M. A. PSYCHOLOGY

SEMESTER - I (CBCS)

PSYCHOLOGY PAPER- COURSE II

(CORE COURSE)

RESEARCH METHODOLOGY FOR PSYCHOLOGY

SUBJECT CODE : PAPSY102

Prof. Suhas Pednekar
Vice Chancellor
University of Mumbai, Mumbai.

Prof. Ravindra D. Kulkarni
Pro Vice-Chancellor,
University of Mumbai.

Prof. Prakash Mahanwar
Director
IDOL, University of Mumbai.

Programe Co-ordinator : Prof. Anil R. Bankar,
Head, Faculty of Humanities and Social Sciences,
IDOL, University of Mumbai

Course Co-ordinator : **Dr. Naresh Tambe**
Assistant Professor (Psychology), IDOL,
University of Mumbai.

Editor: Compiled SLM by : **Dr. Naresh Tambe**
Assistant Professor (Psychology), IDOL,
University of Mumbai.

May 2022, Print I

Published by
Director
Institute of Distance and Open Learning,
University of Mumbai,
Vidyanagari, Mumbai - 400 098.

DTP COMPOSED AND PRINTED BY
Mumbai University Press
Vidyanagari, Santacruz (E), Mumbai - 400098.

CONTENTS

*****

Syllabus
PSYCHOLOGY
MA Semester System (CBCS), Revised Course, 2022.23
Semester I: Course II
Core Course: 4 credits, 60 hrs.
RESEARCH METHODOLOGY FOR PSYCHOLOGY:PAPSY102

Objectives:

1. To acquaint learners with methodology of quantitative and qualitative psychological research.
2. To acquaint learners with philosophy, ethics, design, and evaluation of research in psychology.
3. To enable learners to design research.

Unit 1.Philosophy and ethics of psychological research.

a. Epistemological positions in psychological research: scientific realism, logical positivism;Ockham's razor;
b. Popper and Kuhn's contribution: theory dependence of observation; understanding theory: components and connections – concepts, constructs, variables and hypothesis; Duhem–Quine thesis; Quine's critique of empiricism
c. Ethical standards of psychological research: planning, conduction and reporting research
d. Proposing and reporting quantitative research

Unit 2. Research settings and Mehtods of Data collection

a. Observation and Interview method
b. Questionnaire
c. Survey research
d. Other non-experimental methods

Unit 3. Experimental and Quasi-Experimental Methods

a. Independent groups designs
b. Repeated measures designs
c. Complex designs
d. Quasi-experimental designs and program evaluation

Unit 4. Qualitative Research

a. Philosophy and conceptual foundations; proposing and reporting qualitative research
b. Grounded theory
c. Interpretive phenomenological analysis; discourse analysis
d. Narrative analysis; conversation analysis

Books for study

1. Shaughnessy, J. J., Zechmeister, E. B. &Zechmeister, J. (2012). *Research methods in psychology*. (9th ed..). NY: McGraw Hill.
2. Elmes, D. G. (2011). *Research Methods in Psychology* (9thed.). Wadsworth Publishing.
3. Goodwin, J. (2009). *Research in Psychology: Methods in Design* (6thed.). Wiley.

4. McBurney, D. H. (2009). *Research methods*. (8th Ed.). Wadsworth Publishing.

5. Forrester, M. A. (2010). *Doing Qualitative Research in Psychology: A Practical Guide*. Sage.

Books for reference:

1. Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis (Introducing Qualitative Methods series)*. New Delhi: Sage Publications.
2. Dominowski, R. L. (1980). *Research methods*. N.J.: Engelwood Cliffs, Prentice-Hall.
3. Embreston, S. E., & Raise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum
4. Hambleton, R. K., &Swaminathan H. (1985). *Item Response theory: Principles and Applications*. Boston: Kluwer
5. Hoyle, R. (1995). *Structural equation modeling: concepts, issues and applications*. Thousand Oaks, CA: Sage.
6. Hulin, C. L., Drasgow, F. & Parsons, C.K. (1983). *Item response theory: application to psychological measurement*. Homewood,Il: Dow Jones-Irwin.
7. Kerlinger, F. N. (1995). *Foundations of behavioural research*. New Delhi: Surjeet Publication.
8. Lewis-Beck, M. S. (1994). *International handbook of quantitative applications of social sciences*. Sage: Topan/London.
9. McBurney, D. H. (2001). *How to Think Like a Psychologist: Critical Thinking in Psychology* (2nd Edition). Prentice Hall.
10. Monette, D. R., Sullivan, T. J., &DeJong, C. R. (1994). *Applied psychological research: Tools for human services*. (3rded.). California: Harcourt Brace College Publisher.
11. Morse, J. M. (1994). *Critical issues in qualitative research methods*. Sage Publications.
12. Nunnally, J. & Bernstein, I. (1994). *Psychometric Theory* (3rded.). New York: McGraw Hill.
13. Robinson, P. W. (1976). *Fundamentals of experimental designs: A comparative approach*. Engelwood-Cliff: Prentice Hall.
14. Smith, J. A. (2008). *Qualitative Psychology: A Practical Guide to Research Methods*. Sage.
15. Strauss A L and Glaser, B. G. (1967).*The Discovery of Grounded Theory: Strategies for Qualitative Research*. Aldine Transaction.
16. Wiling, C. (2008). *Introducing Qualitative Research in Psychology* (2nded). Open University Press.
17. American Psychological Association. (2009). *Publication Manual of the American Psychological Association* (6thed.). APA.
18. American Psychological Association. (2009). *Concise Rules of APA Style (Concise Rules of the American Psychological Association (APA) Style)*. APA.

Evaluation:

Internal evaluation: 40 marks

Semester end examination: 60 marks

Paper pattern: 7 questions to be set of 15 marks each, out of which 4 are to be attempted. One of them could be short notes question, which could combine more than one unit.

*****

1

PHILOSOPHY AND ETHICS OF PSYCHOLOGICAL RESEARCH

Unit Structure

1.1 INTRODUCTION

Psychologists develop theories and conduct psychological research to answer questions about behavior and mental processes; these answers can impact individuals and society. The scientific method, a means to gain knowledge, refers to the ways in which questions are asked and the logic and methods used to gain answers. Two important characteristics of the scientific method are an empirical approach and a skeptical attitude.

1.2 EPISTEMOLOGICAL POSITIONS IN PSYCHOLOGICAL RESEARCH: SCIENTIFIC REALISM, LOGICAL POSITIVISM; OCKHAM'S RAZOR

Apart from any philosophical interest that we may have in science because of its status and influence on our lives, science is important to philosophy because it seems to offer answers to fundamental philosophical questions.

One such question is ‘how can we have knowledge as opposed to mere belief or opinion?’, and one very general answer to it is ‘follow the scientific method’. So, for example, whatever any of us may believe, rightly or wrongly, about whether smoking causes cancer or traffic fumes cause asthma, a government will not act unless there is scientific evidence supporting such beliefs (of course, they may still not act even when there is evidence).

Similarly, in all the examples mentioned above, respect is accorded to the views of scientists because their conclusions are supposed to have been reached on the basis of proper methods of gathering and assessing evidence, and hence are supposed to be justified.

1.2.1 Epistemology:

The branch of philosophy that inquires into knowledge and justification is called epistemology. The central questions of epistemology include: what is knowledge as opposed to mere belief?; can we be sure that we have any knowledge?; what things do we in fact know?. The first of these is perhaps the most fundamental epistemological question.

Each of us has many beliefs, some true and some false. If I believe something that is, as a matter of fact, false (suppose, for example, that I believe that the capital city of Australia is Sydney) then I cannot be said to know it. In logical terminology we say a necessary condition, that is a condition that must be satisfied, for somebody knowing some proposition is that the proposition is true.

In other words, if somebody knows some proposition then that proposition is true. (The converse obviously does not hold; there are lots of propositions that are true but which nobody knows, for example, there is a true proposition about how many leaves there are on the tree outside my window, but I presume nobody has bothered to find out what it is.) Where someone believes something that turns out to be false (no matter how plausible it seemed) then we would say that they thought they knew it but that in fact they did not.

Suppose too that another necessary condition for somebody knowing some proposition is that he or she believes that proposition. We now have two necessary conditions for knowledge; knowledge is at least true belief, but is that enough? Consider the following example: suppose that I am very prone to wishful thinking and every week I believe that my numbers will come up on the lottery, and suppose that one particular week my numbers do in fact come up; then I had a belief, that my numbers would come up, and it was a true belief, but it was not knowledge because I had no adequate reason to believe that my numbers would come up on that particular week rather than on all the other weeks when I believed they would come up, but when they did not. Hence, it may be the case that I believe something, and that it is true, but that I do not know it. So it seems that for something someone believes to count as knowledge, as well as that belief being true, something else is required. My belief about the lottery in the example above did not count as knowledge because I lacked an adequate reason to believe that I would win that week; we would say that my belief was not justified. The traditional view in epistemology has been that knowledge can only be claimed when we have an adequate justification for our beliefs, in other words, knowledge is *justified* true belief. Although recently this 'tripartite' definition of knowledge has been the subject of much criticism and debate, justification is still often regarded as necessary for knowledge. This brings us to the issue of what justification amounts to and, as suggested above, justification is often thought to be provided by following scientific methods for testing or arriving at our beliefs (the word science comes from the Latin word *scientia*, which means knowledge).

1.2.2 Scientific realism:

Many of the entities postulated by modern science, such as genes, viruses, atoms, black holes, and most forms of electromagnetic radiation, are unobservable (at least with the unaided senses). So, whatever the scientific method is and however scientific knowledge is justified, we can ask whether we ought to believe what science tells us about reality beyond the appearances of things. Roughly speaking, scientific realism is the view that we should believe in the unobservable objects postulated by our best scientific theories. Of course, many of those who defend scientific realism also defend the rationality of scientific theory change against sceptics and relativists.

However, some ancient and modern critics of scientific realism have not questioned the success or even the progress of scientific inquiry. Many antirealists about scientific knowledge in the history of philosophy are happy to agree with realists that science is the paradigm of rational inquiry, and that it has produced a cumulative growth of empirical knowledge. However, antirealists of various kinds place limits on the extent and nature of scientific knowledge Hence, the issue of scientific realism is more subtle than many of the polarised debates of science wars, and it is important not to confuse the former with questions about the rationality of science.

The disputes about scientific realism are closely related to those about other kinds of realism in philosophy, some of which will be explained in this chapter, but the reader – especially one with a good deal of scientific knowledge – may already be feeling impatient. Isn't it just obvious that plenty of unobservables described by scientific theories exist; after all, don't scientists manipulate things like atoms and invisible radiation when they design microchips and mobile phone networks? In fact, is it really correct to describe atoms as unobservable? After all, don't we now see photographs of crystal lattices made with microscopes that use electrons instead of light to generate images? Is there really any room for reasonable doubt that atoms exist when so many different parts of science describe how they behave and give rise to everything from the characteristic glow of the gas in a neon light on a billboard, to the way that haemoglobin in red blood cells absorbs oxygen in our lungs? Even if we decide that atoms are now observable, the issue of principle returns when we ask about the existence of the entities that supposedly make up atoms, and so on. Furthermore, scientists of the past claimed to be manipulating and observing theoretical entities that no longer feature in our best scientific theories, so why should we have such faith that we have it right this time? These and other arguments for and against scientific realism will be the subject of the chapters that follow. First, in this chapter, I will explain the background of the contemporary debate, and the different components of scientific realism. We begin with the distinction between appearance and reality.

1.2.3 Logical positivism:

The term 'positivism' was coined by a French philosopher called Auguste Comte (1798–1857) who argued that societies pass through three stages – namely the theological, the metaphysical and the scientific.In the theological stage, people explain phenomena such as thunder, drought and disease by invoking the actions of gods, spirits and magic. In the metaphysical stage, they resort to unobservable forces, particles and so on. The scientific stage is achieved when pretensions to explain why things happen, or to know the nature of things in themselves, are renounced; the proper goal of science is simply the prediction of phenomena. He aimed to complete the transition of European thought to the scientific stage by advancing the scientific study of society and social relations (sociology), and established a system of rituals celebrating scientists and science, to replace the traditional calendar of Saint's Days and religious festivals.

Positivism has its roots in empiricism, especially in Hume's attempt to separate the meaningful from meaningless

In general, positivists:

(a) emphasise verification/falsification;

(b) regard observation/experience as the only source of knowledge (empiricism);

(c) are anti-causation;

(d) are anti-theoretical entities;

(e) downplay explanation;

(f) are, in general, anti-metaphysics.

Logical positivism was originally centred around a group of scientists, mathematicians and philosophers called the Vienna Circle, which met in the 1920s. Many of the Vienna Circle were Jewish and/or socialists. The rise of fascism in Nazi Germany led to their dispersal to America and elsewhere, where the ideas and personalities of logical positivism had a great influence on the development of both science and philosophy.

The difference between logical positivism and logical empiricism is a matter of scholarly dispute. The most influential of those classified as logical positivists or empiricists include Moritz Schlick (1882–1936), Carl Hempel (1905–1997), Carnap, Reichenbach (although he was in Berlin, not Vienna), and Ayer (he visited the Circle and brought some of its ideas to Britain). They all adopted the empiricism of Hume and Mach and Comte's aspiration for a fully scientific intellectual culture. What was new about them was that they exploited the mathematical logic, recently developed by Gottlob Frege (1848–1925) and Russell among others, to provide a framework within which theories could be precisely formulated. The idea was that if the connections between ideas and associated experiences could be made precise, then it would be possible to separate meaningless metaphysical mumbo-jumbo from empirical science. is a more fundamental principle of simplicity that is often claimed to be essential to science, namely Occam's razor, which is roughly the prescription not to invoke more entities in order to explain something than is absolutely necessary. (This kind of simplicity is called ontological parsimony.) According to Ockham's razor, whenever we have two competing hypotheses, then if all other considerations are equal, the simpler of the two is to be preferred. Hume's empiricism means that he thinks that, because the two hypotheses entail exactly the same thing with respect to what we are able to observe, then all other considerations that are worth worrying about are indeed equal.

1.2.4 Ockham's razor:

The principle was, in fact, invoked before Ockham by Durandus of Saint-Pourçain, a French Dominican theologian and philosopher of dubious orthodoxy, who used it to explain that abstraction is the apprehension of some real entity, such as an Aristotelian cognitive species, an active intellect, or a disposition, all of which he spurned as unnecessary. Likewise, in science, Nicole d'Oresme, a 14th-century French physicist, invoked the law of economy, as did Galileo later, in defending the simplest hypothesis of the heavens. Other later scientists stated similar simplifying laws and principles.

Ockham, however, mentioned the principle so frequently and employed it so sharply that it was called "Occam's razor" (also spelled Ockham's razor). He used it, for instance, to dispense with relations, which he held to be nothing distinct from their foundation in things; with efficient causality, which he tended to view merely as regular succession; with motion, which is merely the reappearance of a thing in a different place; with psychological powers distinct for each mode of sense; and with the presence of ideas in the mind of the Creator, which are merely the creatures themselves.

In science, Occam's razor is used as a heuristic to guide scientists in developing theoretical models rather than as an arbiter between published models. In physics, parsimony was an important heuristic in Albert Einstein's formulation of special relativity,[45][46] in the development and application of the principle of least action by Pierre Louis

Maupertuis and Leonhard Euler, and in the development of quantum mechanics by Max Planck, Werner Heisenberg and Louis de Broglie.

When scientists use the idea of parsimony, it has meaning only in a very specific context of inquiry. Several background assumptions are required for parsimony to connect with plausibility in a particular research problem. The reasonableness of parsimony in one research context may have nothing to do with its reasonableness in another. It is a mistake to think that there is a single global principle that spans diverse subject matter.

1.3 POPPER AND KUHN'S CONTRIBUTION: THEORY DEPENDENCE OF OBSERVATION; UNDERSTANDING THEORY: COMPONENTS AND CONNECTIONS – CONCEPTS, CONSTRUCTS, VARIABLES AND HYPOTHESIS; DUHEM–QUINE THESIS; QUINE'S CRITIQUE OF EMPIRICISM

Karl Popper had a considerable influence on philosophy of science during the twentieth century and many scientists took up his ideas. As a result, he was made a member of the Royal Society of London, which is one of the most prestigious scientific associations. In fact, Popper's falsificationism is probably now more popular among scientists than it is among philosophers. Popper also played an important role in the intellectual critique of Marxism, and his books *The Poverty of Historicism* and *The Open Society and Its Enemies* are still widely read by political theorists today. His interest in philosophy of science began with the search for a demarcation between science and pseudo-science. He tried to work out what the difference was between theories he greatly admired in physics, and theories he thought were unscientific in psychology and sociology, and soon came to the conclusion that part of the reason why people erroneously thought that mere pseudo-sciences were scientific was that they had a mistaken view about what made physics scientific.

