing the flow of a given cohort of children through a cycle of education and measures for analysing the effect of repetition and dropout on the internal efficiency and on the extent of wastage of school system.

Area Charts

Area charts (Figure 14.19) are much like line charts, but they display different colours in the areas below the lines. This colourful and visual display distinguishes the data more clearly. One of the limitations of area charts is that data series with smaller values that are plotted in the back of an area chart may be completely or partially concealed behind data series with larger values that are plotted in front of them.

Organisational Chart

Organisational charts are diagrams that reveal the overall structure of the workforce of an organisation or company. Through an organisational chart, you can present the formal direct or indirect relationships between the positions in an organisation. The chart also shows how different departments are connected (Figure 14.20).

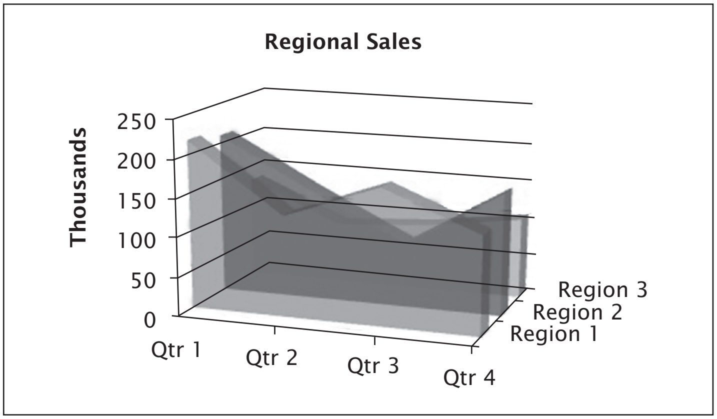


FIGURE 14.19 Area Chart

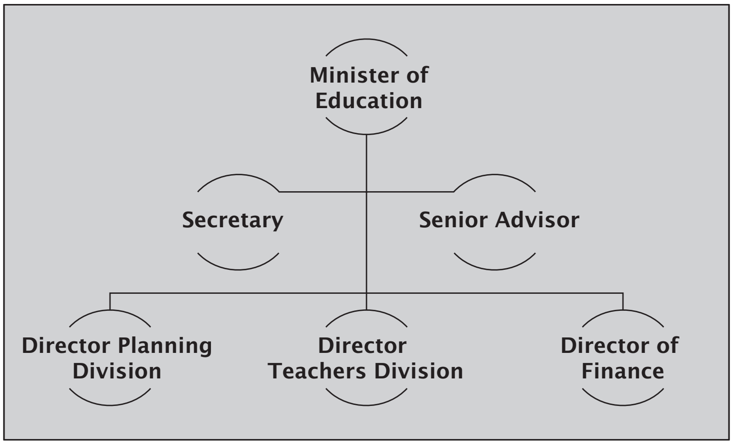


FIGURE 14.20 Organisational Chart: Structure of the Ministry of Education

Organisational charts are types of graphs that show four types of relationships: line, lateral, staff and functional.

- Line relationships exist between superiors and subordinates.
- Lateral relationships exist between different departments of similar rank.
- Staff relationships exist between a managerial assistant and a line manager.
- Functional relationships exist between a specialist and a line manager.

An organisational chart usually looks like a pyramid, with the president, chairman or director-general in the top rectangle and levels of subordinates in descending rectangles according to rank listed further. Each rectangle size corresponds to the level of authority. Thus, superiors have larger rectangles than subordinates. Peers have equally sized rectangles. Solid lines between rectangles highlight a direct relationship and dashed lines signify an indirect relationship. Arrows represent the direction of communication links and flow between the several rectangles (components) of organisational charts.

An organisational chart has its advantages. It encourages structure in an organisation and defines the roles of the management. It also reveals those parts of an organisation or company that need improvement.

An organisation chart also has limitations. It does not show anything about the managerial style. Moreover, organisational charts need to be changed every time an employee leaves or joins the company.

Bubble Chart

A bubble chart, shown by Figure 14.21, is a variation of a scatter chart in which the data points are replaced with bubbles, and an additional dimension of the data is represented in the size of the bubbles.

Bubble charts are the best means for the three-dimensional representation of the data. Along with the basic axis of a plane graph, the size of the bubble represents the third dimension. The bubble charts are growing in their popularity to analyse the teachers' efficiencies or performances. It is the precise method to compare a set of three entities.

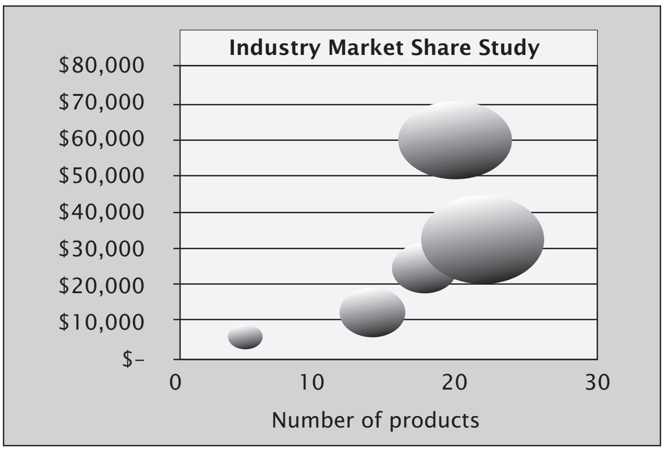


FIGURE 14.21 Bubble Chart

Radar Chart

A radar chart, also known as a spider chart or a star chart because of its appearance, plots the values of each category along a separate axis that starts in the centre of the chart and ends on the outer ring. For example from Figure 14.22, we can quickly conclude the following.

Country B is significantly worse than any other country as far as the PTR is concerned. Whereas Country E is significantly worse in terms of dropout rate in primary level of education. Country D is rated significantly better than any other country in terms of the two primary education indicators.

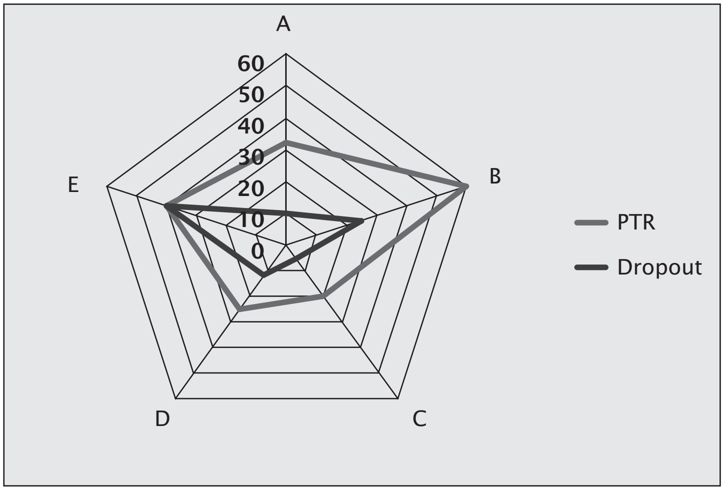


FIGURE 14.22 Radar Chart (Pupil–Teacher Ratio and Dropout Rates: Primary Education)

Source: Authors.

Drawing Graphs Using Computer Packages

You can use various computer software packages, including PowerPoint and Excel, to draw graphs.

However, be sure that these applications are somewhat limited in the type of charts that they can draw, and you may not find the results entirely what you expected. What you really need is a basic understanding of graphs and charts so that you can compare what the computer has generated to what you want to depict.

You can use computer applications to make your overly complicated graphs simple and easy to understand. A three-dimensional exploding pie chart may look 'cool' but does it help you or others to visualise the data? You should always make sure that your graphs and charts are simple and are neatly and clearly formatted.

