Research Methods

FOR BUSINESS AND MARKETING

George Self

EDITION 3 | JUNE 2021

George Self: Research Methods For Business and Marketing, June 2021

Foreword

I have taught BASV 316, Introductory Methods of Analysis, online for the University of Arizona in Sierra Vista since 2010 and enjoy working with students on research methodology. I wanted a textbook that presented research methods that students could use for their research projects.

I searched for an excellent open educational resource book since those are available free of charge and can be modified to meet the objectives of my class. Bhattacherjee (Bhattacherjee, 2012), Blackstone (Blackstone, 2012), and Price (Price, Chiang, & Jhangiani, 2015) all released books about research in the social sciences that were reasonably close to what I needed, so I combined those books and rewrote them to emphasize business research. The result is provided to my students free of charge.

Three goals shaped the choices made about the topics covered by the text and how those topics are presented.

- The topics must have relevance for business students.
- Both qualitative and quantitative research methods are given roughly equal attention since both types of research are used in business.
- The text is engaging and readable.

This book is published under a Creative Commons Attribution-NonCommercial-ShareAlike license, just like the three books that provided its foundation. I hope that students can use this book to learn about business research, and other instructors can modify and use it for their classes.

I will finish this forward with a note about the word "data." That word is plural – the singular form is "datum." Even though widespread usage tends to treat "data" as singular, it is a plural noun in this book. A phrase like "the data are clear…" may look a bit odd, but it is grammatically correct.

— George Self

Contents

Background

Research methods are grounded in philosophy, statistics, sociology, and many other disciplines. The chapters in this section introduce these background concepts.

1: Introduction

In general, people want to know about things. Most people are curious about the world around them, but business owners are interested in specifically how people can be persuaded to purchase. Understanding how one person can walk past a fruit display without even the slightest thought about a purchase while another cannot seem to walk by without checking out the product is valuable information for the store owner. In general, business owners are eager to know about people and what drives their behavior.

Objectives
• Identify the various sources of knowledge. • Define "science." • Describe the scientific method and relate that to business research. • Identify the three types of science research (exploratory, descriptive, and explanatory). • Describe specific considerations for business research.

Knowing

This book aims to teach students how research can be used to help business owners make good decisions. More specifically, the book examines the ways that researchers come to understand the impetus that drives purchases. Research methods are a systematic process of inquiry designed to learn something of value about a business problem. Before considering research methods, though, it is helpful to contemplate other sources of knowledge.

Different Sources of Knowledge

As an introduction to research, it is helpful to consider familiar sources of knowledge briefly.

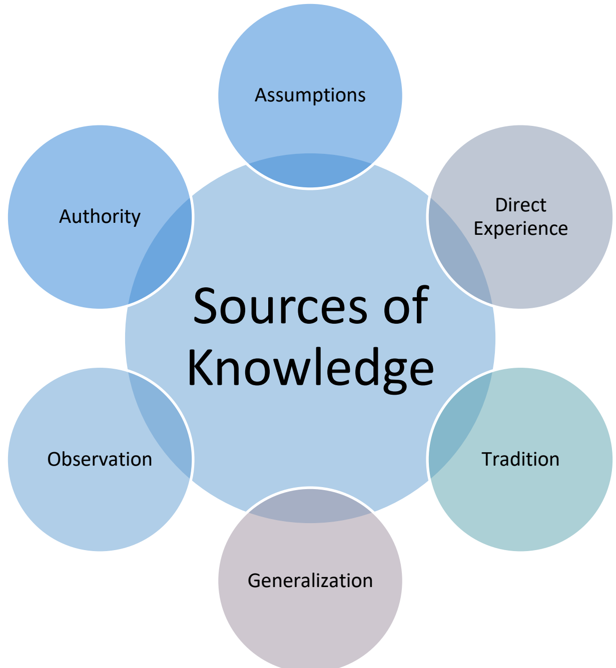


FIGURE 1: SOURCES OF KNOWLEDGE

Assumptions. Many people assume that children without siblings are somewhat spoiled and unpleasant. Many people believe that the social skills of only children will not be as well developed as those of people who were reared with siblings. However, sociological research shows that children who grow up without siblings are about the same as those with siblings when developing good social skills (Bobbitt-Zeher & Downey, 2013). Researchers precisely consider these types of assumptions that "everyone knows" when investigating their worlds. Sometimes the assumptions are correct, and other times not so much.