Popper's solution to the problem of induction is simply to argue that it does not show that scientific knowledge is not justified, because science does not depend on induction at all. Popper pointed out that there is a logical asymmetry between confirmation and falsification of a universal generalisation. The problem of induction arises because no matter how many positive instances of a generalisation are observed it is still possible that the next instance will falsify it. However, if we take a generalisation such as all swans are white, then we need only observe one swan that is not white to falsify this hypothesis.

Popper argued that science is fundamentally about falsifying rather than confirming theories, and so he thought that science could proceed without induction because the inference from a falsifying instance to the falsity of a theory is purely deductive. (Hence, his theory of scientific method is called *falsificationism*.)

Kuhn was a physicist who became interested in the history of science and especially the Copernican revolution. The standard view that he found presented in textbooks and in historical and philosophical works, was that the Copernican revolution, and especially the argument between Galileo and the Catholic Church, was a battle between reason and experiment on the one hand, and superstition and religious dogma on the other. Many historians and scientists suggested that Galileo and others had found experimental data that were simply inconsistent with the Aristotelian view of the cosmos. Kuhn realised that the situation was considerably more complex, and he argued that the history of this and other revolutions in science was incompatible with the usual inductivist and falsificationist accounts of the scientific method. Kuhn's book *The Structure of Scientific Revolutions (1962)* offered a radically different way of thinking about scientific methodology and knowledge, and changed the practice of history of science. His philosophy of science has influenced academia from literary theory to management science, and he seems single-handedly to have caused the widespread use of the word 'paradigm'.

According to Kuhn, the evaluation of theories depends on local historical circumstances, and his analysis of the relationship between theory and observation suggests that theories infect data to such an extent that no way of gathering of observations can ever be theoryneutral and objective. Hence, the degree of confirmation an experiment gives to a hypothesis is not objective, and there is no single logic of theory testing that can be used to determine which theory is most justified by the evidence. He thinks, instead, that scientists' values help determine, not just how individual scientists develop new theories, but also which theories the scientific community as a whole regards as justified

1.3.1 Theory dependence of observation:

The idea that observation is theory dependent is central to many debates in the philosophy of science. This concept is a reaction to the idea that disputes about the way the world is can be easily and simply resolved by simply 'looking at the facts', or performing some sort of experiment or observation. The problem with this answer is that there is no 'neutral' vantage point from which such facts can be gathered or interpreted. Rather, empirical evidence is always interpreted within the context of one's preexisting ideas, conceptions, and expectations, which can often have a dramatic effect on how observations are understood or what they are taken to mean.

Historians of science have given many examples of instances where proponents of rival theories have interpreted the same empirical evidence in very different ways, in accordance with their theoretical commitments. An interesting illustrative case can be found in a popular drawing called the 'duck-rabbit', a sketch which can be interpreted as either a drawing of a duck or of a rabbit, depending on the 'theory' one applies in interpreting the pattern of lines. While in this particular case both interpretations are equally 'correct', in many cases scientific and philosophical disputes, however, it is often unclear whether one, both, or none of the differing interpretations of the relevant facts are correct.

Another problem with observation is that there are always far too many empirical facts for us to consider all of them. One must have some way of selecting which facts are 'relevant' and which are not, an activity which naturally requires the use of some theory, a theory which in turn may be widely disputed. For example, scientists do not spent their time counting the number of blades of grass on every lawn, even though this would result in the collection of more facts, because we have reason to think that this fact has no real importance or significance. If, however, we believed that the number of blades of grass on a field was a form of communication from an extraterrestrial race, or a sign from some divine being, then our attitude towards the significance and meaning of the very same facts would doubtless be very different. This dispute could not simply be resolved by 'looking at the facts', because which facts we regard as relevant would depend upon which theory we accepted.

How to resolve such problems has been the subject of considerable philosophical attention, and remains an ongoing problem for any attempt to provide a comprehensive philosophical underpinning for scientific inquiry.

1.3.2 Understanding theory: components and connections – concepts, constructs, variables and hypothesis:

A theory is a method we use to give us understanding. One of the major purposes of a theory is to provide an answer to the question '*why?*'. Asking, '*why?*', to increase your knowledge of a subject area and realign your thoughts and opinions is an essential skill for anybody who wants to learn and develop.

'*Why*' is one of the very first questions that children ask:

- *"Can you get ready for bed now?" … "Oh* ***why****?"*
- *"****Why*** *is snow cold?"*
- *"****Why*** *do I have to go to school tomorrow?"*
- *"****Why*** *is the sky blue?"*

Questions like these, from children, can be endless. Often finding or providing suitable explanations can be exhausting and frustrating – perhaps we resort to saying, "*Well it just is!*" At the basis of such questions however, are a child's first attempts to understand the world around them, and develop their own theories of why things are the way they are.

Defining 'theory', therefore, has to take into account the 'why?' question, but a theory is deeper than that. The points below go some way to helping with a definition.

- A theory is an attempt to explain why and so to provide understanding.
- A theory is not just 'any' explanation - a theory comes into being when a series of ideas come to be held and accepted by a wider community of people.
- A theory is not necessarily factually based – how we understand and provide explanations arises from our cultural background and how we view the world.

Components: One lesson is that the reason a "good" theory should be testable, be coherent, be economical, be generalizable, and explain known findings is that all of these characteristics serve the primary function of a theory–to be generative of new ideas and new discoveries.

The components of theory are **concepts** (ideally well defined) and **principles**.

A **concept** is a symbolic representation of an actual thing - tree, chair, table, computer, distance, etc. Concept is a world that expresses an abstraction formed by generalizations from particulars e.g., weight, achievement

Construct is the word for concepts with no physical referent - democracy, learning, freedom, etc. Language enables conceptualization. Construct has the added meaning of having been deliberately and consciously invented or adopted for a special scientific purpose

A **principle** expresses the relationship between two or more concepts or constructs.

In the process of theory development, one derives principles based on oneÕs examining/questioning how things/concepts are related.

Concepts and principles serve two important **functions**:

1) They help us to understand or explain what is going on around us.

2) They help us predict future events (Can be causal or correlational)

A **Problem** is an interrogative sentence or statement about the relationship between tow or more variables. Do teachers' comments cause improvement in student performance?

A research problem is a specific issue, difficulty, contradiction, or gap in knowledge that you will aim to address in your research. You might look for practical problems aimed at contributing to change, or theoretical problems aimed at expanding knowledge.

Bear in mind that some research will do both of these things, but usually the research problem focuses on one or the other. The type of research problem you choose depends on your broad topic of interest and the type of research you want to do.

Why is the research problem important? Your topic is interesting and you have lots to say about it, but this isn't a strong enough basis for academic research. Without a well-defined research problem, you are likely to end up with an unfocused and unmanageable project.

You might end up repeating what other people have already said, trying to say too much, or doing research without a clear purpose and justification. You need a problem in order to do research that contributes new and relevant insights.

Whether you're planning your thesis, starting a research paper or writing a research proposal, the research problem is the first step towards knowing exactly what you'll do and why.

In research, **variables** are any characteristics that can take on different values, such as height, age, temperature, or test scores.

Researchers often manipulate or measure independent and dependent variables in studies to test cause-and-effect relationships.

The **independent** variable is the cause. Its value is independent of other variables in your study.

The **dependent** variable is the effect. Its value depends on changes in the independent variable.

Example:

You design a study to test whether changes in room temperature have an effect on math test scores.

Your independent variable is the temperature of the room. You vary the room temperature by making it cooler for half the participants, and warmer for the other half.

Your dependent variable is math test scores. You measure the math skills of all participants using a standardized test and check whether they differ based on room temperature.

A **Hypothesis** is a conjectural or declarative sentence or statement of the relation between two or more variables. Teachers' reinforcement would have significant impact on students performance.

As a researcher, we never know the outcome prior to the research work but we will have certain assumptions on how the end results will be. Based on our hunch and curiosity, we will test it by collecting information that will enable us to conclude whether our assumptions are right. hypothesis has several functions:

(a) Enhance the objectivity and purpose of a research work;

(b) Provide a research with focus and tells a researcher the specific scope of a research problem to investigate;

(c) Help a researcher in prioritising data collection, hence providing focus on the study; and

(d) Enable the formulation of theory for a researcher to specifically conclude what is true and what is not.

Generally, there is only one type of hypothesis, that is, research hypothesis. Research hypothesis forms the basis of investigation for a researcher. However,recent conventions in the scientific field and inquiries stated that hypothesis can be classified into two main categories ÀÛÝ research hypothesis and alternate hypothesis. Alternate hypothesis is a convention among the scientific community. The main function of an alternate hypothesis is to explicitly specify the relationship that will be considered true in case the research hypothesis proves to be wrong. We can see that in a way, alternate hypothesis is the opposite of research hypothesis. As you may come across a null hypothesis, hypothesis of no differences, these are all formulated as alternate hypothesis.

1.3.3 Duhem–Quine thesis:

The Quine-Duhem thesis is a form of the thesis of the underdetermination of theory by empirical evidence. The basic problem is that individual theoretical claims are unable to be confirmed or falsified on their own, in isolation from surrounding hypotheses. For this reason, the acceptance or rejection of a theoretical claim is underdetermined by observation. The thesis can be interpreted in a more radical form that tends to be associated with the epistemic holism of Willard V. O. Quine or in a more restricted form associated with Pierre Duhem. It is primarily an epistemic thesis about the relation between evidence and theory, though in Quine's case it also has semantic overtones connected with his rejection of the analytic-synthetic distinction.

Although a bundle of hypotheses (i.e. a hypothesis and its background assumptions) *as a whole* can be tested against the empirical world and be falsified if it fails the test, the Duhem–Quine thesis says it is impossible to isolate a single hypothesis in the bundle. One solution to the dilemma thus facing scientists is that when we have rational reasons to accept the background assumptions as true (e.g. scientific theories via evidence) we will have rational—albeit nonconclusive—reasons for thinking that the theory tested is probably wrong if the empirical test fails.

As popular as the Duhem–Quine thesis may be in philosophy of science, in reality Pierre Duhem and Willard Van Orman Quine stated very different theses. Duhem believed that only in the field of physics can a single individual hypothesis not be isolated for testing. He says in no uncertain terms that experimental theory in physics is not the same as in fields like physiology and certain branches of chemistry. Also, Duhem's conception of "theoretical group" has its limits, since he states that not all concepts are connected to each other logically. He did not include at all a priori disciplines such as logic and mathematics within the theoretical groups in physics, since they cannot be tested.

Quine, on the other hand, in "Two Dogmas of Empiricism", presents a much stronger version of underdetermination in science. His theoretical group embraces all of human knowledge, including mathematics and logic. He contemplated the entirety of human knowledge as being one unit of empirical significance. Hence all our knowledge, for Quine, would be epistemologically no different from ancient Greek gods, which were posited in order to account for experience.

Quine even believed that logic and mathematics can also be revised in light of experience, and presented quantum logic as evidence for this. Years later he retracted this position; in his book Philosophy of Logic, he said that to revise logic would be essentially "changing the subject". In classic logic, connectives are defined according to truth values. The connectives in a multi-valued logic, however, have a different meaning than those of classic logic. As for quantum logic, it is not even a logic based on truth values, so the logical connectives lose the original meaning of classic logic. Quine also notes that deviant logics usually lack the simplicity of classic logic, and are not so fruitful.

1.3.4 Quine's critique of empiricism:

In his seminal paper "Two Dogmas of Empiricism" (1951), Quine rejected, as what he considered the first dogma, the idea that there is a sharp division between logic and empirical science. He argued, in a vein reminiscent of the later Wittgenstein, that there is nothing in the logical structure of a language that is inherently immune to change, given appropriate empirical circumstances. Just as the theory of special relativity undermines the fundamental idea that events simultaneous to one observer are simultaneous to all observers, so other changes in what human beings know can alter even their most basic and ingrained inferential habits.

The other dogma of empiricism, according to Quine, is that associated with each scientific or empirical sentence is a determinate set of circumstances whose experience by an observer would count as disconfirming evidence for the sentence in question. Quine argued that the evidentiary links between science and experience are not, in this sense, "one to one." The true structure of science is better compared to a web, in which there are interlinking chains of support for any single part. Thus, it is never clear what sentences are disconfirmed by "recalcitrant experience"; any given sentence may be retained, provided appropriate adjustments are made elsewhere. Similar views were expressed by the American philosopher Wilfrid Sellars (1912–89), who rejected what he called the "myth of the given": the idea that in observation, whether of the world or of the mind, any truths or facts are transparently present. The same idea figured prominently in the deconstruction of the "metaphysics of presence" undertaken by the French philosopher and literary theorist Jacques Derrida (1930–2004).

If language has no fixed logical properties and no simple relationship to experience, it may seem close to having no determinate meaning at all. This was in fact the conclusion Quine drew. He argued that, since there are no coherent criteria for determining when two words have the same meaning, the very notion of meaning is philosophically suspect. He further justified this pessimism by means of a thought experiment concerning “radical translation”: a linguist is faced with the task of translating a completely alien language without relying on collateral information from bilinguals or other informants. The method of the translator must be to correlate dispositions to verbal behaviour with events in the alien’s environment, until eventually enough structure can be discerned to impose a grammar and a lexicon. But the inevitable upshot of the exercise is indeterminacy. Any two such linguists may construct “translation manuals” that account for all the evidence equally well but that “stand in no sort of equivalence, however loose.” This is not because there is some determinate meaning—a unique content belonging to the words—that one or the other or both translators failed to discover. It is because the notion of determinate meaning simply does not apply. There is, as Quine said, no “fact of the matter” regarding what the words mean.

1.4 ETHICAL STANDARDS OF PSYCHOLOGICAL RESEARCH: PLANNING, CONDUCTION AND REPORTING RESEARCH

Researchers must consider ethical issues before they begin a research project. Ethical problems can be avoided only by planning carefully and consulting with appropriate individuals and groups prior to doing the research. The failure to conduct research in an ethical manner undermines the entire scientifi c process, impedes the advancement of knowledge, and erodes the public’s respect for scientifi c and academic communities (see Figure 3.2). It can also lead to signify cant legal and fi nancial penalties for individuals and institutions. An important step that researchers must take as they begin to do psychological research is to gain institutional approval.

Prior to conducting any study, the proposed research must be reviewed to determine if it meets ethical standards. Institutional Review Boards (IRBs) review psychological research to protect the rights and welfare of human participants. Institutional Animal Care and Use Committees (IACUCs) review research conducted with animals to ensure that animals are treated humanely.

- Risk/benefit ratio: A subjective evaluation of the risks and benefits of a research project is used to determine whether the research should be conducted.
- Potential risks in psychological research include risk of physical injury, social injury, and mental or emotional stress.
- To protect participants from social risks, information they provide should be anonymous, or if that is not possible, the confidentiality of their information should be maintained.

Informed Consent:

- Researchers and participants enter into a social contract, often using an informed consent procedure.
- Researchers are ethically obligated to describe the research procedures clearly, and answer any questions participants have about the research.
- Research participants must be allowed to withdraw their consent at any time without penalties.
- Individuals must not be pressured to participate in research.

Deception in psychological research:

- Deception in psychological research occurs when researchers withhold information or intentionally misinform participants about the research. By its nature, deception violates the ethical principle of informed consent.
- Deception is considered a necessary research strategy in some psychological research.

Debriefing:

- Debriefing informs participants about the nature of the research and their role in the study and educates them about the research process. The prime goal of debriefing is to have individuals feel good about their participation
- Researchers are ethically obligated to explain to participants their use of deception as soon as is feasible.
- Debriefing allows researchers to learn how participants viewed the procedures, allows potential insights into the nature of the research findings, and provides ideas for future research.

Research with animal:

- Animals are used in research to gain knowledge that will benefit humans, for example, by helping to cure diseases.
- Researchers are ethically obligated to acquire, care for, use, and dispose of animals in compliance with current federal, state, and local laws and regulations, and with professional standards.
- The use of animals in research involves complex issues and is the subject of much debate.

Reporting of psychological research:

- Investigators attempt to communicate their research findings in peer reviewed scientific journals, and the APA Code of Ethics provides guidelines for this process.
- Decisions about who should receive publication credit are based on the scholarly importance of the contribution.
- Ethical reporting of research requires recognizing the work of others by using proper citations and references; failure to do so may result in plagiarism

1.5 PROPOSING AND REPORTING QUANTITATIVE RESEARCH

The purpose of the research proposal (it's job, so to speak) is to convince your research supervisor, committee or university that your research is suitable (for the requirements of the degree program) and manageable (given the time and resource constraints you will face).

The most important word here is "convince" – in other words, your research proposal needs to sell your research idea (to whoever is going to approve it). If it doesn't convince them (of its suitability and manageability), you'll need to revise and resubmit. This will cost you valuable time, which will either delay the start of your research or eat into its time allowance (which is bad news).