Whatever type of graph or chart you choose to draw, once you have the knack of reading them, you will almost certainly find that the old saying is right: a picture really can tell a thousand words. It means that your graph should be an effective way of showing several numbers together and demonstrating the relationships or differences between them.

Maps

Statistical map is another way to display the geographic distribution of data. For example, let us say we are studying the geographic distribution of literacy rates in India. A statistical map (Figure 14.23) would be a great way to visually display our data. On our map, each category is represented by a different colour or shade, and the states are then shaded depending on their classification into the different categories.

Maps can display geographical data at various levels of disaggregation—on the level of cities, counties, city blocks, census tracts, countries, states or other units. This choice depends on your topic and the questions you are exploring.

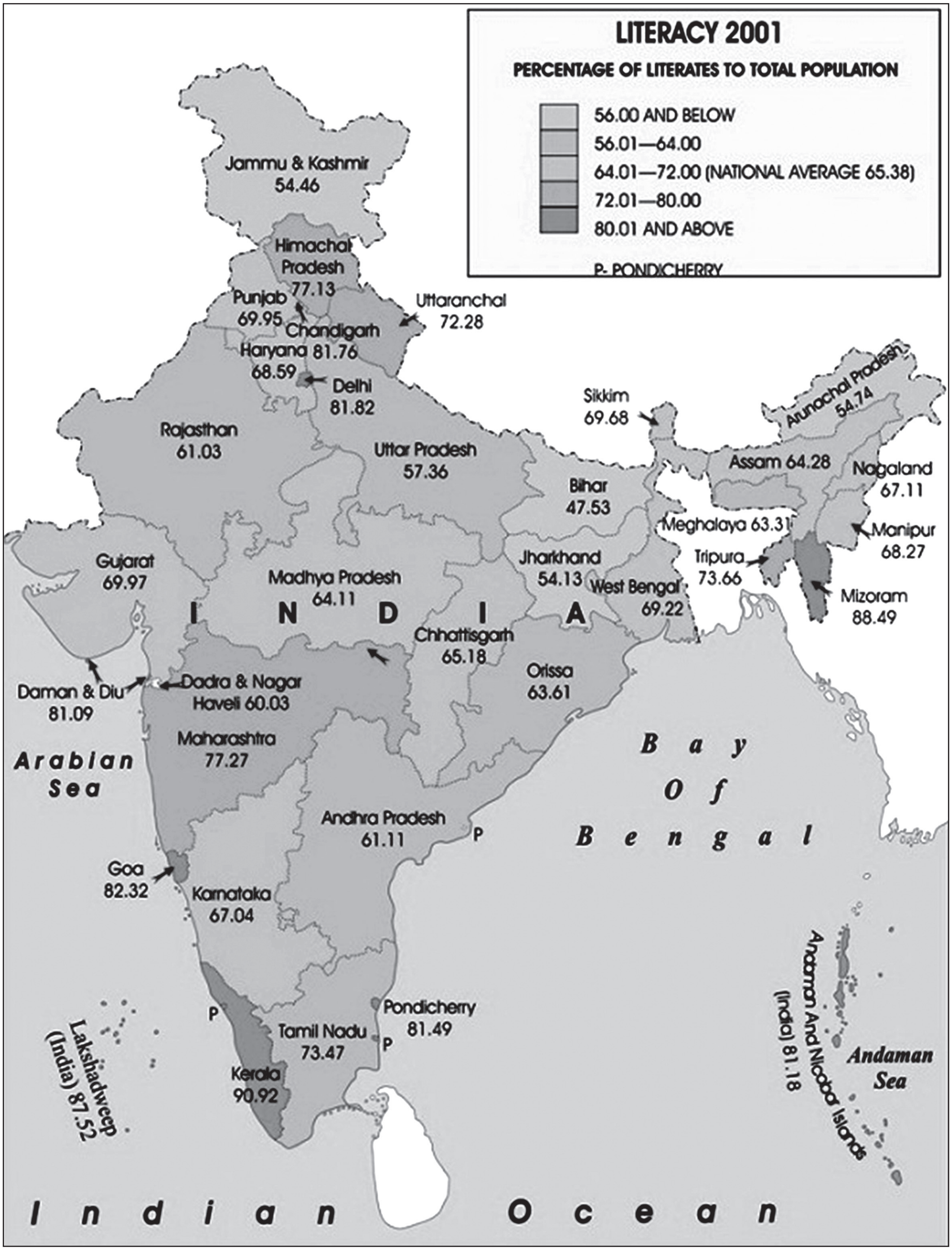


FIGURE 14.23 Map: Literacy Rates in India, 2011

Source: Government of India, Census 2011.

As for all data visualisations, the most important suggestion to ensure that your message gets across is: Keep your map simple. You should always be careful not to distract or confuse your audience by displaying too much information or too many visually conflicting elements.

In summary, graphs and maps are a popular educational tool. They help you to visually present and display, interpret and understand complex topics. They also assist in reading comprehension by allowing you to track main ideas, facts, plot, setting and characters. For special education researchers, these tools can help them to express and show an understanding of concepts that may be difficult for them to show with traditional written or essay assessments.

Finding, modifying and printing graphs and maps are easily accessible via Excel, SPSS and several similar data analysis and presentation packages. These database computer programmes can easily be adapted to assist all type of learners, topics and desired learning outcomes. There are several sites available now which students can use to create their own graphic organisers—they can edit, print and share via the Internet.

However, effective graph design is not always an intuitive process. When you create a graph, you design something that communicates through visual perception. Knowing something about how we perceive and interpret visual stimuli—objects made visible through light, possessing a particular combination of attributes such as colour, size and location in space—is a necessary conceptual foundation to effective graph design (Tufte 1983). Which visual attributes can reliably encode quantitative data? Which visual attributes can be used to make something stand out above others? Which colours work well together? What are the limits of short-term memory and how do limits apply to graph design? Why is it difficult to accurately compare the sizes of the slices in a pie chart?

We have provided in this section of the chapter answers to these questions in non-technical terms and many more that are relevant to graph design. You do not need to be a scientist or a graphic designer yourself to become familiar with these concepts and how they can be used to present data effectively.

A serious limitation or perhaps the common mistake that we have touched very briefly is only a sampling of a much larger list of graphs and map. You can perhaps identify many more without straining your brain in the least. The purpose here is simply to get you thinking about the importance of effective data presentation and noticing how some of your own design practices might need improvement.

Writing the Research Report

The thesis or dissertation writing is the final stage of a research project.[2] It provides us with the opportunity to show that we have gained the necessary skills and knowledge in order to organise and conduct a research project. Your dissertation should demonstrate and reveal that you are skilled in recognising and identifying an area or areas suitable for research: setting research objectives; locating, organising and critically analysing the relevant secondary data and authoritative and relevant literature; devising an appropriate research methodology; analysing the primary data selected and drawing on the literature in the field; drawing conclusions and, if appropriate, making relevant recommendations and suggestion on key of areas for further research.

[2] Before you begin writing, you should get the two essential writing manuals: Strunk Jr. and White (1999) and Turabian (2007). Get them and read them.

A 'dissertation' is a 'formal' document. There are specific 'rules' that govern the way in which it is presented. It must have chapters that provide an introduction, a literature review, a rationale and justification of the data selected for analysis and research methodology, analysis of the data and, finally, conclusions and recommendations.