Direct Experience. One source of knowledge is direct experience. Mark Twain observed that "… the cat that sits down on a hot stove-lid … will never sit down on a hot stove-lid again…" (Twain, 2014). Direct experience may be a source of accurate information, but only for those who experience it. The problem is that the observation is not deliberate or formal; instead, it comes as an accidental by-product of life. Even worse, the lesson learned may be wrong. Without a systematic process for observing and evaluating those observations, any conclusions drawn are suspect.

Tradition. Another source of knowledge is tradition. An urban legend concerns a woman who cut both ends off a ham for years before putting it in the oven (Snopes, 2005). She baked ham that way because that was how her mother did it, so clearly, that was the way it was supposed to be done. Her mother was the authority, after all. After years of tossing cuts of delicious ham into the trash, however, she learned why her mother cut the ends off the ham before cooking; her baking pan was too small for the ham. Tradition may not be the best source of knowledge.

Generalization. Often a broad pattern is observed, and people conclude that the pattern is valid for all instances. Generalizations can be the source of prejudice where the actions of a few bad actors may bias peoples' knowledge of the whole.

Observation. People rely on their informal observations of their worlds. Occasionally, someone will decide to "investigate" something, perhaps an odd sound, and they assume their observations are fact. Unfortunately, these types of observations are not systematic and may easily lead to incorrect conclusions.

Authority. Many people rely on the government, teachers, and other authority figures to dispense knowledge. Unfortunately, authority figures may or may not be a source of accurate knowledge, and it may be challenging to know which information can be trusted.

While people come to know what they know in many ways, some are more reliable than others. Formal research aims to ascertain an accurate answer to the questions people have—to provide a reliable source of knowledge.

What Is Science?

Most research methods used for business and marketing are based on various social science methods, and this section of the book describes how that research is conducted.

FIGURE 2: TWO TYPES OF SCIENCES

Many students assume that "science" is a craft practiced by highly educated experts wearing white lab coats and pouring mysterious, smoking liquids into test tubes. Unfortunately, that is not an accurate definition of "science." Etymologically, the word "science" is derived from the Latin word *Scientia*, which means knowledge. "Science," then, is a systematic and organized body of knowledge acquired using a specific, rigorous method in any field of inquiry. The sciences can be grouped into two broad categories: natural and social. Natural science is the science of naturally occurring objects or phenomena, such as light, objects, matter, earth, celestial bodies, or the human body. Natural sciences are further classified into the physical sciences, earth sciences, life sciences, and others. In contrast, social science studies people or collections of people, like groups, firms, societies, or economies. Social sciences can be classified into disciplines such as psychology (the science of human behaviors), sociology (the science of social groups), and economics (the science of markets and economies).

Their purpose also classifies sciences. Basic sciences, also called pure sciences, explain the most basic objects and forces, relationships between them, and laws governing them. Examples of the basic sciences include physics, mathematics, and biology. Applied sciences, also called practical sciences, apply scientific knowledge from basic sciences to a physical environment. For instance, engineering is an applied science that applies the laws of physics and chemistry to practical applications such as building more durable bridges or fuel-efficient combustion engines, and medicine is an applied science that applies the laws of biology to relieving human ailments.

Scientific knowledge is a generalized body of laws and theories acquired using the scientific method to explain a phenomenon or behavior of interest. Closely related to laws and theories are hypotheses.

Laws are observed patterns of phenomena or behaviors that are based on repeated experimental observations. They are generalized rules that explain observations and are, typically, theories that have been repeatedly tested and believed to be accurate. For example, the three Newtonian Laws of Motion describe 1) what happens when an object is in a state of rest or motion, 2) what force is needed to move a stationary object or stop a moving object, 3) and what happens when two objects collide. Collectively, the three laws constitute the basis of classical mechanics—a theory of moving objects.

Theories are systematic explanations of an underlying phenomenon or behavior. Theories are typically based on hypotheses that have been tested and found to be accurate. However, the testing has been incomplete or lacks the rigor to classify the theory as a law. It is important to note that theories are not “guesses” but are instead the result of experimental observations found to be true in the tested instances. It is also important to note that theories can be falsifiable; that is, there are ways to prove that the theory is not valid. For example:

- the theory of optics explains the properties of light and how it behaves in different media;
- the electromagnetic theory explains the properties of electricity and how to generate it;
- quantum mechanics explains the properties of subatomic particles;
- thermodynamics explains the properties of energy and mechanical work.

None of these theories have been tested in all possible situations, so they are not laws, but there has been enough experimentation to establish a theory.