A good dissertation or thesis proposal needs to cover the "what", the "why" and the "how" of the research. Let's look at each of these in a little more detail:

WHAT – Your research topic

Your proposal needs to clearly articulate your research topic. This needs to be specific and unambiguous. Your research topic should make it clear exactly what you plan to research and in what context. Here's an example:

Topic: An investigation into the factors which impact female Generation Y consumer's likelihood to promote a specific makeup brand to their peers: a British context

What's being investigated – factors that make people promote a brand of makeup

Who it involves – female Gen Y consumers

In what context – the United Kingdom

So, make sure that your research proposal provides a detailed explanation of your research topic. It should go without saying, but don't start writing your proposal until you have a crystal-clear topic in mind, or you'll end up waffling away a few thousand words.

WHY – Your justification

As we touched on earlier, it's not good enough to simply propose a research topic – you need to justify why your topic is original. In other words, what makes it unique? What gap in the current literature does it fill? If it's simply a rehash of the existing research, it's probably not going to get approval – it needs to be fresh.

But, originality alone is not enough. Once you've ticked that box, you also need to justify why your proposed topic is important. In other words, what value will it add to the world if you manage to find answers to your research questions?

For example, let's look at the sample research topic we mentioned earlier (factors impacting brand advocacy). In this case, if the research could uncover relevant factors, these findings would be very useful to marketers in the cosmetics industry, and would, therefore, have commercial value. That is a clear justification for the research.

So, when you're crafting your research proposal, remember that it's not enough for a topic to simply be unique. It needs to be useful and value-creating – and you need to convey that value in your proposal. If you're struggling to find a research topic that makes the cut, watch our video covering how to find a research topic.

HOW – Your methodology

It's all good and well to have a great topic that's original and important, but you're not going to convince anyone to approve it without discussing the practicalities – in other words:

How will you undertake your research?

Is your research design appropriate for your topic?

Is your plan manageable given your constraints (time, money, expertise)?

While it's generally not expected that you'll have a fully fleshed out research strategy at the proposal stage, you will need to provide a high-level view of your research methodology and some key design decisions. Here are some important questions you'll need to address in your proposal:

Will you take a qualitative or quantitative approach?

Will your design be cross-sectional or longitudinal?

How will you collect your data (interviews, surveys, etc)?

How will you analyse your data (e.g. statistical analysis, qualitative data analysis, etc)?

So, make sure you give some thought to the practicalities of your research and have at least a basic understanding of research methodologies before you start writing up your proposal. The video below provides a good introduction to methodology.

Reporting quantitative research:

A quantitative analysis can give people the necessary information to make decisions about policy and planning for a program or organization. A good quantitative analysis leaves no questions about the quality of data and the authority of the conclusions. Whether in school completing a project or at the highest levels of government evaluating programs, knowing how to write a quality quantitative analysis is helpful. A quantitative analysis uses hard data, such as survey results, and generally requires the use of computer spreadsheet applications and statistical know-how.

Step 1:

Explain why the report is being written in the introduction. Point out the need that is being filled and describe any prior research that has been conducted in the same field. The introduction should also say what future research should be done to thoroughly answer the questions you set out to research. You should also state for whom the report is being prepared.

Step 2:

Describe the methods used in collecting data for the report. Discuss how the data was collected. If a survey was used to collect data, tell the reader how it was designed. You should let the reader know if a survey pilot test was distributed first. Detail the target population, or the group of people being studied. Provide the sample size, or the number of people surveyed. Tell the reader if the sample was representative of the target population, and explain whether you collected enough surveys. Break down the data by gender, race, age and any other pertinent subcategory. Tell the reader about any problems with data collection, including any biases in the survey, missing results or odd responses from people surveyed.

Step 3:

Create graphs showing visual representations of the results. You can use bar graphs, line graphs or pie charts depending to convey the data. Only write about the pertinent findings, or the ones you think matter most, in the body of the report. Any other results can be attached in the appendices at the end of the report. The raw data, along with copies of a blank survey should be in the appendices as well. The reader can refer to all the data to inform his own opinions about the findings.

Step 4:

Write conclusions after evaluating all the data. The conclusion can include an action item for the reader to accomplish. It can also advise that more research needs to be done before any solid conclusions can be made. Only conclusions that can be made based on the findings should be included in the report.

Step 5:

Write an executive summary to attach at the beginning of the report. Executive summaries are quick one to two page recaps of what is in the report. They include shorter versions of the introductions, methods, findings and conclusions. Executive summaries serve to allow readers to quickly understand what is said in the report.

1.6 REFERENCES

1. Shaughnessy, J. J., Zechmeister, E. B. &Zechmeister, J. (2012). *Research methods in psychology*. (9th ed..). NY: McGraw Hill.

2. Elmes, D. G. (2011). *Research Methods in Psychology* (9thed.). Wadsworth Publishing.

3. Goodwin, J. (2009). *Research in Psychology: Methods in Design* (6thed.). Wiley.

4. McBurney, D. H. (2009). *Research methods*. (8th Ed.). Wadsworth Publishing.

*****

2

RESEARCH SETTINGS AND MEHTODS OF DATA COLLECTION

Unit Structure

2.1 INTRODUCTION

Collecting data involves gathering the information obtained from your measures to help test your research question. As such, data collection methods are specific to your project and will be different from those used by other researchers. The types of measures used (e.g. questionnaires, online experiments, etc.) and who you will be recruiting as your participants (e.g. local students, international individuals, clinical groups, etc.) will determine your data collection methods. For instance, you may decide to administer an online survey via Qualtrics or conduct experiments online using Inquisit or OpenSesame.

Regardless of the method of research, data collection will be necessary. The method of data collection selected will primarily depend on the type of information the researcher needs for their study; however, other factors, such as time, resources, and even ethical considerations can influence the selection of a data collection method. All of these factors need to be considered when selecting a data collection method because each method has unique strengths and weaknesses. We will discuss the uses and assessment of the most common data collection methods: observation, surveys, archival data, and tests.

Data collection is the procedure of collecting, measuring and analyzing accurate insights for research using standard validated techniques. A researcher can evaluate their hypothesis on the basis of collected data. In most cases, data collection is the primary and most important step for research, irrespective of the field of research. The approach of data collection is different for different fields of study, depending on the required information.

The most critical objective of data collection is ensuring that information-rich and reliable data is collected for statistical analysis so that data-driven decisions can be made for research.

The research setting can be seen as the physical, social, and cultural site in which the researcher conducts the study. In qualitative research, the focus is mainly on meaning-making, and the researcher studies the participants in their natural setting.

The environment within which studies are run has important consequences for experimental design, the type of data that can be collected and the interpretation of results. so, for example running a study in an experimental laboratory may allow you to control variables in a way you cannot do in field work, and the results may be criticised for not reflecting real life. It is often important to conduct complementary studies in various research settings in order to build arguments for the generalisability of findings.

Data collection techniques include interviews, observations (direct and participant), questionnaires, and relevant documents. The use of multiple data collection techniques and sources strengthens the credibility of outcomes and enables different interpretations and meanings to be included in data analysis. This is known as triangulation (Flick, 2014).

How often the data is collected:

It relies on particular event happenings or even every movement of the subjects life. Therefore often researchers use sampling to gather information through various observation. It is much needed to make sure that the sample of the data is representative of the subjects overall behaviour. Students and new researchers face difficulties to make a decision Survey questionnaire development for dissertation help will give you a confident of your thesis data in safe hands.

A representative sample obtained through:

Time Sampling: Taking a sample from different interval of time randomly, this type is entirely on the time interval report.

Situation Sampling:situation sampling is taking down the readings based on the movement of the subject. This type doesn’t include any time interval the information is taken when there is a need.

2.2 OBSERVATION AND INTERVIEW METHOD

2.2.1 Observation:

The observational method involves the watching and recording of a specific behavior of participants. In general, observational studies have the strength of allowing the researcher to see for themselves how people behave. However, observations may require more time and man-power than other data collection methods, often resulting in smaller samples of participants. Researchers may spend significant time waiting to observe a behavior, or the behavior may never occur during observation. It is important to remember that people tend to change their behavior when they know they are being watched (known as the Hawthorne effect).

Observations may be done in a naturalist setting to reduce the likelihood of the Hawthorne effect. During naturalistic observations, the participants are in their natural environment and are usually unaware that they are being observed. For example, observing students participating in their class would be a naturalist observation. The downside of a naturalistic setting is that the research doesn’t have control over the environment. Imagine that the researcher goes to the classroom to observe those students, and there is a substitute teacher. The change in instructor that day could impact student behavior and skew the data.

If controlling the environment is a concern, a laboratory setting may be a better choice. In the laboratory environment, the researcher can manage confounding factors or distractions that might impact the participants’ behavior. Of course, there are expenses associated with maintaining a laboratory setting, increasing the cost of the study, that would not be associated with naturalist observations. And, again, the Hawthorne effect may impact behavior.

Observation allows researchers to experience a specific aspect of social life and get a firsthand look at a trend, institution, or behavior. Participant observation involves the researcher joining a sample of individuals without interfering with that group’s normal activities in order to document their routine behavior or observe them in a natural context. Often researchers in observational studies will try to blend in seamlessly with the sample group to avoid compromising the results of their observations.

Observational research is a type of descriptive research that differs from most other forms of data gathering in that the researcher’s goal is not to manipulate the variables being observed. While participants may or may not be aware of the researchers’ presence, the researchers do not try to control variables (as in an experiment), or ask participants to respond to direct questions (as in an interview or survey based study). Instead, the participants are simply observed in a natural setting, defined as a place in which behavior ordinarily occurs, rather than a place that has been arranged specifically for the purpose of observing the behavior. Unlike correlational and experimental research which use quantitative data, observational studies tend to use qualitative data.

For example, social psychologists Roger Barker and Herbert Wright studied how a sample of children interacted with their daily environments. They observed the children go to school, play with friends, and complete daily chores, and learned a great deal about how children interact with their environments and how their environments shape their character. Similarly, anthropologist Jane Goodall studied the behavior of chimpanzees, taking careful notes on their tool making, family relationships, hunting, and social behavior. Her early work served as the basis for future research on chimpanzees and animal behavior in general.

Advantages of Observational Studies:

By observing events as they naturally occur, patterns in behavior will emerge and general questions will become more specific. The hypotheses that result from these observations will guide the researcher in shaping data into results.

One advantage of this type of research is the ability to make on-the-fly adjustments to the initial purpose of a study. These observations also capture behavior that is more natural than behavior occurring in the artificial setting of a lab and that is relatively free of some of the bias seen in survey responses. However, the researcher must be careful not to apply his or her own biases to the interpretation. Researchers may also use this type of data to verify external validity, allowing them to examine whether study findings generalize to real world scenarios.

There are some areas of study where observational studies are more advantageous than others. This type of research allows for the study of phenomena that may be unethical to control for in a lab, such as verbal abuse between romantic partners. Observation is also particularly advantageous as a cross-cultural reference. By observing people from different cultures in the same setting, it is possible to gain information on cultural differences.

Disadvantages of Observational Studies:

While observational studies can generate rich qualitative data, they do not produce quantitative data, and thus mathematical analysis is limited. Researchers also cannot infer causal statements about the situations they observe, meaning that cause and effect cannot be determined. Behavior seen in these studies can only be described, not explained.

There are also ethical concerns related to observing individuals without their consent. One way to avoid this problem is to debrief participants after observing them and to ask for their consent at that time. Overt observation, where the participants are aware of the researcher's presence, is another option to overcome this problem. However, this tactic does have its drawbacks. When subjects know they are being watched, they may alter their behavior in an attempt to make themselves look more admirable.

This type of research can also be very time consuming. Some studies require dozens of observation sessions lasting for several hours and sometimes involving several researchers. Without the use of multiple researchers, the chances of observer bias increase; because behavior is perceived so subjectively, it is possible that two observers will notice different things or draw different conclusions from the same behavior.

Observation without Intervention:

- The goals of naturalistic observation are to describe behavior as it normally occurs and to examine relationships among variables.
- Naturalistic observation helps to establish the external validity of laboratory findings.
- When ethical and moral considerations prevent experimental control, naturalistic observation is an important research strategy.

Observation with Intervention:

- Most psychological research uses observation with intervention.
- The three methods of observation with intervention are participant observation, structured observation, and the fi eld experiment.
- Whether "undisguised" or "disguised," participant observation allows researchers to observe behaviors and situations that are not usually open to scientific observation.
- If individuals change their behavior when they know they are being observed ("reactivity"), their behavior may no longer be representative of their normal behavior.
- Often used by clinical and developmental psychologists, structured observations are set up to record behaviors that may be diffi cult to observe using naturalistic observation.
- In a field experiment, researchers manipulate one or more independent variables in a natural setting to determine the effect on behavior.

Structured Observation:

There are a variety of observational methods using intervention that are not easily categorized. These procedures differ from naturalistic observation because researchers intervene to exert some control over the events they are observing. The degree of intervention and control over events is less, however, than that seen in field experiments (which we describe briefly in the next section and in more detail in Chapter 6). We have labeled these procedures structured observation. Often the observer intervenes in order to cause an event to occur or to "set up" a situation so that events can be more easily recorded.

2.2.2 Interviews:

Interviews are a type of qualitative data in which the researcher asks questions to elicit facts or statements from the interviewee. Interviews used for research can take several forms:

Informal Interview: A more conversational type of interview, no questions are asked and the interviewee is allowed to talk freely. General interview guide approach: Ensures that the same general areas of information are collected from each interviewee. Provides more focus than the conversational approach, but still allows a degree of freedom and adaptability in getting the information from the interviewee. Standardized, open-ended interview: The same open-ended questions are asked to all interviewees. This approach facilitates faster interviews that can be more easily analyzed and compared.Closed, fixed-response interview (Structured): All interviewees are asked the same questions and asked to choose answers from among the same set of alternatives.

The interview may be regarded either as an alternative to other survey methods or as a supplementary source of information. Although it is more costly in both time and money than the questionnaire, it is also more flexible. Additional information over and above initial plans can be readily obtained and ambiguity and misunderstanding eliminated immediately.

One of the greatest strengths of the interview—direct verbal communication—is also a source of weakness because variability is so common in social interactions. For an interview to be successful, rapport is generally required. It is most readily established when the interviewer is nonjudgmental, supportive, and understanding. However, these very characteristics lead to variability in social interaction among those interviewed. We could achieve sufficient control over social interactions so that the interviews are more homogeneous. However, this would inevitably lead to a sterile interview situation. This, in turn, would result in less rapport, which, we have noted, is important for a good interview.

Other problems beset the interview, especially when there is more than one interviewer. Different interviewers may vary in the way they ask questions or interpret responses, or in the way respondents react to them. Interviewer differences are common. How do we assess the comparability of different interviewers? If you reflect a moment, you'll realize that the situation is similar to using several raters in noninterview settings and determining the interrater reliability. In the present case, we are asking whether there is inter-interviewer reliability.

One way to achieve greater inter-interviewer reliability is to standardize the interview procedures. While this standardization increases the interview reliability, it decreases its flexibility. Because of these weaknesses, the interview might best be reserved as an exploratory method to generate ideas and hypotheses that can later be tested by the use of other methods.

When **personal interviews** are used to collect survey data, respondents are usually contacted in their homes or in a shopping mall and trained interviewers administer the questionnaire. The personal interview allows greater flexibility in asking questions than does the mail survey. During an interview the respondent can obtain clarifi cation when questions are unclear, and the trained interviewer can follow up incomplete or ambiguous answers to open ended questions. The interviewer controls the order of questions and can ensure that all respondents complete the questions in the same order. Traditionally, the response rate to personal interviews has been higher than that for mail surveys.

Telephone Interviews:

The prohibitive cost of personal interviews and difficulties supervising interviewers have led survey researchers to turn to telephone or Internet surveys. Phone interviewing met with considerable criticism when it was first used because of serious limitations on the sampling frame of potential respondents. Many people had unlisted numbers, and the poor and those in rural areas were less likely to have a phone. By 2000, however, more than 97% of all U.S. households had telephones (U.S. Census Bureau, 2000), and households with unlisted numbers could be reached using random-digit dialing. The random-digit dialing technique permits researchers to contact efficiently a generally representative sample of U.S. telephone owners. Telephone interviewing also provides better access to dangerous neighborhoods, locked buildings, and respondents available only during evening hours (have you ever been asked to complete a telephone survey during dinner?). Interviews can be completed more quickly when contacts are made by phone, and interviewers can be better supervised when all interviews are conducted from one location. The telephone survey, like the other survey methods, is not without its drawbacks. A possible selection bias exists when respondents are limited to those who have telephones and the problem of interviewer bias remains. There is a limit to how long respondents are willing to stay on the phone, and individuals

2.3 QUESTIONNAIRE

The questionnaire is more than simply a list of questions or forms to be completed. When properly constructed, a questionnaire can be used as a scientific instrument to obtain data from large numbers of individuals. Construction of a useful questionnaire that minimizes interfering problems requires experience, skill, thoughtfulness, and time. A major advantage of the questionnaire is that data can be obtained on large numbers of participants quickly and relatively inexpensively. Further, the sample can be very large and geographically representative. Often, anonymity can be easily maintained; that is, identifying information is not associated with the data. When constructed properly, a questionnaire provides data that can be organized easily, tabulated, and analyzed. Because of these apparent advantages, the use of the questionnaire is a popular method.