It is important to remind that the master level dissertation or doctoral thesis is distinguished from other forms of writing by its attempt to analyse situations in terms of the 'bigger picture'. It solicits answers, explanations, makes comparisons and arrives at generalisations, which can be used to enlarge theory. As well as explaining, 'what can be done', it addresses the underlying 'why'. The most successful dissertations are those, which are specific and narrowly focused.

In fact, writing a dissertation or doctoral thesis is a massive undertaking and as such it should not be taken lightly. There are similarities and differences to perform each task. For instance, the master's dissertation is something like running a 100 m race—the course is usually very quick and there is not as much time for thinking as the researcher may perhaps want. By comparison, writing a PhD thesis is like running a marathon, working on the same topic for 3–4 years is laborious and can be quite exhausting. But in many ways the approach to both of these tasks is quite similar.

This section of the chapter looks at a number of different aspects of dissertation/thesis writing to assist researchers in the process of running their race—be it a 100 m race (sprint) or a marathon. It attempts to help them realise that whatever type of dissertation/thesis they are writing. It is in fact more like a 400 m race—it can feel neither too short nor too long if the researcher paces himself or herself and take the appropriate steps.

The main aim of the section is to provide key guidelines on the dissertation process. The discussion further only offers suggestions; it by no means guarantees the production of a fine piece of work, but these are suggestions, which, through time, have been found to be both practical and effective.

When to Start Writing the Dissertation Report?

A key question in report writing is: when to start writing the report. Quite often many of you undertake the task of writing the report once you have completed the data analysis and prepared the several tools (tables, charts, graphs, maps and so on) of presenting the relevant outcomes revealed by the data analysis.

There are no hard and fast rules for starting your dissertation. In fact, it all depends on the scope of the research study you are describing and on the duration of your course. In some cases, your research study may be relatively short and you may not be able to write much of your thesis before completing the study. However, in some other instances your study may be relatively long, especially if you are doing a PhD, and you will need to keep writing the thesis while conducting your research. But regardless of the nature of your research study and the scope of your course, you should start writing your thesis or at least some of its sections as early as possible, and there are a number of good reasons for this (Figure 14.24).

Layout of Dissertation

Dissertations need to demonstrate knowledge and understanding beyond undergraduate level and should also reach a level of scope and depth beyond that taught in the class. You must present your dissertation in an appropriate academic style and format to ensure that the precise aims of the dissertation are met and the readers understand its underlying messages. Your dissertation should state precisely its

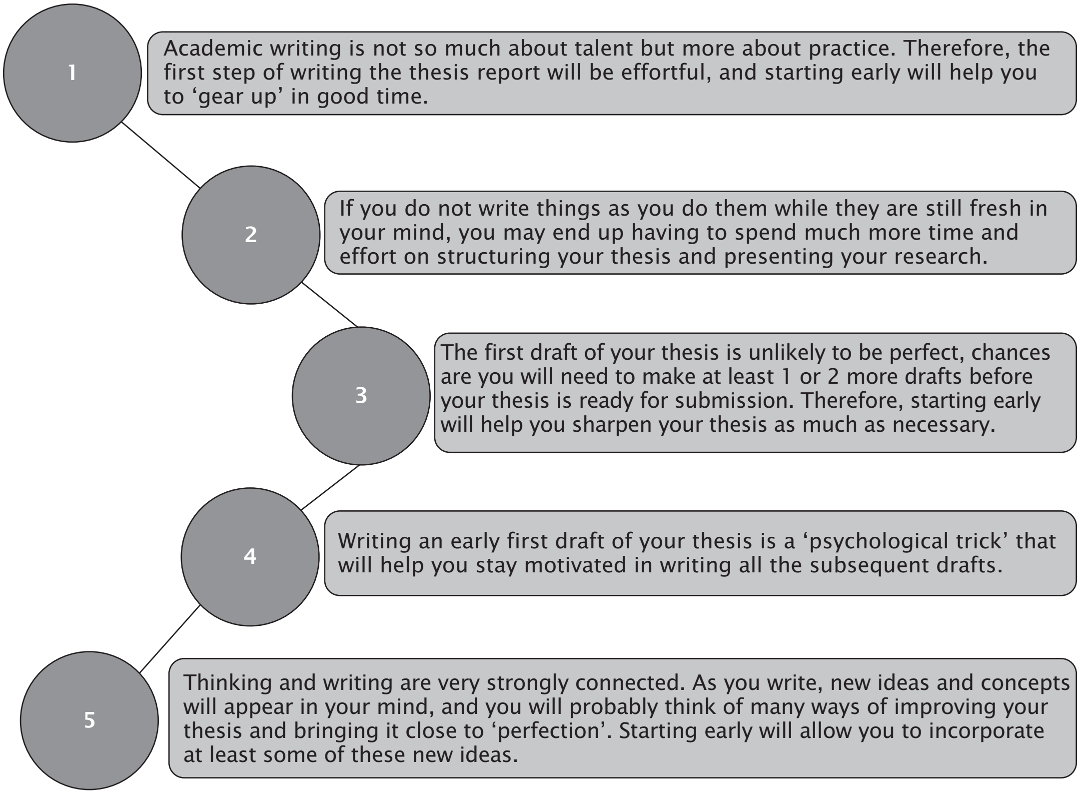


FIGURE 14.24 Reasons for Writing Early the Dissertation

Source: Adapted from POSTGRAD PROGRAM.

aims and objectives, and ensure that these are achievable within the scope of the dissertation framework. Academic style does not just refer to the clarity of expression, grammar, use of citation and referencing but it essentially relates to a clearly structured approach to the justification and validation of facts, theories and opinions presented to form a precise argument.

Writing a dissertation is a unique experience and there is no general consensus on what is the best way to structure it. Almost all universities and research organisations have very specific dissertation structures and designs; thus, as a researcher or postgraduate student, you will probably decide what kind of structure suits your research project best and meets the university standards and requirements. To some extent, all postgraduate dissertations are unique and the guidelines given here do not necessarily apply to everyone. Thus, prior to undertaking the task of writing your dissertation and in the event of you not being sure about any aspect of the process of writing up your work, it is advisable to consult and discuss with your supervisor as well as to read other theses of previous postgraduate students in your university library.

Generally, researchers follow a more conventional way of structuring a master's dissertation or postgraduate thesis where they write it in the form of a book consisting of chapters. Although the number of chapters used is relative to the specific research project and to the course duration, a thesis organised into chapters would typically contain the components shown in Table 14.2.

The layout of a research report refers to the list of various parts of the report/thesis. Usually, a good and meaningful research report should consist of the following three components:

- Preliminary pages
- Main text
- End matters

The following explains each of them in more details.

Preliminary Pages

Preliminary pages of your report should contain title of the report, acknowledgement, certificate page, list of publications and table of contents. Your acknowledgement should thank those who have helped you during your course of investigation. Acknowledgement should also be made at the time of public viva voce also.

Declaration in the certificate page by the researcher is generally done using phrases such as:

> I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of any university or institute of higher learning, except where due acknowledgment has been made in the text.

Your research report should have an abstract. It is a necessary part of any scientific and non-scientific research report. In a research article, it should appear next to the author's name and affiliation. However, in the case of master's dissertation or PhD thesis, before its submission an elaborated abstract of the thesis called synopsis has to be submitted to the institution where registration for PhD degree is made.

An abstract is the synopsis of your thesis. It carries the brief details about the report and conveys its essence. Your abstract should contain a very short statement of the problem under investigation, methodology and procedures followed in the work of your hypothesis; a brief summary of your findings and results of the study and a very brief conclusion. In fact, the abstract can act as a tool to control the flow of ideas in your thesis. It can help you link in a logical manner the reasons for the research and aims of the work. It should contain answers to the questions: What was done in the research project? Why is it of interest? How was it done? What were the outcomes of the work done? What is the significance of the results? You should make sure that your original contribution has been clearly emphasised in the abstract. The abstract of a master's dissertation should be one page and for PhD thesis it could be about three or four pages.