Hypotheses are a well-guessed explanation of some phenomena or a prediction about what will happen in the future. Hypotheses are generally the beginning of an investigation that will either support or reject the hypotheses. For example, a researcher may hypothesize that products in red boxes sell better than products in blue boxes. An experiment can be set up to test the hypothesis where the same product is sold in two identical boxes, except that one is red and the other blue. If more red boxes were sold at the end of the experiment, the hypothesis is supported, but if there was no difference in sales or the blue boxes sold better, the hypothesis is rejected.

The pure science of economics and its applied science of business includes a body of both laws and theories. For example:

- **Law of Supply and Demand**. While this is often described as a “model,” it is usually categorized as law since it is true in repeated observations. This law states that there is a relationship between a product’s demand and its supply.
- **Law of Diminishing Returns**. This law states that at some point, increasing a single production factor will yield less profit-per-unit produced. In other words, the return on the investment is not worth the increased cost.
- **The Theory of Collective Property Rights** states that groups can manage shared resources, like water. It won the 2009 Nobel Prize for economics.
- **The Theory of Marginalism** attempts to explain the discrepancy in the value of goods by looking at their secondary, or marginal, utility. The diamonds are more expensive than water because of the marginal “satisfaction” of owning diamonds compared to water, even though water is far more utilitarian.

Scientific research aims to discover laws and postulate theories that can explain natural or social phenomena and build scientific knowledge. It is essential to understand that this knowledge may be imperfect or even far from the truth. It is essential to understand that scientific knowledge is based on explaining a particular phenomenon, and some explanations tend to fit the observations better than others. The progress of science is marked by the change over time from more flawed theories to better ones through enhanced observations using more accurate instruments and more informed logical reasoning.

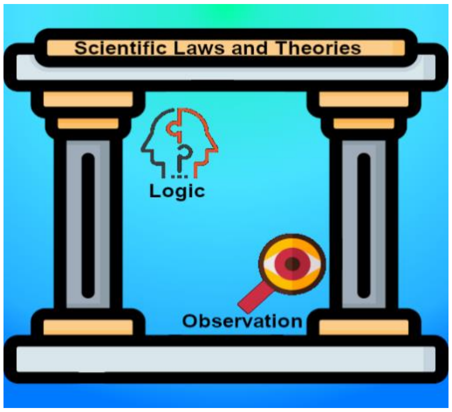


FIGURE 3: PILLARS OF SCIENCE

Scientific laws and theories are based on the two pillars of logic and observation—and nothing else. In science, logic and observation are interrelated, and one cannot exist without the other. Logic provides meaning and significance to observation, while observation helps validate or refine logical concepts. A theory is built from logic, and observation provides evidence to support the theory. Any other means of knowledge acquisition, such as faith or authority, cannot be considered science.

Scientific Research

Scientific research moves between theory and evidence, each reinforcing the other. Theory drives the research of some phenomenon, but evidence uncovered by the researcher further refines the underlying theory. Relying solely on evidence for making inferences while ignoring theory is not scientific research; it is a simple observation, but theory without evidence is no more than guessing. The application of theory and evidence leads to two primary types of scientific research: theoretical and empirical. Theoretical research uses evidence to develop an abstract concept. It is inductive, while empirical research tests theories to see how well they match the evidence and usually is deductive. Hence, inductive research is sometimes called theory-building, while deductive research is called theory-testing. Note that theory testing aims to assess a theory then refine, improve, and extend it.

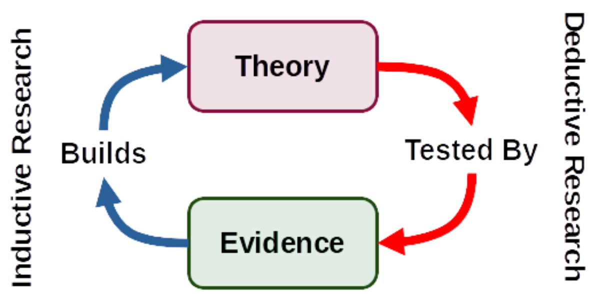


FIGURE 4: SCIENTIFIC RESEARCH MODEL

Figure 4 illustrates how inductive and deductive research are two halves of a continuously iterating research cycle. It is essential to understand that theory-building (inductive research) and theory testing (deductive research) are critical for advancing science[1]. Elegant theories are not valuable if they do not match reality, and mountains of data are useless unless they contribute to meaningful theories. Rather than viewing these two processes in a circular relationship, as shown in Figure 4, perhaps they can be

[1] Inductive and deductive research methods are described more thoroughly in Chapter 2: Foundations.

better viewed as a spiral. Each iteration between theory and data contributes to improved observations of the phenomena and the resulting improved theory. Though both inductive and deductive research is essential for advancing science, inductive (theory-building) research is more valuable when there are few prior theories. At the same time, deductive (theory-testing) research is more productive when there are competing theories.