There are two broad classes of questionnaires: descriptive and analytical. Descriptive questionnaires are usually restricted to factual information, often biographical, which is usually accessible by other means. Job application forms and U.S. Census questionnaires are typically of this type. Analytical questionnaires deal more with information related to attitudes or opinions.

The results of a questionnaire are about as useful as the care and thought that went into its preparation and dissemination. Just as in normal social intercourse, the way questions are formulated and posed may present problems. They may be ambiguous; they may suggest the answer that the researcher “wants”; they may contain loaded words. Ambiguity is relatively easy to eliminate. A pilot project, limited to a small number of respondents, will usually uncover sources of ambiguity of which the researcher was unaware. These may then be corrected. Table 6.7 illustrates several examples of ambiguous and leading survey questions and also suggests improved versions of the questions.

As much as we might wish it to be, completing questionnaires is not a neutral task, devoid of feelings and emotions. Often respondents are somewhat apprehensive about how they will appear in the researcher's eye. They want to look good and do well. Consequently, their responses may reflect their interpretations of the investigator's desires rather than their own beliefs, feelings, or opinions. This is referred to as demand characteristics. We will say more about this later. Obviously, questions should be stated in a neutral way and not in a way that suggests a particular response. A fundamental requirement is that the question should be answerable. If respondents are given answers from which to choose, the options should be clear and independent. Also, different results can occur when open-ended or closed-ended questions are used. In some cases, the questionnaire is sensitive to position effects. Respondents are more likely to skip items placed toward the end of a questionnaire, and the answers are also slightly different when answered.

More attention has been given to response bias than to other sources of possible bias and contamination. As we noted earlier, results can be markedly affected by the sample on which they are based. The problem of sampling bias is compounded in mailed surveys because of the low return rates. The actual sample on which the data analyses are based is generally a subsample of the original sample. Low returns make it difficult to assess the representativeness of the final sample. It is safe to assume that it is biased and that those who participated in the survey are different in some way from those who did not. How important is this difference? It may be considerable, or it may be trivial. Because its importance cannot be assessed, any generalizations based on low returns must be restricted. For this reason, it is important to know the return rate on survey research. Unfortunately, some studies fail to provide this information. Other things being equal, the higher the return rate, the better the survey.

A number of factors affect return/response rates. Some are quite costly, so that economic factors must be balanced against the greater generality permitted by higher rates of return. Methods to increase return rate include follow-up contacts, general delivery and pickup, use of closed-ended rather than open-ended questions wherever possible, use of rewards for participation, and limiting the length of time needed to complete the survey.

Instruments and Inventories are questionnaires that have stood the test of time. That is, they were designed to measure particular attributes and have been demonstrated to do so with validity and reliability. Examples include personality tests, aptitude tests, and achievement tests. Personality tests measure some state or trait of an individual. Examples include the Minnesota Multiphasic Personality Inventory (MMPI), Beck Depression Inventory (BDI), California Psychological Inventory (CPI), and the

Sixteen Personality Factors Questionnaire (16PF). Aptitude tests measure some skill or ability. Examples include the Stanford–Binet Intelligence Scale, the Wechsler Adult Intelligence Scale (WAIS-III), the Wechsler Intelligence Scale for Children (WISC-III), and the Graduate Record Examination (GRE). Achievement tests measure competence in a particular area. Examples include the Stanford Achievement tests that students take as they progress through K–12 grades in school; state licensing exams for teachers, counselors, lawyers, physicians and other professionals; and the major field achievement test that psychology majors at some universities take just prior to graduation.

If you consider a research project in which a questionnaire might be used, it would be wise to determine whether an instrument or inventory already exists to measure the variable of interest. Don't reinvent the wheel. If someone else has already invested the time and effort to develop a measure with known validity and reliability, use it. One of the characteristics of science is that we make information public and continue to build upon what others have done.

2.4 SURVEY RESEARCH

Most surveys involve asking a standard set of questions to a group of participants. In a highly structured survey, subjects are forced to choose from a response set such as "strongly disagree, disagree, undecided, agree, strongly agree"; or "0, 1-5, 6-10, etc." One of the benefits of having forced-choice items is that each response is coded so that the results can be quickly entered and analyzed using statistical software. While this type of survey typically yields surface information on a wide variety of factors, they may not allow for an in-depth understanding of human behavior.

Of course, surveys can be designed in a number of ways. Some surveys ask open-ended questions, allowing each participant to devise their own response, allowing for a variety of answers. This variety may provide deeper insight into the subject than forced-choice questions, but makes comparing answers challenging. Imagine a survey question that asked participants to report how they are feeling today. If there were 100 participants, there could be 100 different answers, which is more challenging and takes more time to code and analyze.

Surveys are useful in examining stated values, attitudes, opinions, and reporting on practices. However, they are based on self-report, and this can limit accuracy. For a variety of reasons, people may not provide honest or complete answers. Participants may be concerned with projecting a particular image through their responses, they may be uncomfortable answering the questions, inaccurately assess their behavior, or they may lack awareness of the behavior being assessed. So, while surveys can provide a lot of information for many participants quickly and easily, the self-reporting may not be as accurate as other methods.

The survey method of data collection is a type of descriptive research, and is likely the most common of the major methods. Surveys have limited use for studying actual social behavior but are an excellent way to gain an understanding of an individual's attitude toward a matter.

Similar to an interview, a survey may use close-ended questions, open-ended questions, or a combination of the two. "Closed-ended questions" are questions that limit the person taking the survey to choose from a set of responses. Multiple choice, check all that apply, and ratings scale questions are all examples of closed-ended questions. "Open-ended questions" are simply questions that allow people to write in their own response.

Surveys are a highly versatile tool in psychology. Although a researcher may choose to only administer a survey to sample of individuals as their entire study, surveys are often used in experimental research as well. For example, a researcher may assign one group of individuals to an experimental condition in which they are asked to focus on all the negative aspects of their week to induce a negative mood, while he assigns another group of people to a control group in which they read a book chapter. After the mood induction, he has both groups fill out a survey about their current emotions. In this example, the mood induction condition is the independent (manipulated) variable, while participants' responses on the emotion survey is the dependent (measured) variable.

Cross-Sectional Design:

- In the cross-sectional design, one or more samples are drawn from the population(s) at one time.
- Cross-sectional designs allow researchers to describe the characteristics of a population or the differences between two or more populations, and correlational findings from cross-sectional designs allow researchers to make predictions.

Longitudinal Design

- In the longitudinal design, the same respondents are surveyed over time in order to examine changes in individual respondents.
- Because of the correlational nature of survey data, it is difficult to identify the causes of individuals' changes over time.
- As people drop out of the study over time (attrition), the final sample may no longer be comparable to the original sample or represent the population.

There are primarily three modes of data collection that can be employed to gather feedback – Mail, Phone, and Online. The method actually used for data-collection is really a cost-benefit analysis:

Mail Surveys:

Pros: Can reach anyone and everyone – no barrier

Cons: Expensive, data collection errors lag time

Phone Surveys:

Pros: High degree of confidence in the data collected, reach almost anyone

Cons: Expensive, cannot self-administer, need to hire an agency

Web/Online Surveys:

Pros: Cheap, can self-administer, very low probability of data errors

Cons: Not all your customers might have an email address/be on the internet, customers may be wary of divulging information online.

Multi-Mode Surveys:

Surveys, where the data is collected via different modes (online, paper, phone etc.), is also another way of going. It is fairly straightforward and easy to have an online survey and have data-entry operators to enter in data (from the phone as well as paper surveys) into the system. The same system can also be used to collect data directly from the respondents.

Writing Great Questions for data collection:

Writing great questions can be considered by an art. Art always requires a significant amount of hard work, practice, and help from others.

Avoid loaded or leading words or questions:

A small change in content can produce effective results. Words such as could, should, might are all used for almost the same purpose, but may produce a 20% difference in agreement to a question. For example, "The management could.. should.. might.. have shut the factory".

Intense words such as – prohibit or action, which represent control or action also produce similar results. For example, "Do you believe that Donald Trump should prohibit insurance companies from raising rates?".

Sometimes the content is just biased. For instance, "You wouldn't want to go to Rudolpho's Restaurant for the organization's annual party, would you?"

Misplaced questions:

Questions should always have reference to the intended context, questions placed out of order or without its requirement should be avoided. Generally, a funnel approach should be implemented – generic questions should be included in the initial section of the questionnaire as a warm-up and specific ones should follow and towards the end, demographic or geographic questions should be included.

Mutually non-overlapping response categories:

Multiple choice answers should be mutually unique in order to provide distinct choices. Overlapping answer options frustrate the respondent and make interpretation difficult at best. Also, the questions should always be precise.

For example: "Do you like water juice?"

This question is vague. In which terms is the liking for orange juice is to be rated? – Sweetness, texture, price, nutrition etc.

Avoid the use of confusing/unfamiliar words:

Asking about industry related terms such as caloric content, bits, bytes, mbs, and other such terms and acronyms can be confusing for respondents. Ensure that the audience understands your language level, terminology and above all, the question you ask.

Non-directed questions give respondents excessive leeway:

What suggestions do you have for improving our shoes? The question is about quality in general, but the respondent may offer suggestions about texture, the type of shoes or variants.

Never force questions:

There will always be certain questions which cross certain privacy rules and since privacy is an important issue for most people, these questions should either be eliminated from the survey or not kept as mandatory. Survey questions about income, family income and status, religious, and political beliefs etc. should always be avoided as they are considered to be intruding and respondents can choose not to answer them.

Unbalanced answer options in scales:

Unbalanced answer options in scales such as Likert Scale and Semantic Scale may be appropriate for some situations and biased in others. When analyzing a pattern in eating habits, a study used a quantity scale that made obese people appear in the middle of the scale with the polar ends reflecting a state where people starve and an irrational amount to consume. There are cases where we usually would not expect poor service such as hospitals.

Questions which cover two points:

What is the fastest and most convenient ISP for your location? The fastest ISP would be expensive and the less expensive ones will most likely be slow. To understand both factors, two separate questions should be asked.

Dichotomous questions:

Dichotomous questions are used in case you want a distinct answer, for example – Yes/No, Male/Female. For example, the question "Do you think Hillary Clinton will win the election?" – The answer can either be Yes or No.

Avoid the use of long questions:

The use of long questions will definitely increase the time taken for completion which will generally lead to an increase in the survey dropout rate. Multiple choice questions are the longest and most complex and open-ended questions are the shortest and easiest to answer.

Advantages of Surveys:

The benefits of this method include its low cost and its large sample size. Surveys are an efficient way of collecting information from a large sample and are easy to administer compared with an experiment. Surveys are also an excellent way to measure a wide variety of unobservable data, such as stated preferences, traits, beliefs, behaviors, and factual information. It is also relatively simple to use statistical techniques to determine validity, reliability, and statistical significance.

Surveys are flexible in the sense that a wide range of information can be collected. Since surveys are a standardized measure, they are relatively free from several types of errors. Only questions of interest to the researcher are asked, codified, and analyzed. Survey research is also a very affordable option for gathering a large amount of data.

Disadvantages of Surveys:

The major issue with this method is its accuracy: since surveys depend on subjects' motivation, honesty, memory, and ability to respond, they are very susceptible to bias. There can be discrepancies between respondents' stated opinions and their actual opinions that lead to fundamental inaccuracies in the data. If a participant expects that one answer is more socially acceptable than another, he may be more motivated to report the more acceptable answer than an honest one.

When designing a survey, a researcher must be wary of the wording, format, and sequencing of the questions, all of which can influence how a participant will respond. In particular, a researcher should be concerned with the reliability of their survey. "Reliability" concerns the degree to which the survey questions are likely to yield consistent results each time. A survey is said to have high reliability if it produces similar results each time. For example, a reliable measure of emotion is one that measures emotion the same way each time it is used. However, for a survey to be useful, it needs to be not only reliable, but valid. If a measure is has high "validity", this means that it is in fact measuring the concept it was designed to measure (in this case, emotion). It is important to note that a survey can be reliable, but not valid (and vice versa). For example, just because our emotion survey is reliable, and provides us with consistent results each time we administer it, does not necessarily mean it is measuring the aspects of emotion we want it to. In this case, our emotion survey is reliable, but not necessarily valid.

Structured surveys, particularly those with closed-ended questions, may have low validity when researching affective variables. Survey samples tend to be self-selected since the the respondents must choose to complete the survey. Surveys are not appropriate for studying complex social phenomena since they do not give a full sense of these processes.

Information not gathered as part of a controlled experiment or from random assignment of study subjects. Nonexperimental data are commonly used in social science research, particularly when gathering experimental data would be too costly or unethical. Because the researcher cannot control assignment of subjects to the treatment and control groups, nonexperimental data are more difficult than experimental data to analyze and interpret. Examples of nonexperimental data include survey data, administrative records, and standardized test scores. They also are known as observational data.

Case study provides a systematic and scientific way of perceiving or examining the events, collecting data, analysing information, and preparing a report. As a result the researcher may gain a sharpened understanding of why the instance happened as it did, and what might become important to look at more extensively in future research. Case studies lend themselves to both generating and testing hypotheses. In other words, case study should be defined as a research strategy, an inquiry that investigates a phenomenon within its real-life context. Case study research means single and multiple case studies, can include quantitative evidence, relies on multiple sources of evidence and benefits from the prior development of theoretical propositions. Case studies are based on evidence of quantitative and qualitative research. Single subject-research provides the statistical framework for making inferences from quantitative case-study data. According to Lamnek (2005) “The case study is a research approach, situated between concrete data taking techniques and methodologic paradigms.” In the past years, case study method was used in the field of clinical psychology to examine the patient’s previous history regarding the person’s mental health status. To know about the patient’s physical and mental health, and to make an accurate diagnosis, it is very important to know about the patient’s past and present health related as well as environmental related problems and issues.

Aside from consulting the primary origin or source, data can also be collected through a third party, a process common with secondary data. It takes advantage of the data collected from previous research and uses it to carry out new research.

Secondary data is one of the two main types of data, where the second type is the primary data. These 2 data types are very useful in research and statistics, but for the sake of this article, we will be restricting our scope to secondary data. Sources of secondary data include books, personal sources, journals, newspapers, websitess, government records etc. Secondary data are known to be readily available compared to that of primary data. It requires very little research and needs for manpower to use these sources.

Archival Records:

- Archival records are the public and private documents describing the activities of individuals, groups, institutions, and governments, and comprise running records and records of specifi c, episodic events.
- Archival data are used to test hypotheses as part of the multimethod approach, to establish the external validity of laboratory findings, and to assess the effects of natural treatments.
- Potential problems associated with archival records include selective deposit, selective survival, and the possibility of spurious relationships. Consider for a moment all of the data about you that exist in various records: birth certificate; school enrollment and grades; credit/debit card purchases; driver's license, employment and tax records; medical records; voting history;e-mail, texting, and cell phone accounts; and if you're active on sites such as

Facebook and Twitter, countless entries describing your daily life. Now multiply this by the millions of other people for whom similar records exist and you will only touch upon the amount of data "out there." Add to this all of the data available for countries, governments, institutions, businesses, media, and you will begin to appreciate the wealth of data available to psychologists to describe people's behavior using archival records. Archival records are the public and private documents describing the activities of individuals, groups, institutions, and governments. Records that are continuously kept and updated are referred to as running records. The records of your academic life (e.g., grades, activities) are an example of running records, as are the continuous records of sports teams and the stock market. Other records, such as personal documents (e.g., birth certificates, marriage licenses), are more likely to describe specific events or episodes, and are referred to as episodic records

2.6 REFERENCES

1. Shaughnessy, J. J., Zechmeister, E. B. &Zechmeister, J. (2012). *Research methods in psychology*. (9th ed..). NY: McGraw Hill.
2. Elmes, D. G. (2011). *Research Methods in Psychology* (9thed.). Wadsworth Publishing.
3. Goodwin, J. (2009). *Research in Psychology: Methods in Design* (6thed.). Wiley.
4. McBurney, D. H. (2009). *Research methods*. (8th Ed.). Wadsworth Publishing.
5. Forrester, M. A. (2010). *Doing Qualitative Research in Psychology: A Practical Guide*. Sage.

*****

3

EXPERIMENTAL AND QUASI-EXPERIMENTAL METHODS

Unit Structure

3.1 INTRODUCTION

We introduced you to the four goals of research in psychology: description, prediction, explanation, and application. Psychologists use observational methods to develop detailed descriptions of behavior, often in natural settings. Survey research methods allow psychologists to describe people's attitudes and opinions. Psychologists are able to make predictions about behavior and mental processes when they discover measures and observations that co-vary (correlations). Description and prediction are essential to the scientific study of behavior, but they are not sufficient for understanding the causes of behavior. Psychologists also seek explanation—the "why" of behavior. We achieve scientific explanation when we identify the causes of a phenomenon.