Table of contents should give title of the chapters, section headings, title of appendices and their page numbers. All the preliminary pages should be numbered with lower case roman numbers.

Preliminary pages in a dissertation generally contain the following elements:

- Title page
- Abstract
- Declaration
- Dedication
- Acknowledgement

- Table of contents with page numbers
- List of tables, list of figures or list of illustrations, with titles and page numbers (if applicable)
- List of abbreviations (if applicable)
- List of symbols (if applicable)

Title Page

The title page is the opening page of the dissertation. It should include the research title which has been approved by the faculty, name of candidate according to the registration records, his or her identity number (if applicable) and the statement '…submission of research report/dissertation/thesis for the fulfilment of the Degree of Master of …/Doctor of Philosophy' and finally the year of submission. An example of the title page is shown in Figure 14.25.

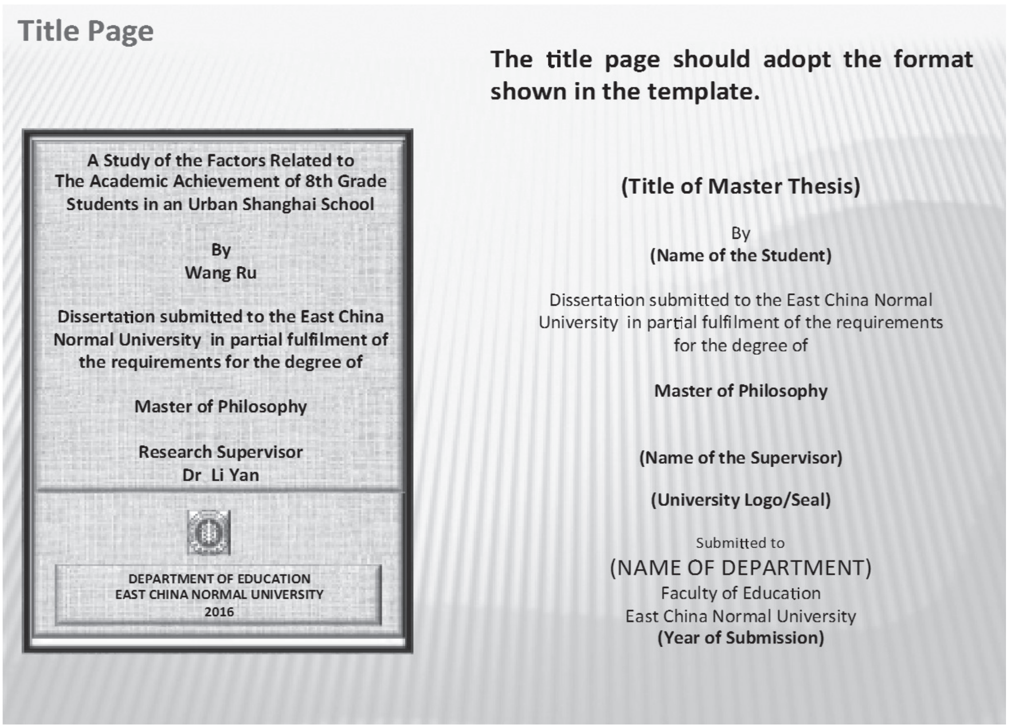


FIGURE 14.25 Example of the Layout of the Title Page of the Dissertation

Source: Authors.

Abstract

An abstract of a dissertation is the brief summary of the research study including background, methodology and findings. A dissertation should contain an abstract of up to 350 words (Box 14.1). The abstract should provide sufficient information about the results of the research so that reading the full dissertation is not necessary.

An abstract should always include the following:

- Aim and objectives: What are the main themes, ideas or areas of theory being investigated?
- Boundaries: What is the context and background to this dissertation? In what areas of theory or business practice should the reader concentrate their attention?
- Methodology: What was/were the main method(s) employed to generate the results?
- Results: What were your main findings?
- Conclusions: What are the main conclusions that you arrive at when viewing the entire dissertation?
- Recommendations (if appropriate): What solutions do you offer in answer to the problems posed in the research objectives?

BOX 14.1: Example of an Abstract of a Dissertation

Abstract

This study determines the Information and Communication Technology (ICT) profile of elementary teachers in Shanghai. Specifically it looks into the thinking process and preparation of teachers towards ICT integration in the teaching–learning process. The study makes use of exploratory descriptive research design and utilises quantitative measures in assessing teachers thinking process and preparation of teachers towards ICT integration. The respondents of the study are elementary teachers who have knowledge about computers. The result of the study shows that only 26 per cent of the respondents have formal training on educational technology and 47 per cent have attended trainings, seminars and workshops related to ICT. However, very few utilise the use of computer in teaching and are mostly limited to the use of word processing, spreadsheet and presentation software. In terms of the thinking process, the study shows that it facilitates teachers' thinking towards ICT integration. Out of the four variables under study—teachers' constructivist belief, teachers' self-efficacy, attitudes towards computer and computer self-efficacy—the study shows that elementary teachers have slight adherence to the constructivist beliefs and shows little confidence on their teaching ability while they hold neutral behaviour towards the use of computer and their computer self-efficacy.

Declaration

Declaration in the certificate page by the scholar is generally done using phrases such as shown in Box 14.2.

BOX 14.2: Certificate of Declaration

Declaration

'I hereby declare that this submission is my own work and that, to the best my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of any university or institute of higher learning, except where due acknowledgement has been made in the text'.

Signature

Dedication (Optional)

You have an option to include a dedication. A dedication is a message from the author prefixed to a work in tribute to a person, group or cause. Most dedications are short statements of tribute beginning with 'To…' such as 'To my family' (Box 14.3).

BOX 14.3: Example of Dedication

Dedication

To my mentor and friend, I could not have done this without you.
Thank you for all your support along the way.

Acknowledgement

Most research reports, dissertations or theses have their subsection to convey appreciation to those who have been involved in the study.

Acknowledgements contain the author's statement of gratitude to and recognition of the people and institutions that helped the author's research and writing.

Acknowledgments should be brief, simple, modest and given only to substantial assistance provided by the guide, head of the department, staff of the department, agencies which provided financial support (scholarship), collaborators and institutions where part of the work has been carried out. Acknowledgments made for routine participation by members of the researcher's family, librarian, friends and clerical helpers are normally considered superfluous. Box 14.4 shows an example of an acknowledgement.

BOX 14.4: Example of an Acknowledgement

Acknowledgement

I would like to express my gratitude to my supervisor Professor Li-peng for the useful comments, remarks and engagement through the learning process of this master's thesis. Furthermore, I would like to thank Dr Li Yan for introducing me to the topic as well for the support on the way. Also, I like to thank the participants in my survey, who have willingly shared their precious time during the process of interviewing. I would like to thank my loved ones, who have supported me throughout entire process, both by keeping me harmonious and helping me putting pieces together. I will be grateful forever for your love.

Signature

Table of Contents

The table of contents lists the chapters and figures contained in the thesis. The contents page should contain the chapter headings, appendices, references and the pages on which they can be found. Separate listing should be given for all the figures, tables and abbreviations. All the preliminary pages should be numbered with lower case roman numbers. Table 14.2 shows an example of table of contents.

You will find this easier to create these if you know how to generate these lists using the indexing and style functions within Microsoft Wword.