The chapters in this book are arranged to help facilitate the understanding of both inductive and deductive research.

1. Part 1, Background, contains information pertinent to all research methods.
    a. Chapter 2, Foundations, describes the philosophical underpinnings for research.
    b. Chapter 3, Ethics, takes a deep dive into the ethical principles that govern all research projects.
    c. Chapter 4, Research Design, outlines the process used to design a research project regardless of the type of research involved.
2. Part 2, Quantitative Methods, contains information most appropriate for deductive (theory-testing) research.
    a. Chapter 5, Defining and Measuring Concepts, teaches how concepts can be measured.
    b. Chapter 6, Data, describes the various types of data collected and how those data are analyzed.
    c. Chapter 7, Sampling, provides essential information about how to sample a large population.
    d. Chapter 8, Survey Research, contains background and numerous tips on creating and conducting a survey.
    e. Chapter 9, Experimental Research, shows how experiments are used to verify a theory.
3. Part 3, Qualitative Methods, contains information most appropriate for inductive (theory-building) research.
    a. Chapter 10, Interviews, shows how to conduct in-depth interviews to “dig down” to the root causes of some observation.
    b. Chapter 11, Field Research, investigates the methods used to “go into the field” and seek answers to questions.
    c. Chapter 12, Unobtrusive Research, teaches how to use artifacts like documents and photographs to build theories.

   d. Chapter 13, Interpretive Research, describes how to design research that interprets case studies of populations.
4. Part 3, Mixed Methods, contains only Chapter 14, which shows how to combine both deductive and inductive methods in a single research project.
5. Part 4, Reporting, contains only Chapter 15, which teaches how to report research in live presentations and written papers.

Theory building and theory testing are complex in business and marketing, given the imprecise nature of the theoretical concepts and many unaccounted factors that can influence the phenomenon of interest. It is also challenging to refute theories that do not work. For instance, Karl Marx's theory of communism as an effective economic engine withstood for decades before it was finally discredited as inferior to capitalism in promoting growth. Communist economies like the Soviet Union and China eventually moved toward capitalism with private enterprises focused on maximizing profits. However, the recent collapse of the mortgage and financial industries in the United States demonstrates that capitalism also has its flaws and is not effective in fostering economic growth and social welfare as previously assumed. Unlike theories in the natural sciences, marketing theories are rarely perfect, which provides numerous opportunities for researchers to improve those theories or build alternative theories.

Therefore, conducting scientific research requires two sets of skills, methodological and theoretical, to operate in the empirical and theoretical levels. Methodological skills ("know-how") are relatively standard, invariant across disciplines, and quickly acquired through various educational programs. However, theoretical skills ("know-what") are considerably harder to master, requiring years of observation and reflection, and are tacit skills that cannot be taught but instead learned through experience. The most outstanding scientists in the history of humanity, such as Galileo, Newton, and Einstein, were master theoreticians. They are honored for the theories they postulated that transformed the course of science.

Scientific Method

If science is knowledge acquired through a scientific method, it is essential to define what is meant by the "scientific method." The scientific method refers to a standardized set of techniques for building scientific knowledge, such as making valid observations, interpreting results, and generalizing those results. The scientific method allows researchers to test preexisting theories and prior findings independently and impartially and then subject them to open debate, modifications, or enhancements. The scientific method must satisfy four characteristics:

- **Replicability**. Others could independently replicate a scientific study and obtain similar, if not identical, results.
- **Precision**. Theoretical concepts, which are often hard to measure, must be defined with enough precision that others can use those definitions to measure and test those concepts.
- **Falsifiability**. A theory must be stated in a way that it can be disproven. Theories that cannot be tested or falsified are not scientific, and any knowledge gained is not scientific knowledge. A theory specified in imprecise terms or whose concepts are not measurable cannot be tested and is not scientific. Sigmund Freud’s work on psychoanalysis falls into this category and is therefore not considered a “theory” even though it may have practical utility in treating specific ailments.
- **Parsimony**. When there are multiple explanations of a phenomenon, scientists must always accept the simplest or most logically straightforward explanation. This concept is called parsimony or “Occam’s razor.” Parsimony prevents scientists from pursuing overly complex or outlandish theories with an endless number of concepts and relationships that may explain a little bit of everything but nothing in particular.