We will explore how the experimental method is used to test psychological theories as well as to answer questions of practical importance. As we have emphasized, the best overall approach to research is the multimethod approach. We can be more confident in our conclusions when we obtain comparable answers to a research question after using different methods. Our conclusions are then said to have convergent validity. Each method has different shortcomings, but the methods have complementary strengths that overcome these shortcomings. The special strength of the experimental method is that it is especially effective for establishing cause-and-effect relationships. In this chapter we discuss the reasons researchers conduct experiments and we examine the underlying logic of experimental research. Our focus is on a commonly used experimental design—the random groups design. We describe the procedures for forming random groups and the threats to interpretation that apply specifically to the random groups design. Then we describe the procedures researchers use to analyze and interpret the results they obtain in experiments, and also explore how researchers establish the external validity of experimental findings. We conclude the chapter with consideration of two additional designs involving independent groups: the matched groups design and the natural groups design.

3.1.1 Why Psychologists Conduct Experiments:

- Researchers conduct experiments to test hypotheses about the causes of behavior.
- Experiments allow researchers to decide whether a treatment or program effectively changes behavior.

One of the primary reasons that psychologists conduct experiments is to make empirical tests of hypotheses they derive from psychological theories. For example, Pennebaker (1989) developed a theory that keeping in thoughts and feelings about painful experiences might take a physical toll. According to this “inhibition theory,” it’s physically stressful to keep these experiences to oneself.

3.1.2 Experimental Methods:

Pennebaker and his colleagues conducted many experiments in which they assigned one group of participants to write about personal emotional events and another group to write about superficial topics. Consistent with the hypotheses derived from the inhibition theory, participants who wrote about emotional topics had better health outcomes than participants who wrote about superficial topics. Not all the results, however, were consistent with the inhibition theory. For example, students asked to dance expressively about an emotional experience did not experience the same health benefits as students who danced and wrote about their experience. Pennebaker and Francis (1996) did a further test of the theory and demonstrated that cognitive changes that occur through writing about emotional experiences were critical in accounting for the positive health outcomes.

Our brief description of testing the inhibition theory illustrates the general process involved when psychologists do experiments to test a hypothesis derived from a theory. If the results of the experiment are consistent with what is predicted by the hypothesis, then the theory receives support. On the other hand, if the results differ from what was expected, then the theory may need to be modified and a new hypothesis developed and tested in another experiment.

Testing hypotheses and revising theories based on the outcomes of experiments can sometimes be a long and painstaking process, much like combining the pieces to a puzzle to form a complete picture. The self-correcting interplay between experiments and proposed explanations is a fundamental tool psychologists use to understand the causes of the ways we think, feel, and behave.

Well-conducted experiments also help to solve society's problems by providing vital information about the effectiveness of treatments in a wide variety of areas. This role of experiments has a long history in the fi eld of medicine (Thomas, 1992). For example, near the beginning of the 19th century, typhoid fever and delirium tremens were often fatal. The standard medical practice at that time was to treat these two conditions by bleeding, purging, and other similar "therapies." In an experiment to test the effectiveness of these treatments, researchers randomly assigned one group to receive the standard treatment (bleeding, purging, etc.) and a second group to receive nothing but bed rest, good nutrition, and close observation. Thomas (1992) describes the results of this experiment as "unequivocal and appalling" (p. 9): The group given the standard medical treatment of the time did worse than the group left untreated. Treating such conditions using early-19th-century practices was worse than not treating them at all! Experiments such as these contributed to the insight that many medical conditions are self-limited: The illness runs its course, and patients recover on their own.

3.2 INDEPENDENT GROUPS DESIGNS

Independent measures design, also known as between-groups, is an experimental design where different participants are used in each condition of the independent variable. This means that each condition of the experiment includes a different group of participants.

This should be done by random allocation, which ensures that each participant has an equal chance of being assigned to one group or the other.

Independent measures involve using two separate groups of participants; one in each condition.

The independent groups design is an experimental design whereby two groups are exposed to different experimental conditions. Usually half of the participants are assigned to the experimental group where they are exposed to a condition where the independent variable is manipulated. The other half are assigned to a control group for comparison, where no such manipulation occurs.One advantage of using this design is that there are no order effects which affect the outcomes of the experiment. These happen when participants take part in both conditions of the experiment, and their performance differs across conditions as a result. For example, the practice at doing a memory task felt after the first condition could lead to better performance on the second memory task, irrespective of the manipulation of the independent variable.One disadvantage of this design is differences between the experimental and control groups may be due to individual differences between participants., rather than the effect of the independent variable. For example, due to chance, one group may have a better working memory than the other, and when given a memory task, that group will perform better, regardless of the independent variable manipulation, due to pre-disposed advantage. This could be mitigated with random sampling of participants.

Example: a series of five experiment using independent-groups designs, the researchers examined the hypothesized effects of the color red on men’s perceptions of women. In each study, the participants were shown a photograph of a woman. While the woman depicted remained the same, the background color was varied across different conditions. Thus, independent groups comparisons were made for red background vs. backgrounds that were white, gray, blue, or green. After a brief view (5 seconds) of the picture, each participant assessed the woman shown for (general) attractiveness, intelligence, likeability, kindness, and several measures of sexual desirability. In one of the five experiments, a small sample of women also assessed the attractiveness of the women shown. The participants included 172 men and 32 women, all college undergraduates.

The researchers found statistically significant effects of the color red on men’s perceptions of sexual attractiveness of women. Interestingly, the color red had no effect on women evaluating other women or on men’s evaluation of women’s nonsexual attributes, such as intelligence, likeability, or kindness. The results provide strong support for the hypothesized “red effect.” Even a brief (5-second) glimpse of red enhances men’s attraction to women. Similar results have been reported for other animals. The researchers discuss their results and implications for studies in interpersonal and sexual attraction.

The logic of the experimental method and the application of control techniques that produce internal validity can be illustrated in an experiment investigating girls’ dissatisfaction with their body, conducted in the United Kingdom by Dittmar, Halliwell, and Ive (2006). Their goal was to determine whether exposure to very thin body images causes young girls to experience negative feelings about their own body. Many experiments conducted with adolescent and adult participants demonstrate that women report greater dissatisfaction about themselves after exposure to a thin female model compared to other types of images. Dittmar and her colleagues sought to determine whether similar effects are observed for girls as young as 5 years old. The very thin body image they tested was the

Barbie doll. Anthropological studies that compare the body proportions of Barbie to actual women reveal that the Barbie doll has very unrealistic body proportions, yet Barbie has become a sociocultural ideal for female beauty (see Figure 6.1).

In the experiment small groups of young girls (51D 2–61D 2 years old) were read a story about "Mira" as she went shopping for clothes and prepared to go to a birthday party. As they heard the story, the girls looked at picture books with six scenes related to the story. In one condition of the experiment, the picture books had images of Barbie in the scenes of the story (e.g., shopping for a party outfit, getting ready for the party). In a second condition the picture books had similar scenes but the fi gure pictured was the "Emme" doll. The Emme fashion doll is an attractive doll with more realistic body proportions, representing a U.S. dress size 16 (see Figure 6.2). Finally, in the third condition of the experiment the picture books did not depict Barbie or Emme (or any body) but, instead, showed neutral images related to the story (e.g., windows of clothes shops, colorful balloons). These three versions of the picture books (Barbie, Emme, neutral) represent three levels of the independent variable that was manipulated in the experiment. Because different groups of girls participated in each level of the independent variable, the experiment is described as an independent groups design.

Manipulation:

Dittmar et al. (2006) used the control technique of manipulation to test their hypotheses about girls' body dissatisfaction. The three conditions of the independent variable allowed these researchers to make comparisons relevant to their hypotheses. If they tested only the Barbie condition, it would be impossible to determine whether those images influenced girls' body dissatisfaction in any way. Thus, the neutral-image condition created a comparison—a way to see if the girls' body dissatisfaction differed depending on whether they looked at a thin ideal vs. neutral images. The Emme condition added an important comparison. It is possible that any images of bodies might influence girls' perceptions of themselves. Dittmar and her colleagues tested the hypothesis that only thin body ideals, as represented by Barbie, would cause body dissatisfaction.

At the end of the story, the young girls turned in their picture books and completed a questionnaire designed for their age level. Although Dittmar and her colleagues used a number of measures designed to assess the girls' satisfaction with their body, we will focus on one measure, the Child Figure Rating Scale. This scale has two rows of seven line drawings of girls' body shapes ranging from very thin to very overweight. Each girl was asked first to color in the figure in the top row that most looks like her own body right now (a measure of perceived actual body shape). Then, on a second row of the same figures, the girls were asked to color in the figure that shows the way they most want to look (ideal body shape). Girls were told they could pick any of the figures and that they could choose the same figure in each row. A body shape dissatisfaction score, the dependent variable, was computed by counting the number of figures between each girl's actual shape and her ideal shape. A score of zero indicated no body shape dissatisfaction, a negative score indicated a desire to be thinner, and a positive score indicated a desire to be bigger.

Holding Conditions Constant:

In Dittmar et al.'s experiment, several factors that could have affected the girls' attitudes toward their body were kept the same across the three conditions. All of the girls heard the same story about shopping and attending a birthday party, and they looked at their picture books for the same amount of time. They all received the same instructions throughout the experiment and received the exact same questionnaire at the conclusion.

Researchers use holding conditions constant to make sure that the independent variable is the only factor that differs systematically across the groups. If the three groups had differed on a factor other than the picture books, then the results of the experiment would have been uninterpretable. Suppose the participants in the Barbie condition had heard a different story, for example, a story about Barbie being thin and popular. We wouldn't know whether the observed difference in the girls' body dissatisfaction was due to viewing the images of Barbie or to the different story. When the independent variable of interest and a different, potential independent variable are allowed to c-ovary, a confounding is present. When there are no confoundings, an experiment has internal validity.

Holding conditions constant is a control technique that researchers use to avoid confoundings. By holding constant the story the girls heard in the three conditions, Dittmar and her colleagues avoided confoundings by this factor. In general, a factor that is held constant cannot possibly co-vary with the manipulated independent variable. More importantly, a factor that is held constant does not change, so it cannot possibly covary with the dependent variable either. Thus, researchers can rule out factors that are held constant as potential causes for the observed results.

It is important to recognize, however, that we choose to control only those factors we think might influence the behavior we are studying—what we consider plausible alternative causes. For instance, Dittmar et al. held constant the story the girls heard in each condition. It is unlikely, however, that they controlled factors such as the room temperature to be constant across the conditions because room temperature probably would not likely affect body image (at least when varying only a few degrees). Nevertheless, we should constantly remain alert to the possibility that there may be confounding factors in our experiments whose influence we had not anticipated or considered.

Balancing:

Clearly, one key to the logic of the experimental method is forming comparable (similar) groups at the start of the experiment. The participants in each group should be comparable in terms of various characteristics such as their personality, intelligence, and so forth (also known as individual differences). The control technique of balancing is required because these factors often cannot be held constant. The goal of random assignment is to establish equivalent groups of participants by balancing, or averaging, individual differences across where R1, R2, and R3 refer to the random assignment of subjects to the three independent conditions of the experiment; X1 is one level of an independent variable (e.g., Barbie), X2 is a second level of the independent variable (e.g., Emme), and X3 is a third level of the independent variable (e.g., neutral images). An observation of behavior (O1) in each group is then made.

Block Randomization:

Block randomization balances subject characteristics and potential confoundings that occur during the time in which the experiment is conducted, and it creates groups of equal size. Common procedure for carrying out random assignment is block randomization. First, let us describe exactly how block randomization is carried out, and then we will look at what it accomplishes. Suppose we have an experiment with five conditions (labeled, for convenience, as A, B, C, D, and E). One "block" is made up of a random order of all five conditions:

One block of conditions '! Random order of conditions A B C D E C A E B D

In block randomization, we assign subjects to conditions one block at a time. In our example with five conditions, five subjects would be needed to complete the first block with one subject in each condition. The next five subjects would be assigned to one of each of the five conditions to complete a second block, and so on. If we want to have 10 subjects in each of five conditions, then there would be 10 blocks in the block-randomized schedule. Each block would consist of a random arrangement of the fi ve conditions. This procedure is illustrated below for the first 11 participants.

Threats to Internal Validity:

- Randomly assigning intact groups to different conditions of the independent variable creates a potential confounding due to preexisting differences among participants in the intact groups.
- Block randomization increases internal validity by balancing extraneous variables across conditions of the independent variable.
- Selective subject loss, but not mechanical subject loss, threatens the internal validity of an experiment.
- Placebo control groups are used to control for the problem of demand characteristics, and double-blind experiments control both demand characteristics and experimenter effects.

3.3 REPEATED MEASURES DESIGNS

Repeated Measures design is an experimental design where the same participants take part in each condition of the independent variable. This means that each condition of the experiment includes the same group of participants. Repeated Measures design is also known as within groups, or within-subjects design.

Thus far we have considered experiments in which subjects participate in only one condition of the experiment. They are randomly assigned to one condition in the random groups and matched groups designs, or they are selected to be in one group in natural groups designs. These independent groups designs are powerful tools for studying the effects of a wide range of independent variables.

There are times, however, when it is more effective to have each subject participate in all the conditions of an experiment. These designs are called repeated measures designs (or within-subjects designs). In an independent groups design, a separate group serves as a control for the group given the experimental treatment. In a repeated measures design, subjects serve as their own controls because they participate in both the experimental and control conditions.

We begin this chapter by exploring the reasons why researchers choose to use a repeated measures design. We then describe one of the central features of repeated measures designs. Specifically, in repeated measures designs, participants can undergo changes during the experiment as they are repeatedly tested. Participants may improve with practice, for example, because they learn more about the task or because they become more relaxed in the experimental situation. They also may get worse with practice—for example, because of fatigue or reduced motivation. These temporary changes are called practice effects. We described in Chapter 6 that individual differences among participants cannot be eliminated in the random groups design, but they can be balanced by using random assignment. Similarly, the practice effects that participants experience due to repeated testing in the repeated measures designs cannot be eliminated. Like individual differences in the random groups design, however, practice effects can be balanced, or averaged, across the conditions of a repeated measures design experiment.

When balanced across the conditions, practice effects are not confounded with the independent variable and the results of the experiment are interpretable. Our primary focus in this chapter is to describe the techniques that researchers can use to balance practice effects. We also introduce data analysis procedures for repeated measures designs. We conclude the chapter with a consideration of problems that can arise in repeated measures designs.

Researchers choose to use a repeated measures design in order to

(1) Conduct an experiment when few participants are available,

(2) Conduct the experiment more efficiently,

(3) Increase the sensitivity of the experiment, and (4) study changes in participants' behavior over time.

Researchers gain several advantages when they choose to use a repeated measures design. First, repeated measures designs require fewer participants than an independent groups design, so these designs are ideal for situations in which only a small number of participants is available. Researchers who do experiments with children, the elderly or special populations such as individuals with brain injuries frequently have a small number of participants available.

3.3.1The role of practice effects in repeated measures designs:

- Repeated measures designs cannot be confounded by individual differences variables because the same individuals participate in each condition (level)of the independent variable.
- Participants' performance in repeated measures designs may change across conditions simply because of repeated testing (not because of the independent variable); these changes are called practice effects.
- Practice effects may threaten the internal validity of a repeated measures experiment when the different conditions of the independent variable are presented in the same order to all participants.
- There are two types of repeated measures designs (complete and incomplete) that differ in the specific ways in which they control for practice effects.

The repeated measures designs have another important advantage in addition to the ones we have already described. In a repeated measures design, the characteristics of the participants cannot confound the independent variable being manipulated in the experiment. The same participants are tested in all the conditions of a repeated measures design, so it is impossible to end up with brighter, healthier, or more motivated participants in one condition than in another condition. Stated more formally, there can be no confounding by individual differences variables in repeated measures designs. The absence of the potential for confounding by individual differences variables is a great advantage of the repeated measures designs. This does not mean, however, that there are no threats to the internal validity of experiments that are done using repeated measures designs.

3.3.2 Balancing Practice Effects in the Complete Design:

- Practice effects are balanced in complete designs within each participant using block randomization or ABBA counterbalancing.
- In block randomization, all of the conditions of the experiment (a block) are randomly ordered each time they are presented.
- In ABBA counterbalancing, a random sequence of all conditions is presented, followed by the opposite of the sequence.

- Block randomization is preferred over ABBA counterbalancing when practice effects are not linear, or when participants' performance can be affected by anticipation effects.