TABLE 14.2 Example of Table of Contents

TABLE OF CONTENTS

Source: Authors.

Lists of Tables, Figures and Illustrations

This list should contain the titles of tables, figures or illustrations, together with their page numbers, which are listed in the text. For example, figures in Chapter 3 should be numbered sequentially: Figure 3.1, Figure 3.2 and so on. Similarly, for tables, the numbering system should be according to chapter, for example, tables in Chapter 3 are numbered sequentially: Table 3.1, Table 3.2 and so on.

List of Abbreviations and Symbols

The symbols and abbreviations must be in accordance with international convention. For example:

MOE:	Ministry of Education, China
NCERT:	National Council of Educational Research and Training, India
NGOs:	Non-governmental Organisations
OECD:	Organisation for Economic Co-operation and Development, Paris
UNO:	United Nations Organization, New York
UNESCO:	United Nations Educational, Scientific and Cultural Organization, Paris
WB:	World Bank, Washington DC

List of Appendices

This list is optional. It should contain the titles of appendices placed in the supplementary section.

Appendices may be used to provide relevant supporting evidence for reference but should only be used if necessary. You may wish to include in appendices evidence which confirms the originality of your work or illustrates points of principle set out in the main text, questionnaires and interview guidelines. Only subsidiary material should be included in appendices. You should not assume that examiners would read appendices in detail.

Main Text of the Dissertation

The main text should present the details of the research work and results. This part of the thesis should contain the following about the research work:

- Introduction
- Literature review
- Research methodology
- Findings, results and data analysis
- Discussion
- Summary and conclusion

Introduction

The introduction should 'introduce' the thesis. This is not the summary of your thesis. It is neither a brief version of each chapter of the dissertation. It is an introduction to the topic. In general terms, the introduction of your dissertation should include the following:

The Context in Which the Research Has Been Taken

- What is the background, the context, in which the research took place?
- Why is this research topic or issue important?
- Who participated in the research study and/or 'actors' in the area under investigation?

- Are there important trends or pivotal variables of which the reader needs to be made aware or to be informed?
- Is there a clear and succinct statement of the aims and objectives that your dissertation is going to address?
- Have you described a clear and unambiguous exposition of your research aim, the objectives you will address to meet this aim and your research questions?

The Reasons Why This Study Was Carried Out

- Was this study undertaken, for example, in order to test some aspect of professional or business practice or theory or framework of analysis?
- Was the research carried out to fulfil the demands of an organisation or the education system?

The Way the Dissertation Is To Be Organised

While writing your dissertation, you should keep in mind the idea that the intended reader and reviewer have some shared understanding of the area being investigated. However, key concepts and arguments still need to be included; otherwise, the depth of research will be compromised. This will help you not to be tempted to make too many implicit assumptions. The dissertation should be a document that is comprehensive and more importantly 'self-contained' and does not need any additional explanation, interpretation or reference to other documents in order that it may be fully understood.

This short final section of the Introduction should tell the reader what topics are going to be discussed in each of the chapters and how the chapters are related to each other. By doing so, you are, in effect, providing the reader with a 'road map' of the work ahead. Thus, at a glance, they can see (a) where they are starting from, (b) the context in which the journey is taking place, (c) where they are going to end up and (d) the route which they will take to reach their final destination. Such a 'map' will allow the readers to navigate their way through your work much more easily and appreciate to the maximum what you have done.

In principle, this chapter should not be too lengthy although in some subjects or topics the justification of the subject and scope may change the length of this chapter.

Finally in this part of the report, the researcher should bring clearly the importance of the field and the current status of it. It should contain an overview of the problem, its importance and statement about the hypothesis or specific questions to be explored. This should be followed by a preview of the scheme of the following chapters, that is, an outline of plan of the work. Here, the aim of each of the chapters and their contents should be briefly stated. Related and relevant work done by others must be highlighted. Various concepts and definitions of scientific and technical terms necessary for understanding of the research work are to be defined and explained. The details of statistical methods and tools quantities used in the study can be given in a separate chapter. Irrelevant and less informative materials should not be presented at all in this section of your dissertation.

The introductory chapter should be prepared in such a way that it should interest the reader in the subject matter of research. The chapter should not be aimless, confused and lacking in precision. You may write introductory part, if required, in one or two chapters. Basically, the introductory part should cover the following aspects:

- Features of the topic
- Present status of the field

- Some unsolved problems
- Statement of the problem undertaken
- Justification of the present problem
- Preview of the scheme of the following chapters, their interrelationship and definitions of various scientific terms used
- Methodology used

Literature Review

The purpose of a literature review is to show your reader that you have read and have a good grasp and understanding of the main published work concerning a particular topic or question in your field. The review may be in any format including online sources. You can write it as a separate assignment or as one of the introductory sections of your dissertation or thesis. In the latter cases, in particular, the review should be guided by your research objective or by the issue or thesis you are arguing and should provide the framework for your further work.

In the literature review, you should summarise the main viewpoints and important facts that you noticed in your reading as they relate to your topic of research. You should also use the literature review to justify the value of doing research on your topic by highlighting what is already known, what is not yet known and how it is relevant.

The literature review should not be simply descriptive but should also provide a critical analysis of the body of work. It should demonstrate that you understand how it fits together as a whole and how your own research fits with previous studies.

A key aspect of a literature review is what sources you select to include and which you exclude.

According to Caulley (1992), the literature review should:

- Compare and contrast different authors' views on an issue
- Group authors who draw similar conclusions
- Criticise aspects of methodology
- Note areas in which authors are in disagreement
- Highlight exemplary studies
- Highlight gaps in research
- Show how your study relates to previous studies
- Show how your study relates to the literature in general
- Conclude by summarising what the literature says

The purpose of the literature review is to concisely demonstrate your level of understanding of the research related to your project. You should not discuss all of the literature in depth. Rather you should group your literature according to some general topics and only discuss specific studies if they are 'landmark' studies for your area of research (there should be 6–10 of these). Each of these specific discussions should include specific information about the group involved in the research project, data and results reported. Often a review of literature will include several of these in-depth reviews with 'mini-reviews' of studies that came to the same or similar conclusions. The literature review should end with a discussion of how the literature relates to your study. In short, it should:

- Define and limit the problem you are working on
- Place your study in an historical perspective
- Avoid unnecessary duplication
- Evaluate promising research methods
- Relate your findings to previous knowledge and suggest further research

A good literature review, therefore, is a critical review of what other people have written on your topic or close to your topic. It must identify areas of controversy, raise questions and identify areas which need further research. Table 14.3 provides a checklist of questions which you have to ask yourself to decide whether or not a particular piece of work is worth including in your literature review.

TABLE 14.3 Checklist of Questions for Critical Reading

The Author	• Who is the author? What can you find out about him or her? Has he or she written other books, articles and so on? • What is the author's position in the research process, e.g., gender, class, politics, life experience, relationship to research participants?
The Medium	• Where and when was the document produced? What type of document is it? • Is it reporting original research that the author has done, or is it presenting second-hand information about a topic? • Is it formal or informal? • Is it 'authoritative' (e.g., academic, scientific) or 'popular' (newspaper or magazine) article? • How has it been produced? Is it glossy, with lots of pictures, diagrams and so on? • If it is contained on a website, is the website from a reputable organisation, or is the document drawn from some other reputable source?
The Message	• What is being said? • What is not being said? • How is the argument presented? Why? • What use has been made of diagrams, pictures and so on? • Who was or is the intended audience? • Whose interests are being served by this message? Are there political implications, for instance? • What evidence is presented to support the claims that are made? • Does the evidence actually support the claims? Is the evidence presented in enough detail for you to make-up your own mind whether you agree with the claims? • Are there errors or inconsistencies? • What is the significance to your topic and the research that you wish to carry out?