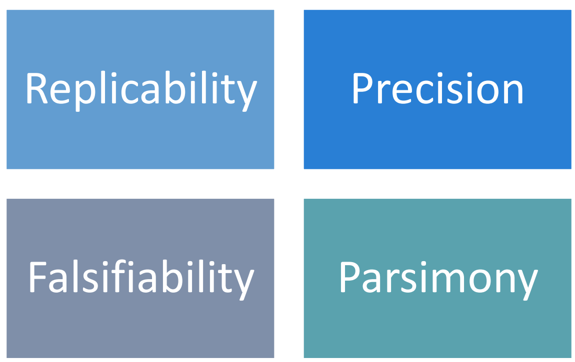


FIGURE 5: THE FOUR CHARACTERISTICS OF THE SCIENTIFIC METHOD

Any branch of inquiry that does not allow the scientific method to test its basic laws or theories cannot be called “science.” For instance, art is not science because artistic ideas (such as the value of perspective) cannot be tested by independent observers using the four characteristics of scientific research. Similarly, music, literature, humanities, and law are also not considered science, even though they are creative and worthwhile endeavors.

As applied to business and marketing, the scientific method includes quantitative research techniques like surveys and experiments that operate primarily at the practical level, analyzing observations to determine if they support the theory. The more challenging part of scientific research is developing theory, which is often a result of qualitative research methods, like interviews and field research. These methods generate new theories by using a "grounded" technique. Both quantitative and qualitative techniques are discussed throughout the book.

Finally, business researchers must bear in mind that the natural sciences are different from the social sciences in several essential respects. The natural sciences are exact, accurate, deterministic, and independent of the person making the observations. For instance, a scientific experiment in physics, like measuring the speed of sound through a particular medium, should always yield the same results, irrespective of the time or place of the experiment. However, the same cannot be said for the social sciences, which tend to be ambiguous. For instance, an economist may want to measure the impact of some factor on a city's economy. Unfortunately, the research outcomes may depend on the background and experience of the researcher, the measurement indices, and the interpretation of those measures. In other words, there is a high degree of variability in all social science research. While natural scientists agree on the speed of light or the earth's gravitational attraction, economists have no agreement on questions like the impact of immigration and how much of a nation's economy should be earmarked for reducing carbon emissions. Researchers in business and marketing must be comfortable handling ambiguity, uncertainty, and error that come with research in such sciences.

Types of Science Research

Depending on the purpose of research, scientific research projects can be grouped into exploratory, descriptive, and explanatory research.

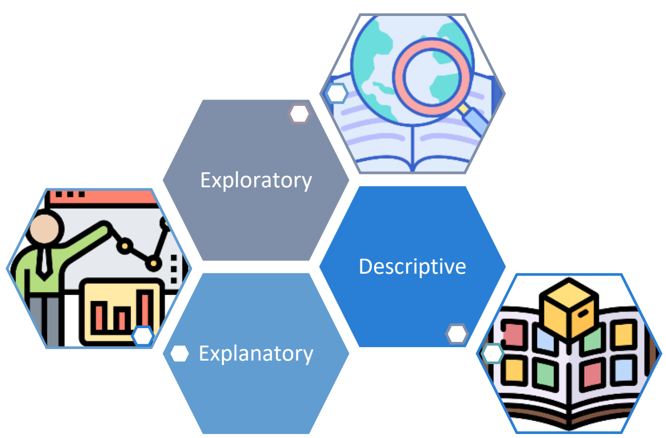


FIGURE 6: THREE TYPES OF SCIENTIFIC RESEARCH

Exploratory

Exploratory research is often conducted in new areas of inquiry, where the goals of the research are:

- to scope out the magnitude or extent of a particular phenomenon, problem, or behavior
- to generate some initial ideas (or "hunches") about that phenomenon
- to test the feasibility of undertaking a more extensive study regarding that phenomenon.

For instance, if the citizens of a country are generally dissatisfied with governmental policies during an economic recession, exploratory research may measure the extent of citizens' dissatisfaction. It would consider how the dissatisfaction is manifested and the presumed causes of such dissatisfaction. Such research may include examining reported economic indicators like gross domestic product (GDP), unemployment, and consumer price index. This research may not lead to a very accurate understanding of the target problem. However, it may be worthwhile in determining the nature and extent of the problem and serve as a valuable precursor to more in-depth research.