3.3.3 Block Randomization:

We introduced block randomization in Chapter 6 as an effective technique for assigning participants to conditions in the random groups design. Block randomization can also be used to order the conditions for each participant in a complete design. For instance, Sackeim et al. administered each of the three versions of their photographs (left composite, original, and right composite) 18 times to each participant. The sequence of 54 trials is broken up into 18 blocks of 3 trials. Each block of trials contains the three conditions of the experiment in random order. In general, the number of blocks in a block randomized schedule is equal to the number of times each condition is administered, and the size of each block is equal to the number of conditions in the experiment.

3.3.4 ABBA Counterbalancing:

In its simplest form, ABBA counterbalancing can be used to balance practice effects in the complete design with as few as two administrations of each condition. ABBA counterbalancing involves presenting the conditions in one sequence (i.e., A then B) followed by the opposite of that same sequence (i.e., B then A). Its name describes the sequences when there are only two conditions (A and B) in the experiment, but ABBA counterbalancing is not limited to experiments with just two conditions. Sackeim et al. could have presented the versions of their photographs according to the ABBA sequence outlined in the top row of Ta ble 7.2 labeled "Condition." Note that in this case it literally would be ABCCBA since there are three conditions.

3.3.5 Balancing Practice Effects in the Incomplete Design

- Practice effects are balanced across subjects in the incomplete design rather than for each subject, as in the complete design.
- The rule for balancing practice effects in the incomplete design is that each condition of the experiment must be presented in each ordinal position (first, second, etc.) equally often.
- The best method for balancing practice effects in the incomplete design with four or fewer conditions is to use all possible orders of the conditions.
- Two methods for selecting specific orders to use in an incomplete design are the Latin Square and random starting order with rotation.
- Whether using all possible orders or selected orders, participants should be randomly assigned to the different sequences.

The preferred technique for balancing practice effects in the incomplete design is to use all possible orders of the conditions. Each participant is randomly assigned to one of the orders. With only two conditions there are only two possible orders (AB and BA); with three conditions there are six possible orders (ABC, ACB, BAC, BCA, CAB, CBA). In general, there are N! (Which is read "N factorial") possible orders with N conditions, where N! Equals N(N - 1) (N - 2) . . . (N - [N - 1]). As we just saw, there are six possible orders with three conditions, which is 3! (3 x 2 x 1 - 6). The number of required orders increases dramatically with increasing numbers of conditions. For instance, for five conditions there are 120 possible orders, and for six conditions there are 720 possible orders. Because of this, the use of all possible orders is usually limited to experiments involving four or fewer conditions.

We have just described the preferred method for balancing practice effects in the incomplete design, all possible orders. There are times, however, when the use of all possible orders is not practical. For example, if we wanted to use the incomplete design to study an independent variable with seven levels, we would need to test 5,040 participants if we used all possible orders—one participant for each of the possible orders of the seven conditions (7! orders). We obviously need some alternative to using all possible orders if we are to use the incomplete design for experiments with fi ve or more conditions.

Practice effects can be balanced by using just some of all the possible orders. The number of selected orders will always be equal to some multiple of the number of conditions in the experiment. For example, to do an experiment with one independent variable with seven levels, we need to select 7, 14, 21, 28, or some other multiple of seven orders to balance practice effects.

3.3.6 The Problem of Differential Transfer:

- Differential transfer occurs when the effects of one condition persist and influence performance in subsequent conditions.
- Variables that may lead to differential transfer should be tested using a random groups design because differential transfer threatens the internal validity of repeated measures designs.
- Differential transfer can be identified by comparing the results for the same independent variable when tested in a repeated measures design and in a random groups design. Researchers can overcome the potential problem of practice effects in repeated measures designs by using appropriate techniques to balance practice effects. There is a much more serious potential problem that can arise in repeated measures designs that is known as differential transfer (Poulton, 1973, 1975, 1982; Poulton & Freeman, 1966). Differential transfer arises when performance in one condition differs depending on the condition that precedes it.

Consider a problem-solving experiment in which two types of instructions are being compared in a repeated measures design. One set of instructions

(A) Is expected to enhance problem solving, whereas the other set of instructions

(B) Serves as the neutral control condition.

It is reasonable to expect that participants tested in the order AB will be unable or unwilling to abandon the approach outlined in the A instructions when they are supposed to be following the B instructions. Giving up the "good thing" participants had under instruction A would be the counterpart of successfully following the admonition "Don't think of pink elephants!" When participants fail to give up the instruction from the first condition (A) while they are supposed to be following instruction B, any difference between the two conditions is reduced. For those participants, after all, condition B was not really tried. The experiment becomes a situation in which participants are tested in an "AA" condition, not an "AB" condition.

In general, the presence of differential transfer threatens internal validity because it becomes impossible to determine if there are true differences between the conditions. It also tends to underestimate differences between the conditions and thereby reduces the external validity of the findings. Therefore, when differential transfer could occur, researchers should choose an independent groups design. Differential transfer is sufficiently common with instructional variables to advise against the use of repeated measures designs for these studies (Underwood & Shaughnessy, 1975). Unfortunately, differential transfer can arise in any repeated measures design. For instance, the effect of 50 units of marijuana may be different if administered after the participant has received 200 units than if administered after the participant has received the placebo (e.g., if the participant has an increased tolerance for marijuana after receiving the 200 dose). There are ways, however, to determine whether differential transfer is likely to have occurred.

The best way to determine whether differential transfer is a problem is to do two separate experiments (Poulton, 1982). The same independent variable would be studied in both experiments, but a random groups design would be used in one experiment and a repeated measures design in the other. The random groups design cannot possibly involve differential transfer because each participant is tested in only one condition. If the experiment using a repeated measures design shows the same effect of the independent variable as that shown in the random groups design, then there has likely been no differential transfer. If the two designs show different effects for the same independent variable, however, differential transfer is likely to be responsible for producing the different outcome in the repeated measures design. When differential transfer does occur, the results of the random groups design should be used to provide the best description of the effect of the independent variable.

3.4 COMPLEX DESIGNS

Complex designs can also be called *factorial designs* because they involve factorial combination of independent variables. *Factorial combination* involves pairing each level of one independent variable with each level of a second independent variable. This makes it possible to determine the effect of each independent variable alone (*main effect*) and the effect of the independent variables in combination (*interaction effect*).

Complex designs may seem a bit complicated at this point, but the concepts will become clearer as you progress through this chapter. We begin with a review of the characteristics of experimental designs that can be used to investigate independent variables in a complex design. We then describe the procedures for producing, analyzing, and interpreting main effects and interaction effects. We introduce the analysis plans that are used for complex designs. We conclude the chapter by giving special attention to the interpretation of interaction effects in complex designs.

3.4.1 Describing effects in a complex design:

- Researchers use complex designs to study the effects of two or more independent variables in one experiment.
- In complex designs, each independent variable can be studied with an independent groups design or with a repeated measures design.
- The simplest complex design is a 2 x 2 design—two independent variables, each with two levels.
- The number of different conditions in a complex design can be determined by multiplying the number of levels for each independent variable
- More powerful and efficient complex designs can be created by including more levels of an independent variable or by including more independent variables in the design.

An experiment with a complex design has, by definition, more than one independent variable. Each independent variable in a complex design must be implemented using either an independent groups design or a repeated measures design according to the procedures described in Chapters 6 and 7. When a complex design has both an independent groups variable and a repeated measures variable, it is called a mixed design.

3.4.2 Main Effects and Interaction Effects:

- The overall effect of each independent variable in a complex design is called a main effect and represents the differences among the average performance for each level of an independent variable collapsed across the levels of the other independent variable.
- An interaction effect between independent variables occurs when the effect of one independent variable differs depending on the levels of the second independent variable.

In any complex factorial design it is possible to test predictions regarding the overall effect of each independent variable in the experiment while ignoring the effect of the other independent variable(s). The overall effect of an independent variable in a complex design is called a main effect. We will examine two main effects Kassin and his colleagues observed in their experiment for two different dependent variables.

Prior to their interrogation of the suspect, student interrogators were given information about interrogation techniques, including a list of possible questions they could ask about the theft. Twelve questions were written as pairs (but presented randomly in the list). One question of the pair was written in such a way that the suspect's guilt was presumed (e.g., "How did you fi nd the key that was hidden behind the VCR?") and the second question in the pair was written so as not to presume guilt (e.g., "Do you know anything about the key that was hidden behind the VCR?"). Student interrogators were asked to select six questions they might later want to ask. Thus, students could select from 0 to 6 questions that presumed guilt. Based on the behavioral confirmation theory, Kassin et al. predicted that interrogators in the guilty-expectation condition would select more guilt-presumptive questions than would interrogators in the innocent-expectation condition. Thus, they predicted a main effect of the interrogator-expectation independent variable.

The data for this dependent variable, number of guilt-presumptive questions selected, are presented in Table 8.1. The overall mean number of guilt presumptive questions for participants in the guilty-expectation condition (3.62) is obtained by averaging the means of the actual-guilt and actual-innocence conditions for interrogators in the guilty-expectation condition: (3.54 + 3.70)/2 = 3.62. Similarly, the overall mean for the innocent-expectation condition is computed to be 2.60: (2.54 + 2.66)/2 = 2.60.1 The means for a main effect represent the overall performance at each level of a particular independent variable collapsed across (averaged over) the levels of the other independent variable. In this case we collapsed (averaged) over the suspect status variable to obtain the means for the main effect of the interrogator expectation variable. The main effect of the interrogator-expectation variable is the difference between the means for the two levels of the variable (3.62 + 2.60 = 1.02). In the Kassin et al. experiment, the main effect of the interrogator-expectation variable indicates that the overall number of guilt-presumptive questions selected was greater when interrogators expected a guilty suspect (3.62) than when they expected an innocent suspect (2.60). Inferential statistics tests confirmed that the main effect of interrogator expectation was statistically significant. This supported the researchers' hypothesis based on behavioral confirmation theory.

Let's now turn to a dependent variable for which there was a statistically significant main effect of the suspect-status independent variable. The researchers also coded the tape-recorded interviews to analyze the techniques used by the interrogators to obtain a confession. Student interrogators were given brief, written instructions regarding the powerful techniques police use to break down a suspect's resistance. Researchers counted the number of interrogator statements that reflected these persuasive techniques, such as building rapport, assertions of the suspect's guilt or disbelief in the suspect's statements, appeals to the suspect's self-interest or conscience, threats of punishment, promises of leniency, and presentation of false evidence.

3.4.3 Complex Designs with Three Independent Variables:

The power and complexity of complex designs increase substantially when the number of independent variables in the experiment increases from two to three. In the two-factor design there can be only one interaction effect, but in the three factor design each independent variable can interact with each of the other two independent variables and all three independent variables can interact together.

Thus, the change from a two-factor to a three-factor design introduces the possibility of obtaining four different interaction effects. If the three independent variables are symbolized as A, B, and C, the three-factor design allows a test of the main effects of A, B, and C; two-way interaction effects of A x B, A x C, B x C; and the three-way interaction effect of A x B x C. The efficiency of an experiment involving three independent variables is remarkable. An experiment investigating discrimination in the workplace will give you a sense of just how powerful complex designs can be. Pingitore, Dugoni, Tindale, and Spring (1994) investigated possible discrimination against moderately obese people in a mock job interview. Participants in the experiment viewed videotapes of job interviews. In one of their experiments they used a 2 x 2 x 2 design. The first independent variable was the weight of the applicant (normal or overweight). The role of the applicant for the job in the videotapes was played by professional actors who were of normal weight. In the moderately obese conditions, the actors wore makeup and prostheses so that they appeared 20% heavier. The second independent variable in the experiment was the sex of the applicant (male or female). The third independent variable was participants' concern about their own body and the importance of body awareness to their self-concept (high or low). This variable was defined using a self-report measure of how participants viewed their body.

A natural groups design was used to study this "body-schema variable." Participants were randomly assigned to evaluate male or female applicants who were normal weight or moderately obese (random groups designs). The dependent variable was the participants' rating on a 7-point scale of whether they would hire the applicant (1 = definitely not hire and 7 = definitely hire). The results of the Pingitore et al. experiment for these three variables are shown in Figure 8.5. As you can see, displaying the means for a three-variable experiment requires a graph with more than one "panel." One panel of the figure shows the results for two variables at one level of the third variable, and the other panel shows results for the same two variables at the second level of the third independent variable.

3.4.4 Interaction Effects and Ceiling and Floor Effects:

When participants' performance reaches a maximum (ceiling) or a minimum (floor) in one or more conditions of an experiment, results for an interaction effect are uninterpretable.

Consider the results of a 3 x 2 experiment investigating the effects of increasing amounts of practice on performance during a physical-fitness test. There were six groups of participants in this plausible but hypothetical experiment. Participants were first given 10, 30, or 60 minutes to practice, doing either easy or hard exercises. Then they took a fi tness test using easy or hard exercises (the same they had practiced). The dependent variable was the percentage of exercises that each participant was able to complete in a 15-minute test period.

The pattern of results in Figure 8.7 looks like a classic interaction effect; the effect of amount of practice time differed for the easy and hard exercises. Increasing practice time improved test performance for the hard exercises, but performance leveled off after 30 minutes of practice with the easy exercises. If a standard analysis was applied to these data, the interaction effect would very likely be statistically significant. Unfortunately, this interaction effect would be essentially uninterpretable. For those groups given practice with the easy exercises, performance reached the maximum level after 30 minutes of practice, so no improvement beyond this point could be shown in the 60-minute group. Even if the participants given 60 minutes of practice had further benefited from the extra practice, the experimenter could not measure this improvement on the chosen dependent variable.

The preceding experiment illustrates the general measurement problem referred to as a ceiling effect. Whenever performance reaches a maximum in any condition of an experiment, there is danger of a ceiling effect. The corresponding name given to this problem when performance reaches a minimum (e.g., zero errors on a test) is a floor effect. Researchers can avoid ceiling and floor effects by selecting dependent variables that allow ample "room" for performance differences to be measured across conditions. For example, in the fitness experiment it would have been better to test participants with a greater number of exercises than anyone could be expected to complete in the time allotted for the test. The mean number of exercises completed in each condition could then be used to assess the effects of the two independent variables without the danger of a ceiling effect. It is important to note that ceiling effects also can pose a problem in experiments that don't involve a complex design. If the fitness experiment had included only the easy exercises, there would still be a ceiling effect in the experiment.

3.5 QUASI-EXPERIMENTAL DESIGNS AND PROGRAM EVALUATION

There are many reasons why researchers do experiments in natural settings. One reason for these "field experiments" is to test the external validity of a laboratory finding (see Chapter 6). That is, we seek to fi nd out if a treatment effect observed in the laboratory works in a similar way in another setting. Other reasons for experimenting in natural settings are more practical.

Research in natural settings is likely to be associated with attempts to improve conditions under which people live and work. The government may experiment with a new tax system or a new method of job training for the economically disadvantaged. Schools may experiment by changing lunch programs, after-school care, or curricula. A business may experiment with new product designs, methods of delivering employee benefits, or flexible work hours. In these cases, as is true in the laboratory, it is important to determine whether the "treatment" caused a change. Did a change in the way patients are admitted to a hospital emergency room cause patients to be treated more quickly and efficiently? Did a college energy conservation program cause a decrease in energy consumption? Knowing whether a treatment was effective permits us to make important decisions about continuing the treatment, about spending additional money, about investing more time and effort, or about changing the present situation on the basis of our knowledge of the results.

Research that seeks to determine the effectiveness of changes made by institutions, government agencies, and other organizations is one goal of program evaluation.

- Quasi-experiments provide an important alternative when true experiments are not possible.
- Quasi-experiments lack the degree of control found in true experiments; most notably, quasi-experiments typically lack random assignment.
- Researchers must seek additional evidence to eliminate threats to internal validity when they do quasi-experiments rather than true experiments.
- The one-group pretest-posttest design is called a pre-experimental design or a bad experiment because it has so little internal validity.

A dictionary will tell you that one definition of the prefix quasi- is "resembling." Quasi-experiments involve procedures that resemble those of true experiments. Generally speaking, quasi-experiments include some type of intervention or treatment and they provide a comparison, but they lack the degree of control found in true experiments. Just as randomization is the hallmark of true experiments, so lack of randomization is the hallmark of quasi experiments. As Campbell and Stanley (1966) explain, quasi-experiments arise when researchers lack the control necessary to perform a true experiment.

Quasi-experiments are recommended when true experiments are not feasible. Some knowledge about the effectiveness of a treatment is more desirable than none. The list of possible threats to internal validity that we reviewed earlier can be used as a checklist in deciding just how good that knowledge is. Moreover, the investigator must be prepared to look for additional kinds of evidence that might rule out a threat to internal validity that is not specifically controlled in a quasi-experiment. For example, suppose that a quasi-experiment does not control for history threats that would be eliminated by a true experiment.