Source: http://www.skillsyouneed.com/learn/literature-review.html (accessed on 17 November 2017).

The literature review should be presented in the form of a précis, a classification, a comparison and a critical analysis of that material which is germane to a full understanding of your research study. Published material such as textbooks, journal articles, conference papers, reports, case studies, the Internet, magazine features or newspaper articles should be drawn. Among all the published material for the literature review, the most important source of academic literature is articles in professional and academic journals. You should also make sure that you are familiar with the most recent publications in journals relevant to your subject area.

The overall structure of your review will depend basically on your own thesis or research area. What you will be required to do is to group together, and compare and contrast the varying opinions of different writers on certain topics. What you must not do is just describe and narrate what one writer says and then go on to give a general overview of another writer, and then another and so on. The structure of your review should be dictated instead by topic areas, controversial and critical issues or by questions to which there are varying approaches and theories. Within each of these sections, you should then discuss and analyse what the different literature argues and link this to your own purpose.

Linking words are important. If you were grouping together writers with similar opinions, you would use words or phrases such as:

'similarly, in addition, also, again'

More importantly, if there is disagreement, you need to indicate clearly that you are aware of this by the use of linkers such as:

'however, on the other hand, conversely, nevertheless'

At the end of the review, you should include a summary of what the literature implies, which again links to your hypothesis or main question.

You should always keep in mind that your literature review leads and justifies the research objectives and questions of your dissertation. Your literature review should not just be a reference list and/or catalogue of authors, frameworks and ideas but should contain a critical evaluation of those authors' work.

The literature review will be in the range of 3,000 to 4,000 words.

Research Methodology

The research methodology will explain to your reader the methods you used to gather the information and data for your study to answer your research question, almost like a recipe. This chapter should be considered as the core of the dissertation or thesis as an error or weakness in this explanation can invalidate the findings, dooming the study as a whole. This chapter is often scrutinised and reviewed very critically by other experts and weighs heavily in examining the dissertation.

- Start the chapter with a clear explanation of the methodology that will be used to address the research question(s) and solve the problem.
- Provide a detailed description of the several components of the methodology.
- Explain what the methods are all about and how you have used the methodology. You should give the concrete reason why you choose to use this specific methodology for answering your research question.

As mentioned earlier, this is the chapter of your dissertation that explains how you carried out your research, where your data come from, what sort of data gathering techniques you used and so forth. Your description on methodology should contain enough information so as to enable someone reading your methodology to create methods very similar to the ones you used to obtain your data. You should not include any questionnaires, reviews, interviews and so on that you used to conduct your research. The main aim of this section is to explain why you chose to use those particular techniques to gather your data.

The information included in the dissertation methodology is similar to the process of creating a science project: you need to present the subject that you aim to examine and explain the way you chose to go about approaching your research. You have seen that there are several different types of research, and research analysis, including primary and secondary research, and qualitative and quantitative analysis, and in your dissertation methodology, thus, you should explain what types of instruments and techniques or methods you have employed in assembling and analysing your data.

The chapter on research methodology is that part of your dissertation where you attempt to justify to the reader the process by which your research questions, which you derived by an analysis of the relevant literature, were answered. Therefore, it is not appropriate to say, for example, 'suitable respondents were sampled using a quota sampling technique and then surveyed using a postal questionnaire' and then leave it at that. It might well be the case that given the problem(s) to be investigated, such a choice of research methods is entirely appropriate. However, if you have not justified your research choices to a reader they could be correct in assuming that you have, by chance, merely guessed at what would work and, more by luck than judgement, arrived at the 'correct' solution to the problem (Bloomberg and Volpe 2012).

You can structure your methodology in several ways. The following four main headings are designed to give you a better idea of what you may want to include, as well as how you might want to present your findings:

- Research overview: it is where you should reiterate the topic of your research.
- Research design: it is where you have to explain how you have set up your project and what each piece of it aims to accomplish.
- Data collection: what method and tools you used to collect the data (surveys, questionnaires, interviews, trials and so on). Here you should not forget to include sample size and any attempt to defeat bias.
- Data analysis: Finally, what do your data mean in the context of your research? Were your results conclusive or not? Remember to include what type of data you were working with (qualitative or quantitative? primary or secondary sources?) and how any variables, spurious or otherwise, factor into your results.

The conclusion of this chapter should provide a summary of the main points that have been covered. The conclusion should also direct the reader as to how the contents of this chapter link in with the contents of the next chapter your findings.

This chapter will be usually between 1,000 and 2,000 words.

Findings, Results and Data Analysis

While writing the results and discussion sections of your dissertation or thesis, you may feel both these sections as the most interesting as well as the most challenging section to write.

This is a narrative presentation of your findings. This is where you present your statistics, tables, figures and so on that show what the specific findings of your study are. Present them in detail. Remember that someone should be able to duplicate your study based solely on this document. This requires considerable description.

It is very important not to try and combine this chapter with the next one. You need to carefully present your results first with no further interpretation. Once you have presented the data, you are ready to move on to the next section.

These two sections can be written as two separate chapters or you can write only one chapter. This will, however, depend on the guidelines and rules of your university and your own preference.

There are advantages to both approaches.

Writing the results and the discussion as separate sections will allow you to focus first on what the results actually show and help you to sort them in your head.

However, many researchers prefer to combine the results with their implications as the two are closely linked. This you have to check and find out from your university's requirements carefully before combining the results and discussions sections.

In the section on 'results', you should highlight your key experimental results and the statistical analysis proving whether or not the results of these are significant.

The section on results should be drafted in the past tense because you are describing what you have done in the past.

Another important aspect that you have to take note is how meaningfully and effectively present the results of your study: tables, figures, graphs or text. You should attempt to use a variety of different methods of presentation and keep your reader in mind: 20 pages of dense tables are difficult to understand, as are five pages of graphs, but a single table and well-chosen graph that illustrates your overall findings will make things much clearer and readable.

You should make sure that each table and figure has a number and a title. Tables and figures should be numbered in separate lists, but presented consecutively by the order in which you mention them in the text. If you have more than about two or three, it is often helpful to provide lists of tables and figures alongside the table of contents at the start of your dissertation.

Based on the graphs, charts and tables, you should summarise your results and make sure that the text and figures are complementary and not repeating the same information. Every table and graph should be referred to in the text. Figures and tables that you consider relatively less significant or need not to be referred to should be moved to an appendix or even removed.

For example, 'details of all the interview participants can be found in Appendix A, with transcripts of each interview in Appendix B'.

Discussion

This chapter should contain the interpretation of the results. The findings of the research should be compared and contrasted with those of previous studies presented in the literature review. The purpose of this chapter is to discuss the findings of the research.

This chapter should begin with a concise restatement of your study's purpose along with any needed background information. You should restate each of your hypotheses. Now that you have presented the results in the previous section, discuss them in this section. What, specifically, do the results mean? How can they be interpreted? Can they be interpreted in multiple ways? What do the findings tell you about your hypothesis? Do not claim more for your results than the data really show. Avoid speculation.

Conclusions

In this chapter, you bring together the work of the dissertation by showing how the initial research plan has been addressed in such a way that conclusions may be formed from the evidence of the dissertation. You should not put any new material or references in this chapter. The statement on your conclusion should clearly reflect the extent to which each of the aims and objectives of your study has been met.

While writing your conclusion, you should be careful not to make claims that are not substantiated from the evidence you have presented in earlier chapters.