The investigator may be able to show that the history threat is implausible based on a logical analysis of the situation or based on evidence provided by a supplementary analysis. If the investigator can show that the history threat is implausible, then a stronger argument can be made for the internal validity of the quasi-experiment. Researchers must recognize the specifi c shortcomings of quasi-experimental procedures, and they must work like detectives to provide whatever evidence they can to overcome these shortcomings. As we begin to consider the appropriate uses of quasi-experiments, we need to acknowledge that there is a great difference between the power of the true experiment and that of the quasi-experiment. Before facing the problems of interpretation that result from quasi-experimental procedures, the researcher should make every effort possible to approximate the conditions of a true experiment.

Perhaps the most serious limitation researchers face in doing experiments in natural settings is that they are frequently unable to assign participants randomly to conditions. This occurs, for instance, when an intact group is singled out for treatment and when administrative decisions or practical considerations prevent randomly assigning participants. For example, children in one classroom or school and workers at a particular plant represent intact groups that might receive a treatment or intervention without the possibility of randomly assigning individuals to conditions. If we assume that behavior of a group is measured both before and after treatment, such an "experiment" can be described as follows:

O1 X O2:

Where O1 refers to the fi rst observation of a group, or pretest, X indicates a treatment, and O2 refers to the second observation, or posttest. This one-group pretest-posttest design represents a pre-experimental design or, more simply, may be called a bad experiment. Any obtained difference between the pretest and posttest scores could be due to the treatment or to any of several threats to internal validity, including history, maturation, testing, and instrumentation threats (as well as experimenter expectancy effects and novelty effects). The results of a bad experiment are inconclusive with respect to the effectiveness of a treatment. Fortunately, there are quasi-experiments that improve upon this pre-experimental design.

3.5.1 The Nonequivalent Control Group Design:

- In the nonequivalent control group design, a treatment group and a comparison group are compared using pretest and posttest measures. If the two groups are similar in their pretest scores prior to treatment but differ in their posttest scores following treatment, researchers can more confidently make a claim about the effect of treatment.

- Threats to internal validity due to history, maturation, testing, instrumentation, and regression can be controlled in a nonequivalent control group design.

The one-group pretest-posttest design can be modified to create a quasi experimental design with greatly superior internal validity if two conditions are met: (1) there exists a group “like” the treatment group that can serve as a comparison group, and (2) there is an opportunity to obtain pretest and posttest measures from individuals in both the treatment and the comparison groups. Campbell and Stanley (1966) call a quasi-experimental procedure that meets these two conditions a nonequivalent control group design. Because a comparison group is selected on bases other than random assignment, we cannot assume that individuals in the treatment and control groups are equivalent on all important characteristics (i.e., a selection threat arises). Therefore, it is essential that a pretest be given to both groups to assess their similarity on the dependent measure. A nonequivalent control group design can be outlined as follows:

O1 X O2

\- - - - - -

O1 O2

The dashed line indicates that the treatment and comparison groups were not formed by assigning participants randomly to conditions.

By adding a comparison group, researchers can control threats to internal validity due to history, maturation, testing, instrumentation, and regression. A brief review of the logic of experimental design will help show why this occurs. We wish to begin an experiment with two similar groups; then one group receives the treatment and the other does not. If the two groups’ posttest scores differ following treatment, we fi rst must rule out alternative explanations before we can claim that treatment caused the difference. If the groups are truly comparable, and both groups have similar experiences (except for the treatment), then we can assume that history, maturation, testing, instrumentation, and regression effects occur to both groups equally. Thus, we may assume that both groups change naturally at the same rate (maturation), experience the same effect of multiple testing, or are exposed to the same external events (history). If these effects are experienced in the same way by both groups, they cannot possibly be used to account for group differences on posttest measures. Therefore, they no longer are threats to internal validity. Thus, researchers gain atremendous advantage in their ability to make causal claims simply by adding a comparison group. These causal claims, however, depend critically on forming comparable groups at the start of the study, and ensuring that the groups then have comparable experiences, except for the treatment. Because this is difficult to realize in practice, as we’ll see, threats to internal validity due to additive effects with selection typically are not eliminated in this design.

As you approach the end of a course on research methods in psychology, you might appreciate learning about the results of a nonequivalent control group design that examined the effect of taking a research methods course on reasoning about real-life events (Vander Stoep & Shaughnessy, 1997). Students enrolled in two sections of a research methods course (and who happened to be using an edition of this textbook) were compared with students in two sections of a developmental psychology course on their performance on a test emphasizing methodological reasoning about everyday events. Students in both kinds of classes were administered tests at the beginning and at the end of the semester. Results revealed that research methods students showed greater improvement than did students in the control group. Taking a research methods course improved students' ability to think critically about real-life events.

With that bit of encouraging news in mind, let us now examine in detail another study using a nonequivalent control group design. This will give us the opportunity to review both the specific strengths and limitations of this quasi experimental procedure.

Sources of Invalidity in the Nonequivalent Control Group Design:

- To interpret the findings in quasi-experimental designs, researchers examine the study to determine if any threats to internal validity are present.
- The threats to internal validity that must be considered when using the nonequivalent control group design include additive effects with selection, differential regression, observer bias, contamination, and novelty effects.
- Although groups may be comparable on a pretest measure, this does not ensure that the groups are comparable in all possible ways that are relevant to the outcome of the study.

3.5.2 Interrupted Time-Series Designs:

- In a simple interrupted time-series design, researchers examine a series of observations both before and after a treatment.
- Evidence for treatment effects occurs when there are abrupt changes (discontinuities) in the time-series data at the time treatment was implemented.
- The major threats to internal validity in the simple interrupted time-series design are history effects and changes in measurement (instrumentation) that occur at the same time as the treatment.

A second quasi-experiment, a simple interrupted time-series design, is possible when researchers can observe changes in a dependent variable for some time before and after a treatment is introduced (Shadish et al., 2002). The essence of this design is the availability of periodic measures before and after a treatment has been introduced. The simple interrupted time-series design can be outlined in the following way:

$O_1\ O_2\ O_3\ O_4\ O_5\ X\ O_6\ O_7\ O_8\ O_9\ O_{10}$

3.5.3 Program evaluation:

- Program evaluation is used to assess the effectiveness of human service organizations and provide feedback to administrators about their services.
- Program evaluators assess needs, process, outcome, and efficiency of social services.
- The relationship between basic research and applied research is reciprocal.
- Despite society's reluctance to use experiments, true experiments and quasi experiments can provide excellent approaches for evaluating social reforms.

Organizations that produce goods have a ready-made index of success. If a company is set up to make microprocessors, its success is ultimately determined by its profi ts from the sale of microprocessors. At least theoretically, the efficiency and effectiveness of the organization can be easily assessed by examining the company's financial ledgers. Increasingly, however, organizations of a different sort play a critical role in our society. Because these organizations typically provide services rather than goods, Posavac (2011) refers to them as human service organizations. For example, hospitals, schools, police departments, and government agencies provide a variety of services ranging from emergency room care to fi re prevention inspections. Because profit-making is not their goal, some other method must be found to distinguish between effective and ineffective agencies. One useful approach to assessing the effectiveness of human service organizations is program evaluation.

According to Posavac (2011), program evaluation is: a methodology to learn the depth and extent of need for a human service and whether the service is likely to be used, whether the service is sufficiently intensive to meet the unmet needs identified, and the degree to which the service is offered as planned and actually does help people in need at a reasonable cost without unacceptable side effects. (p. 1)

The definition of program evaluation includes several components; we will take up each of these components in turn. Posavac emphasizes, however, that the overarching goal of program evaluation is to provide feedback regarding human service activities. Program evaluations are designed to provide feedback to the administrators of human service organizations to help them decide what services to provide to whom and how to provide them most effectively and efficiently. Program evaluation is an integrative discipline that draws on political science, sociology, economics, education, and psychology. We are discussing program evaluation at the end of this chapter on research in natural settings because it represents perhaps the largest-scale application of the principles and methods we have been describing throughout this book.

Posavac (2011) identifies four questions that are asked by program evaluators. These questions are about needs, process, outcome, and effi ciency. An assessment of needs seeks to determine the unmet needs of the people for whom an agency might provide a service. Consider, for example, a city government that has received a proposal to institute a program of recreational activities for senior citizens in the community. The city would fi rst want to determine whether senior citizens actually need or want such a program. If the senior citizens do want such a program, the city would further want to know what kind of program would be most attractive to them. The methods of survey research are used extensively in studies designed to assess needs. Administrators can use the information obtained from an assessment of needs to help them plan what programs to offer. Once a program has been set up, program evaluators may ask questions about the process that has been established. Observational methods are often useful in assessing the processes of a program.

Programs are not always implemented the way they were planned, and it is essential to know what actually is being done when a program is implemented. If the planned activities were not being used by the senior citizens in a recreational program designed specifi cally for them, it might suggest that the program was inadequately implemented. An evaluation that provides answers to questions about process, that is, about how a program is actually being carried out, permits administrators to make adjustments in the delivery of services in order to strengthen the existing program (Posavac, 2011).

The next set of questions a program evaluator is likely to ask involves the outcome. Has the program been effective in meeting its stated goals? For example, do senior citizens now have access to more recreational activities, and are they pleased with these activities? Do they prefer these particular activities over other activities? The outcome of a neighborhood-watch program designed to curb neighborhood crime might be evaluated by assessing whether there were actual decreases in burglaries and assaults following the implementation of the program. It is possible to use archival data like those described in Chapter 4 to carry out evaluations of outcome. For example, examining police records in order to document the frequency of various crimes is one way to assess the effectiveness of a neighborhood-watch program. Evaluations of outcome may also involve both experimental and quasi-experimental methods for research in natural settings. An evaluator may, for example, use a nonequivalent control group design to assess the effectiveness of a school reform program by comparing students' performance in two different school districts, one with the reform program and one without. The final questions evaluators might ask are about the efficiency of the program.

Most often, questions about efficiency relate to the cost of the program. Choices often have to be made among possible services that a government or other institution is capable of delivering. Information about how successful a program is (outcome evaluation) and information about the program's cost efficiency evaluation) are necessary if we want to make informed decisions about continuing the program, how to improve it, whether to try an alternative program, or whether to cut back on the program's services.

3.6 REFERENCES

1. Shaughnessy, J. J., Zechmeister, E. B. &Zechmeister, J. (2012). *Research methods in psychology*. (9th ed..). NY: McGraw Hill.

2. Elmes, D. G. (2011). *Research Methods in Psychology* (9thed.). Wadsworth Publishing.

3. Goodwin, J. (2009). *Research in Psychology: Methods in Design* (6thed.). Wiley.

4. McBurney, D. H. (2009). *Research methods*. (8th Ed.). Wadsworth Publishing.

5. Forrester, M. A. (2010). *Doing Qualitative Research in Psychology: A Practical Guide*. Sage.

*****

4

QUALITATIVE RESEARCH

Unit Structure

4.1 INTRODUCTION

Qualitative research is defined as a market research method that focuses on obtaining data through open-ended and conversational communication.

This method is not only about "what" people think but also "why" they think so. For example, consider a convenience store looking to improve its patronage. A systematic observation concludes that the number of men visiting this store are more. One good method to determine why women were not visiting the store is to conduct an in-depth interview of potential customers in the category.

For example, on successfully interviewing female customers, visiting the nearby stores and malls, and selecting them through random sampling, it was known that the store doesn't have enough items for women and so there were fewer women visiting the store, which was understood only by personally interacting with them and understanding why they didn't visit the store, because there were more male products than female ones.

Qualitative research is based on the disciplines of social sciences like psychology, sociology, and anthropology. Therefore, the qualitative research methods allow for in-depth and further probing and questioning of respondents based on their responses, where the interviewer/researcher also tries to understand their motivation and feelings. Understanding how your audience takes decisions can help derive conclusions in market research.

4.2 PHILOSOPHY AND CONCEPTUAL FOUNDATIONS; PROPOSING AND REPORTING QUALITATIVE RESEARCH

4.2.1 Philosophy And Conceptual Foundations:

All knowledge production is based on a set of philosophical assumptions about the nature of reality (ontology), the nature of knowledge (epistemology), and the ways in which we acquire knowledge (methodology). These are known as paradigms, or worldviews (Kuhn, 1970). They refer to researchers' assumptions about the world and are often implicit or taken for granted. Paradigm assumptions include claims about notions such as subjectivity, objectivity, truth, knowledge, and reality. Paradigms inform the kinds of questions that can be asked and answered through research – and those that cannot. Paradigms guide both the researcher and the research inquiry (Kuhn, 1970; Guba & Lincoln, 1994).

For example, a post-positivist qualitative researcher may favour results presented as description to better fit with the dominant ideas of the post-positivist paradigm. An assumption of this paradigm is that the researcher should stay as close as possible to participants' words and their descriptions of events. Interpretive qualitative researchers, on the other hand, would explicitly engage with interpretation throughout the study process. They want to not only describe a phenomenon but to also explain analytical insights they have gleaned about it through the study. Interpretive researchers might develop a new concept based on a theoretically informed analysis and interpretation of participants' words and their descriptions of events. They might offer an explanation that is analytically or conceptually generalizable beyond the study itself and possibly transferable to other contexts, such as the "discourse of abuse" in Eakin (2005) or the concept of "talk" in Facey (2010).

Researchers can identify their knowledge-producing paradigms or worldview by thinking about whether they acquire knowledge by being "objective" and "unbiased," by being detached, value-free observers, or by acknowledging their subjectivity. They can also consider whether they see themselves as intimately involved in co-producing knowledge, whether they think the research process and the knowledge produced is, or can be, value-free; and whether they can know and produce knowledge about how things really are and how they really work.

Guba and Lincoln (Guba & Lincoln, 1994) propose four paradigms: positivism, post-positivism, critical social, and constructivism/interpretivism. Both (post-)positivist paradigms assume that a stable reality exists "out there," that phenomena such as health and disease exist whether we look for or find them or not, and that what exists as health and disease are real only if they can be observed through or are amenable to the senses. "Stable" means for example that realities such as our understandings of disease are not affected by factors such as social, political, historical, or economic processes; only what is observable can be considered valid, and knowledge is achieved through the accumulation of verified facts.

From a (post-)positivist perspective, metaphysical notions such as one's feelings would be considered valid knowledge only if they could be observed or measured (Guba & Lincoln, 1994; Green & Thorogood, 2004; Denzin & Lincoln, 2011). This philosophy also holds that researchers must be objective, which means they must rid themselves of their biases because these can taint the research process and thus undermine the validity of the knowledge produced. This orientation is more appropriate for research in the natural sciences. In the health sciences, it focuses on prediction of behaviour and functionalist frameworks to explain social relations.

The second paradigm that is very influential in the health sciences is the critical-social paradigm. In this paradigm, reality is shaped by socio-economic, political, historical, and cultural contexts. Researchers acknowledge their subjectivity and as a result, recognize that truths (e.g., research findings) are value-mediated (Guba & Lincoln, 1994). Critical-social theories are concerned with issues of power – underlying power structures and how they impinge on individuals and groups. Within this paradigm, theories such as neo-Marxism, feminism, postcolonialism, poststructuralism, postmodernism, and critical race studies, among others, explore the power relations that shape current social relations. Researchers are involved in advocacy and committed to social justice (Guba and Lincoln, 1994). Their objective is to produce knowledge to promote social change by identifying forms of oppression and supporting the empowerment of disadvantaged groups (Denzin, 2015). They study how we came to have groups of privileged people benefiting from the current power arrangements while others experience unnecessary suffering and deprivation. For example, why does the nursing profession have less social prestige and remuneration than the medical profession despite providing essential health care? Or, why is there a lack of access to dental care for part-time workers and their families?

Constructivism/constructionism/interpretivism lies at the other end of the continuum in Guba and Lincoln's (1994) typology of paradigms. The theories organized under this paradigm that are better known in the health sciences are phenomenology, social constructionism, and symbolic interactionism. This perspective assumes that reality is multiple, contingent, and socially constructed through social interactions. And, unlike (post-)positivism, it has the capacity to include metaphysical considerations. Interpretivism is concerned with meaning and subjective experiences, with understanding phenomena from the perspective of those who experience it (Green & Thorogood, 2004).

Further, where (post-)positivist researchers assert that knowledge and understanding of health and diseases are products of accumulated facts, constructivist/ interpretivist researchers argue that they are social constructions and that our understandings and experiences of them are informed by social, historical, and political contexts (Singer, 2004). For example, TB sufferers in Canada experience their disease as stigma, depression, fear, isolation, and anxiety; as a limitation on their freedom and autonomy; and as an intrusion that is related to surveillance through Directly Observed Therapy (DOT) programs (Bender, 2009; Bender, Peter, Wynn, Andrews & Pringle, 2011). A constructivist/interpretivist researcher would note that their experiences are shaped in part by their social status as new and/or racialized immigrants, the construction of TB as contagion, and the personal moral judgments that inform such understandings of this disease (Bender, Guruge, Hyman & Janjua, 2012). The assertion that diseases are "social constructions" does not mean they do not exist. Diseases objectively do exist, but this perspective prompts us to carefully consider the ways in which we think and talk about them. As the TB example above suggests, prevailing attitudes toward a particular disease have implications for the people diagnosed with that disease.