In case you want to include a list of important recommendations with concrete policy implications, then do so in a separate short chapter. Similarly, you may include a short subsection on any suggestions for further research for colleagues who might wish to undertake research in this area in the future. There should also be a short statement of the limitations of the research.

End Matters

The final section of the report should consist of references, appendices, research instruments such as questionnaires, maps or computer programmes and copies of research publications that came out from the research work done.

References or Bibliography

The moment you launch the survey of available source, the preparation and collection of references preferably with annotations should be undertaken. Keeping systematically the records during your research helps you achieve various objectives. It stores and preserves data for future use. You may consider this information of little significance that may not be of immediate use, but it would help you in future. You should keep all details in files.

It is better that your references from current contents or from journals is noted on a separate card or sheet with the names of authors and the title of the paper/book and so on. If the reference is from a research paper, then its title, journal name, volume number, starting and ending pages of it and year of publication should be noted. As regards a book, publisher's name, place of publication and year of publication must be written down. Instead of using cards or paper sheets, you can store this information in your computer and can retrieve it whenever you want. It is better that you classify the reference and group them separately.

All references used in writing the dissertation (whether direct quotations or paraphrasing) should be included in a reference list/bibliography, compiled in alphabetical order by author. The Harvard system for listing references should be used.

Appendices

Appendices should be included as relevant supporting evidence for reference but should only be used if necessary. You may wish to include in appendices, evidence which confirms the originality of your work or illustrate points of principle set out in the main text, questionnaires and interview guidelines. Only subsidiary material should be included in appendices. You should not assume that examiners in detail would read appendices. You should under no circumstances publish someone else's data without their express written consent and proper acknowledgement.

Appendices can be placed at the end of report after references. They should be numbered in capital alphabets.

Writing and Editing the Report

Once you have your outline sorted out and you have got a pile of research notes together, it is time to start writing. You need not necessarily start at the beginning—in fact, introductions are often easier to write at the end when you know how your argument has developed (Klein 2013).

Get going on the bits you know you will find easy, then use your outline to put them together in the right order. You will find areas that need further research, so be prepared to revisit the library as you are going along.

Your style of writing should communicate your ideas effectively. A well-planned and researched dissertation can be let down by poorly expressed ideas or unclear phrasing. You should prepare two or three drafts, refining your work each time, before you are happy with the end result.

Joe Wolfe (2016) offers the following observations in regard to the editorial process:

> As you write your thesis, your scientific writing is almost certain to improve. Even for native speakers of English who write very well in other styles, one notices an enormous improvement in the first drafts from the first to the last chapter written. The process of writing the thesis is like a course in scientific writing, and in that sense each chapter is like an assignment in which you are taught, but not assessed. Remember, only the final draft is assessed: the more comments your adviser adds to first or second draft, the better.

Style of the Dissertation Report

During your research, you will have consulted and read a number of scholarly articles. It is therefore necessary to select a recommended academic text that you find easy and enjoyable to read. For this, the first task you have to do is to study the structures and work out how arguments are presented. It will be equally interesting to collect good examples of vocabulary and punctuation.

For facilitating your task, you may also consider how techniques are used in similar dissertations for convincing the reader of their argument and see if you can apply them in your own writing.

Dissertation reports are quite lengthy. Thus, subheadings are a useful way of breaking up the text and signalling to the reader what stage you have reached. Pull these subheadings as you move through each draft to ensure they still provide a useful overview of the section. You should avoid repetition.

British and American Format

The general recommendation for writing theses in English in most universities is to use British English. There are scholars and research organisations that publish their work in American journals and are more used to communicating in American English. Regarding the use of British and American English, you should check this with your supervisor before you start wring your dissertation. In principle, you should use British English unless you have reason for writing your thesis in American English.

Structuring Sentences

Your dissertation is an academic work. You are not expected to produce something completely original, but instead, you should show understanding of key issues and theories; evidence of thought and insight;

critical analysis and evaluation; and a demonstration that your findings are appropriate and have some policy implications. Simple description will not suffice and will result in a low grade for your dissertation.

As you write your dissertation, it is worth distinguishing the key points in your discussion from less important supporting ideas. Your aim should be to give full weight to your key points by giving them each a sentence of their own. You can add elaborations and detail in subsequent sentences.

A common mistake that researcher generally makes in writing the dissertation is the use of long and complex sentences. The general belief is that the longer the sentence, the cleverer it sounds. You should understand that every word conveys a unit of meaning on its own, however small, so the more words there are in a sentence, the harder it will be for the reader to understand and grasp its underlying meaning.

Try to avoid too much use of clauses. Start the next point with a new sentence. Connective words and phrases such as however, consequently, but, so and so on can be placed at the start of the new sentence, if required, to indicate its relationship to the previous one and make your workflow..

Your writing style should be an appropriate academic style. Thus, you should attempt to avoid colloquialisms, contractions, phrasal verbs and vagueness. You do not need, however, to use long, over-formal vocabulary: you should aim at all times for clear and concise expression.

Use of personal language ('I', 'my' and so on) should be normally avoided, although opinions on this vary among experts. However, you can only use it when you want to describe what you actually did and when you want to express your personal opinions, probably in your conclusion. Do not refer to yourself as 'we' unless you are describing some sort of group work and do not refer to yourself as 'the author'.

Acknowledging Ideas

Acknowledgement of your sources should be clearly mentioned and highlighted in your dissertation. If you do not do this, particularly at dissertation/postgraduate level, you could be blamed of plagiarism.

Indeed your dissertation should contain your own original though. However, you would want to refer to the views and ideas of other writers as well on the topic.

It is vital that you make sure that your dissertation critically evaluates those ideas and identifies what problems remain in your area of research and what has not yet been explored.

You can also refer to the work of others as evidence to back up your own argument. When doing this, ensure you add a footnote to signpost clearly to the reader the original source of the point you are making.

Referencing

For referencing purpose, you can select one of the two methods—Oxford method or Harvard method. You have to check this with your course tutor or in the course literature of your university. The idea of referencing can be a complicated area, but there are many guides and staff to help you.

If your dissertation has little or no references and a short bibliography, it means you have not carried out a thorough research on your topic of study. It would generally reflect an un-academic research approach and maybe even copying from source material. You should ensure that you have consulted a sufficient number of references to books, articles and sources. Some should be primary sources and you will also have to refer to and quote secondary sources, which are usually academic articles that analyse primary sources.

TABLE 14.4 Example of Listing the References in Dissertation

References (Harvard Style)

- Author, P.S. (2006). Title of a book. London: Publishing House.
- Contributor, A.M. (2006). Title of a chapter in a book. In P.S. Author (Ed.), *Title of book*. London: Publishing House. pp. 101–134.
- Editorial. (2008) Future without technology. *Sunday Times*, April 21, p. 23.
- Davidson, J. (2008). GM foods: The answer to shortages. *Sunday Times*, April 21, p. 31.
- Games people play. www.gamespeopleplay.co.za (Accessed 16/07/2011).
- Masters, P.J. (2007). *Title of dissertation.* Unpublished master's dissertation. Johannesburg: Da Vinci Institute.
- Davis, C. (2012). Interview with the author. Johannesburg, 10 March.
- Republic of South Africa. (1997). *Constitution of the Republic of South Africa,* 1996 (Act 35 of 1997). Cape Town: Government Printer.
- Researcher, L.M. (2005). Title of the article. *Journal of Whatever*, 20(3) pp. 102–108.
- Researcher, M.M., Assistant, P.L. & Professor, A.B. (2006). Title of the article. American Journal of Whatever, 45(3) pp. 91–96.
- Samuels, P.R. (2007). *Personal communication*. Johannesburg, 10 February 2007.
- Scholar, P.C. (2005). *Title of conference paper*. Paper presented at the 12th Annual Conference of the International XYZ Association, London, UK, 12–15 September.
- Scientist, V.C. (2012). *Title of the contribution*. [Full website address] (Date accessed).
- Task Group Y. (2007). *Report on whatever was investigated. Johannesburg*: Eskom Holdings.