Also, in this paradigm, the researcher's values and roles hold primary places in the research process; the researcher is the "orchestrator and facilitator of the inquiry" (Guba & Lincoln, 1994; Denzin & Lincoln, 2011). The researcher and the participant are also inexorably linked in a research relationship. This means, for example, that the research data and by extension research results are co-created in the research process. From this perspective, researchers do not make claims of objectivity, but rather acknowledge and engage their thoughts and feelings during the research process. They "account for themselves" through the ethical and epistemological lens of reflexivity (Denzin and Lincoln, 2011). These reflexive practices not only become a resource that informs the research inquiry and outcomes, they also buttress the rigour or quality of research because they contribute to transparency in research practice and process.

4.2.2. Proposing and reporting Qualitative research:

Writing the proposal of a research work in the present era is a challenging task due to the constantly evolving trends in the qualitative research design and the need to incorporate medical advances into the methodology. The proposal is a detailed plan or 'blueprint' for the intended study, and once it is completed, the research project should flow smoothly. Even today, many of the proposals at post-graduate evaluation committees and application proposals for funding are substandard.

A proposal needs to show how your work fits into what is already known about the topic and what new paradigm will it add to the literature, while specifying the question that the research will answer, establishing its significance, and the implications of the answer. The proposal must be capable of convincing the evaluation committee about the credibility, achievability, practicality and reproducibility (repeatability) of the research design. Four categories of audience with different expectations may be present in the evaluation committees, namely academic colleagues, policy-makers, practitioners and lay audiences who evaluate the research proposal. Tips for preparation of a good research proposal include; ‘be practical, be persuasive, make broader links, aim for crystal clarity and plan before you write’.

A researcher must be balanced, with a realistic understanding of what can be achieved. Being persuasive implies that researcher must be able to convince other researchers, research funding agencies, educational institutions and supervisors that the research is worth getting approval. The aim of the researcher should be clearly stated in simple language that describes the research in a way that non-specialists can comprehend, without use of jargons. The proposal must not only demonstrate that it is based on an intelligent understanding of the existing literature but also show that the writer has thought about the time needed to conduct each stage of the research.

Reporting of qualitative research results should identify the main analytic findings. Often, these findings involve interpretation and contextualization, which represent a departure from the tradition in quantitative studies of objectively reporting results. The presentation of results often varies with the specific qualitative approach and methodology; thus, rigid rules for reporting qualitative findings are inappropriate. However, authors should provide evidence (e.g., examples, quotes, or text excerpts) to substantiate the main analytic findings.

Qualitative research is expansive and occasionally controversial, spanning many different methods of inquiry and epistemological approaches. A “one-size-fits-all” standard for reporting qualitative research can be restrictive, but COREQ and SRQR both serve as valuable tools for developing responsible qualitative research proposals, effectively communicating research decisions, and evaluating submissions. Ultimately, tailoring a set of standards specific to health design research and its frequently used methods would ensure quality research and aid reviewers in their evaluations.

4.3 GROUNDED THEORY

“Grounded theory refers to a set of systematic inductive methods for conducting qualitative research aimed toward theory development. The term grounded theory denotes dual referents: (a) a method consisting of flexible methodological strategies and (b) the products of this type of inquiry. Increasingly, researchers use the term to mean the methods of inquiry for collecting and, in particular, analyzing data. The methodological strategies of grounded theory are aimed to construct middle-level theories directly from data analysis. The inductive theoretical thrust of these methods is central to their logic. The resulting analyses build their power on strong empirical foundations. These analyses provide focused, abstract, conceptual theories that explain the studied empirical phenomena.

Grounded theory has considerable significance because it (a) provides explicit, sequential guidelines for conducting qualitative research; (b) offers specific strategies for handling the analytic phases of inquiry; (c) streamlines and integrates data collection and analysis; (d) advances conceptual analysis of qualitative data; and (e) legitimizes qualitative research as scientific inquiry. Grounded theory methods have earned their place as a standard social research method and have influenced researchers from varied disciplines and professions.

Yet grounded theory continues to be a misunderstood method, although many researchers purport to use it. Qualitative researchers often claim to conduct grounded theory studies without fully understanding or adopting its distinctive guidelines. They may employ one or two of the strategies or mistake qualitative analysis for grounded theory. Conversely, other researchers employ grounded theory methods in reductionist, mechanistic ways. Neither approach embodies the flexible yet systematic mode of inquiry, directed but open-ended analysis, and imaginative theorizing from empirical data that grounded theory methods can foster. Subsequently, the potential of grounded theory methods for generating middle-range theory has not been fully realized."

You should consider using grounded theory when there is no existing theory that offers an explanation for a phenomenon that you are studying. It can also be used if there is an existing theory, but it is potentially incomplete as the data used to derive that theory wasn't collected from the group of participants that you plan on researching.

4.3.1 Benefits of using grounded theory:

Findings accurately represent real world settings:

The theories you develop using grounded theory are derived directly from real world participants in real world settings using methods like in depth interviews and observation, so your findings will more accurately represent the real world. This is in contrast to other research approaches that occur in less natural settings like research labs or focus group tables.

Findings are tightly connected to the data:

Because grounded theory primarily relies on collected data to determine the final outcome, the findings are tightly connected to that data. This is in contrast to other research approaches that rely more heavily on external research frameworks or theories that are further removed from the data.

Great for new discoveries:

Grounded theory is a strong, inductive research method for discovering new theories. You don't go in with any preconceived hypothesis about the outcome, and are not concerned with validation or description. Instead, you allow the data you collect to guide your analysis and theory creation, leading to novel discoveries.

Offers strategies for analysis:

The process of grounded theory describes specific strategies for analysis that can be incredibly helpful. While grounded theory is a very open ended methodology, the analysis strategies enable you to stay structured and analytical in your discovery process.

Data collection and analysis are streamlined:

Data collection and analysis are tightly interwoven. As you collect data, you analyze it, and as you learn from analysis, you continue to collect more data. This helps ensure that the data you collect is sufficient enough to explain the findings that arise from analysis.

Buffers against confirmation bias:

Because data collection and analysis are tightly interwoven, you are truly following what is emerging from the data itself. This provides a great buffer against confirming preconceived beliefs about your topic.

4.3.2 Limitations of grounded theory:

Difficulty recruiting:

Grounded theory relies on an iterative recruiting process called theoretical sampling where you continuously recruit and conduct new rounds of interviews with new participants and previous participants while you analyze data. The recruiting criteria also evolves and changes based on what you learn. Because the recruiting is not predefined, it can be challenging to continuously find the right participants for your study.

Time consuming to collect data:

There is no way to know ahead of time how much data you will need to collect, so you need to be flexible with your time. With grounded theory, you continuously collect and analyze data until you reach theoretical saturation, which is the point at which new data does not contribute new insight to your evolving theory. This means that you are likely to conduct many rounds of data collection before your theory is complete.

Challenges in analysis:

Data analysis occurs on a rolling basis and involves making constant comparisons between different excerpts of data. It can be challenging to keep track of your comparisons and findings as you go. It can be helpful to use a qualitative data analysis software like Delve to help you stay organized during your analysis.

Steps for grounded theory:

1. Determine initial research questions
2. Recruit and collect data (theoretical sampling)
3. Break transcripts into excerpts (open coding)
4. Group excerpts into codes (open coding)
5. Group codes into categories (axial coding)

6. Analyze more excerpts and compare with codes
7. Repeat steps 2-6 until you reach theoretical saturation
8. Define the central idea (selective coding)
9. Write your grounded theory

Grounded theory is not a linear process where you collect data, analyze it, and then you're done. It is an iterative research methodology that involves cycling through the steps iteratively. Part of what made Grounded Theory revolutionary was that it mixed data collection with analysis. It emphasized going back to the field even after conducting some analysis. You will recruit some participants, gather data and analyse it, and go back into the field again with a different recruiting strategy and focus of inquiry. Then you'll incorporate those findings into further rounds of analysis. Grounded theory is deliberately cyclical in nature.

4.4 INTERPRETIVE PHENOMENOLOGICAL ANALYSIS; DISCOURSE ANALYSIS

4.4.1 Interpretative phenomenological analysis (IPA):

Interpretative phenomenological analysis (IPA) is an approach to psychological qualitative research with an idiographic focus, which means that it aims to offer insights into how a given person, in a given context, makes sense of a given phenomenon. Usually, these phenomena relate to experiences of some personal significance, such as a major life event, or the development of an important relationship. It has its theoretical origins in phenomenology and hermeneutics, and key ideas from Edmund Husserl, Martin Heidegger, and Maurice Merleau-Ponty are often cited. IPA is one of several approaches to qualitative, phenomenological psychology. It is distinct from other approaches, in part, because of its *combination* of psychological, interpretative, and idiographic components.

Sometimes IPA studies involve a close examination of the experiences and meaning-making activities of only one participant. Most frequently they draw on the accounts of a small number of people (6 has been suggested as a good number, although anywhere between 3 and 15 participants for a group study can be acceptable· In either case, participants are invited to take part precisely because they can offer the researcher some meaningful insight into the topic of the study; this is called purposive sampling [i.e. it is not randomised]. Usually, participants in an IPA study are expected to have certain experiences in common with one another: the small-scale nature of a basic IPA study shows how something is understood in a given context, and from a shared perspective, a method sometimes called homogeneous sampling. More advanced IPA study designs may draw together samples that offer multiple perspectives on a shared experience (husbands and wives, for example, or psychiatrists and patients); or they may collect accounts over a period of time, to develop a longitudinal analysis.

Data collection:

In IPA, researchers gather qualitative data from research participants using techniques such as interview, diaries, or focus group. Typically, these are approached from a position of flexible and open-ended inquiry, and the interviewer adopts a stance that is curious and facilitative (rather than, say, challenging and interrogative). IPA usually requires personally salient accounts of some richness and depth, and it requires that these accounts be captured in a way that permits the researcher to work with a detailed verbatim transcript.

Data analysis:

Data collection does not set out to test hypotheses, and this stance is maintained in data analysis. The analyst reflects upon their own preconceptions about the data, and attempts to suspend these in order to focus on grasping the experiential world of the research participant. Transcripts are coded in considerable detail, with the focus shifting back and forth from the key claims of the participant, to the researcher's interpretation of the meaning of those claims. IPA's hermeneutic stance is one of inquiry and meaning-making, and so the analyst attempts to make sense of the participant's attempts to make sense of their own experiences, thus creating a double hermeneutic. One might use IPA if one had a research question which aimed to understand what a given experience was like (phenomenology) and how someone made sense of it (interpretation).

Analysis in IPA is said to be 'bottom-up.' This means that the researcher generates codes *from* the data, rather than using a pre-existing theory to identify codes that might be applied *to* the data. IPA studies do not test theories, then, but they are often relevant to the development of existing theories. One might use the findings of a study on the meaning of sexual intimacy to gay men in close relationships, for example, to re-examine the adequacy of theories which attempt to predict and explain safe sex practices. IPA encourages an open-ended dialogue between the researcher and the participants and may, therefore, lead us to see things in a new light.

After transcribing the data, the researcher works closely and intensively with the text, annotating it closely ('coding') for insights into the participants' experience and perspective on their world. As the analysis develops, the researcher catalogues the emerging codes, and subsequently begins to look for patterns in the codes. These patterns are called 'themes'. Themes are recurring patterns of meaning (ideas, thoughts, feelings) throughout the text. Themes are likely to identify both something that *matters* to the participants (i.e. an object of concern, topic of some import) and also convey something of the *meaning* of that thing, for the participants. E.g. in a study of the experiences of young people learning to drive, we might find themes like 'Driving as a rite of passage' (where one key psychosocial understanding of the meaning of learning to drive, is that it marks a cultural threshold between adolescence and adulthood).

Some themes will eventually be grouped under much broader themes called 'superordinate themes'. For example, 'Feeling anxious and overwhelmed during the first driving lessons' might be a superordinate category that captures a variety of patterns in participants' embodied, emotional and cognitive experiences of the early phases of learning to drive, where we might expect to find sub-themes relating to, say, 'Feeling nervous,' 'Worrying about losing control,' and 'Struggling to manage the complexities of the task.' The final set of themes are typically summarised and placed into a table or similar structure where evidence from the text is given to back up the themes produced by a quote from the text.

4.4.2 Discourse analysis:

Discourse analysis is a research method for studying written or spoken language in relation to its social context. It aims to understand how language is used in real life situations.

When you do discourse analysis, you might focus on:

- The purposes and effects of different types of language
- Cultural rules and conventions in communication
- How values, beliefs and assumptions are communicated
- How language use relates to its social, political and historical context

Discourse analysis is a common qualitative research method in many humanities and social science disciplines, including linguistics, sociology, anthropology, psychology and cultural studies. It is also called critical discourse analysis.

Conducting discourse analysis means examining how language functions and how meaning is created in different social contexts. It can be applied to any instance of written or oral language, as well as non-verbal aspects of communication such as tone and gestures.

Materials that are suitable for discourse analysis include:

- Books, newspapers and periodicals
- Marketing material, such as brochures and advertisements
- Business and government documents
- Websites, forums, social media posts and comments
- Interviews and conversations

By analyzing these types of discourse, researchers aim to gain an understanding of social groups and how they communicate.

Unlike linguistic approaches that focus only on the rules of language use, discourse analysis emphasizes the contextual meaning of language.

It focuses on the social aspects of communication and the ways people use language to achieve specific effects (e.g. to build trust, to create doubt, to evoke emotions, or to manage conflict).

Instead of focusing on smaller units of language, such as sounds, words or phrases, discourse analysis is used to study larger chunks of language, such as entire conversations, texts, or collections of texts. The selected sources can be analyzed on multiple levels.

Discourse analysis is a qualitative and interpretive method of analyzing texts (in contrast to more systematic methods like content analysis). You make interpretations based on both the details of the material itself and on contextual knowledge.

4.5 NARRATIVE ANALYSIS; CONVERSATION ANALYSIS

Researchers use narrative analysis to understand how research participants construct story and narrative from their own personal experience. That means there is a dual layer of interpretation in narrative analysis. First the research participants interpret their own lives through narrative. Then the researcher interprets the construction of that narrative.

Narratives can be derived from journals, letters, conversations, autobiographies, transcripts of in-depth interviews, focus groups, or other types of narrative qualitative research and then used in narrative research.

Examples of personal narratives:

Personal narratives come in a variety of forms and can all be used in narrative research.

- Topical stories
- A restricted story about one specific moment in time with a plot, characters, and setting, but doesn't encompass the entirety of a person's life. Example: a research participant's answer to a single interview question

Personal narrative:

- Personal narratives come from a long interview or a series of long narrative interviews that give an extended account of someone's life. Example: a researcher conducting an in-depth interview, or a series of in-depth interviews with an individual over an extended period of time.

Entire life story:

- Constructed from a collection of interviews, observations, and documents about a person's life. Example: a historian putting together the biography of someone's life from past artifacts.

Capturing narrative data:

While humans naturally create narratives and stories when interpreting their own lives, certain data collection methods are more conducive to understanding your research participants' sense of self narrative. Semi-structured interviews, for example, give the interviewee the space to go on narrative tangents and fully convey their internal narratives. Heavily structured interviews that follow a question answer format or written surveys, are less likely to capture narrative data.

Transcribing narrative data:

As mentioned earlier, narrative analysis has dual layers of interpretation. Researchers should not take narrative interviews at face value because they are not just summarizing a research participant's self-narrative. Instead, researchers should actively interpret how the interviewee created that self-narrative. Thus narrative analysis emphasizes taking verbatim transcription of narrative interviews, where it is important to include pauses, filler words, and stray utterances like "um….".

For more information on transcription options, please see our guide on how to transcribe interviews.

Coding in narrative analysis:

There are many methods for coding narrative data. They range from deductive coding where you start with a list of codes, and inductive coding where you do not. You can also learn about many other ways to code in our Essential Guide to Coding Qualitative Data or take our Free Course on Qualitative Data Analysis.

What is narrative research?:

In addition to narrative analysis, you can also practice narrative research, which is a type of study that seeks to understand and encapsulate the human experience by using in depth methods to explore the meanings associated to people's lived experiences. You can utilize narrative research design to learn about these concepts. Narrative analysis can be used in narrative research as well as other approaches such as grounded theory, action research, ethnology and more.

4.6 REFERENCES

1. Shaughnessy, J. J., Zechmeister, E. B. &Zechmeister, J. (2012). *Research methods in psychology*. (9th ed..). NY: McGraw Hill.

2. Elmes, D. G. (2011). *Research Methods in Psychology* (9thed.). Wadsworth Publishing.

3. Goodwin, J. (2009). *Research in Psychology: Methods in Design* (6thed.). Wiley.

4. McBurney, D. H. (2009). *Research methods*. (8th Ed.). Wadsworth Publishing.

5. Forrester, M. A. (2010). *Doing Qualitative Research in Psychology: A Practical Guide*. Sage.

*****