Source: Authors.

Extensive referencing and bibliography indicate wide research, a correct approach and the use of these sources as evidence to back up your argument. Table 14.4 gives you an idea as to how you can prepare and format your list of references or bibliography.

The list of references should include all direct and indirect sources of information that you have used in your dissertation or thesis. You should not include other sources that may perhaps be relevant, but not actually used in the dissertation.

- You should list references in alphabetical order across categories of sources; in the case of more than one publication by the same author, the references should be listed in chronological order.
- You should not differentiate between different sub-categories of sources (e.g., journals and books).
- Your entries must be accurate.
- Author or title of website should be referenced and the link of website should be added behind with the date of your consultation of that website.

Summary

This chapter describes two important aspects of a research study: data presentation and report writing.

The section on data presentation describes the critical role that tables, figures and graphs (or display items) play in enhancing the quality of your research. It showed how scientific tables and graphs could be utilised to represent sizeable numerical or statistical data in a time- and space-effective manner. The section suggested guidelines that researchers can use to present data, clarify interpretations and explain concepts. This section covered when the researcher should use figures and tables, and how to format them such that they serve their purpose.

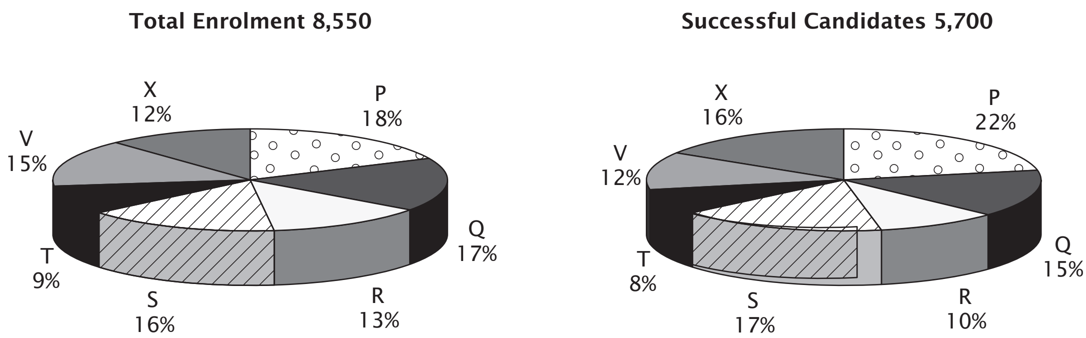
The second part on report writing deals with: How to write a thesis? This chapter provided simple and practical advice on the problems of getting started, getting organised, dividing the huge task into less formidable pieces and working on those pieces. It also explained the practicalities of surviving the ordeal. It included a suggested structure and guidelines to what should go in each section.

In particular, this chapter suggested that before starting the dissertation or thesis, the researcher should start by setting out each chapter, section and subsections. The outline of the report should clearly reflect the logical details of the layout, formatting and the researcher's writing style. The logical structuring of chapters, sections and subsections helps to introduce the reader systematically to the necessary background and makes him or her receptive to the new ideas and conclusions which he or she will be exposed to. The section suggests that each chapter in your dissertation must have a central idea which is introduced, argued and concluded. As you develop a deeper understanding of the area that you are working in, you may realise that there are better ways of organising your report.

Self-test Exercises

Exercise 14.1: The following graph shows the distribution of candidates who were enrolled for MA (economics) entrance exam and the candidates (out of those enrolled) who passed the exam in different colleges of a university.

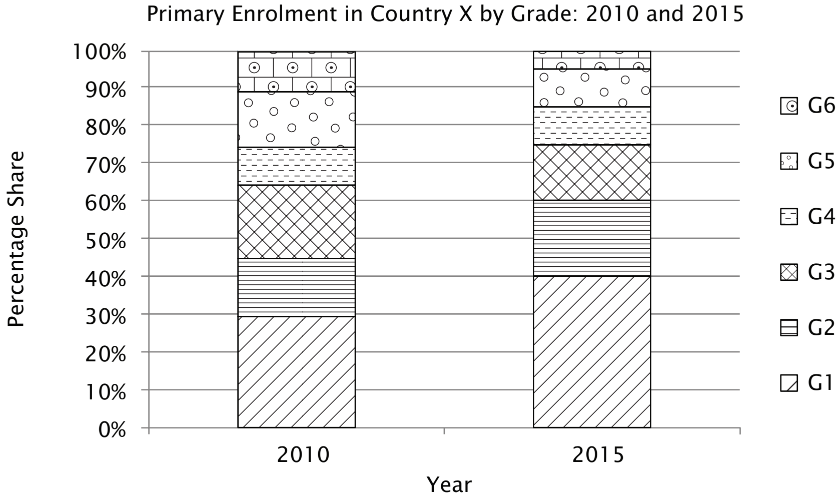
Study the graph carefully and answer the following question:



14.1.1. What percentage of candidates passes the exam from College T out of the total number of candidates enrolled from the same college?

14.1.2. Which College has the highest percentage of candidates passed to the candidates enrolled?

Exercise 14.2: The bar graph given further shows the percentage distribution of enrolment by grade in primary level of education in Country X in 2010 and 2015.



14.2.1. What was the difference in the number of Grade 2 pupils in 2010 and those enrolled in 2015?

14.2.2. What were the total number of pupils enrolled in Grade 1, Grade 2 and Grade 5?

Exercise 14.3:

14.3.1. What is a dissertation proposal and what is it good for?

14.3.2. Which key steps you will take into account before you start writing your dissertation? List these steps in chronological order.

14.3.3. Write on one page an abstract form a hypothetical research.

References

Bloomberg, L.D. and M. Volpe. 2012. *Completing Your Qualitative Dissertation: A Road Map from Beginning to End*. Thousand Oaks, CA: SAGE Publications.

Caulley, D.N. 1992. *Writing a Critical Review of the Literature*. Bundoora: La Trobe University.

Cleveland, S.W. 1994. *The Elements of Graphing Data*. Summit, NJ: Hobart Press.

Hunter, H. *Understand charts and graphs used in business*. Available at http://slideplayer.com/slide/6882344/ (accessed on 12 Feb 2018).

Klein, S. 2013. 'How to Write Your Dissertation'. *The Guardian*. Available at https://www.theguardian.com/education/2013/mar/25/how-to-write-your-dissertation (accessed on 12 February 2018).

Strunk Jr., William and E.B. White. 1999. *The elements of style*, 4th ed. London: Longman Publishing.

Tufte, E.R. 1983. *The Visual Display of Quantitative Information*. Graphics Press. Available at https://www.edwardtufte.com/tufte/books_vdqi (accessed on 17 November 2017).

Turabian, Kate L. 2007. *A Manual for Writers of Research Papers, Theses, and Dissertations*, 7th ed. Chicago, IL: University of Chicago Press.

United Nations Economic Commission for Europe. 2009. Making Data Meaningful, ECE/CES/STAT/NONE/2009/3. DAG Repository.

Wolfe, J. 2016. *How to Write a PhD Thesis*. Sydney: School of Physics, University of New south Wales.

Index