Descriptive

Descriptive research is directed at making careful observations and detailed documentation of a phenomenon of interest. Therefore, these observations must be based on the scientific method and are more reliable than casual observations by untrained people. Examples of descriptive research are the tabulation of demographic statistics by the United States Census Bureau that use validated instruments for estimating factors like employment by sector. Suppose changes are made to the measuring instruments. In that case, estimates are provided with and without the changed instrumentation to allow the readers to make a fair before-and-after comparison regarding population or employment trends. Other descriptive research may include chronicling reports of gang activities among adolescent youth, the persistence of religious, cultural, or ethnic practices in select communities, and the role of technologies in the spread of democracy movements.

Explanatory

Explanatory research seeks explanations of observed phenomena, problems, or behaviors. It attempts to "connect the dots" in research by identifying causal factors and outcomes of the target phenomenon. For example, a research project may seek to understand the reasons behind gang violence to prescribe strategies to overcome such societal ailments. Most academic or doctoral research belongs to the explanation category, though some exploratory and descriptive research may also be needed during the initial phases of a research project. Seeking explanations for observed events requires strong theoretical and interpretation skills, along with intuition, insights, and personal experience.

Specific Considerations for Business/Marketing Research

It is essential to keep in mind that business researchers attempt to explain patterns in customers' habits. A pattern does not explain every single person's experience, a fact that is both fascinating and frustrating. Individuals who create a pattern may not be the same over time and may not know one another, but they collectively create a pattern. Those new to business research may find these patterns frustrating because they expect various patterns to describe a group's characteristics, but that often does not translate into an experience. A pattern can exist among a cohort without a specific individual being 100% true to that pattern.

As an example of patterns and their exceptions, consider the impact of social class on peoples' educational attainment. Ellwood & Kane (Ellwood, Kane, & others, 2000) found that the percentage of children who did not receive postsecondary schooling was four times greater among those in the lowest quartile income than the upper quartile. That is, children from low-income families were less likely than high-income children to go to college. These research findings detected patterns in society, but there are certainly many exceptions. Just because a child grows up in a household with little wealth does not keep that child from pursuing a college degree. People who object to research findings tend to cite evidence from their own experience, insisting that no patterns exist. However, the problem with this response is that objecting to a social pattern because it does not match a specific person's experience misses the point about patterns.

Another matter that business researchers must consider is where they stand on the value of basic instead of applied research. In essence, this has to do with questions about whom and for what purpose research is conducted. It is helpful to think of basic and applied research as resting on either end of a continuum. In marketing, basic research studies marketing for marketing's sake—nothing more, nothing less. Sometimes researchers are motivated to conduct research simply because they happen to be interested in a topic, and the goal may be to learn more about a topic. Applied research lies at the other end of the continuum. In marketing, applied research studies marketing for some purpose beyond a researcher's interest in a topic. Applied research is often client-focused, meaning that researchers investigate a question posed by someone other than themselves.

One final consideration for business and marketing researchers is the difference between qualitative research and quantitative research methods. Qualitative methods generally involve words (like letters, memos, or policies) or pictures, and standard methods used include field research, interviews, and focus groups. Quantitative methods, on the other hand, generally involve numbers and standard methods include surveys, content analysis, and experimentation. While qualitative methods aim to understand a relatively small number of cases, quantitative methods offer less depth but more breadth because they typically focus on many cases.

Sometimes these two methods are presented or discussed to suggest they are somehow in opposition to one another. The qualitative/quantitative debate is fueled by researchers who may prefer one approach over another, either because their research questions are better suited to one approach or because they happened to have been trained in one specific method. While these two methodological approaches differ in goals, strengths, and weaknesses, they attempt to answer a researcher's question and are equally viable. This text operates from the perspective that qualitative and quantitative methods are complementary rather than competing, and both will be covered.

A Sampling of Research

Inductive Methods for Operations Management Research

Mario Binder and John S. Edwards researched operations management in the German automotive industry (Binder & Edwards, 2010). They used an inductive approach "…in which systematic data collection is used to develop theories." Their project conducted 31 interviews, followed by a survey with 110 responses and then a focus group. They state that the project was not "…about presenting the subjective experience of experts per se but about abstracting it into theoretical statements in the form of a set of tentative propositions." This project is the essence of inductive research where observations lead to a new proposition or theory. The researchers verified that they used an iterative process where later interviews were informed by data drawn from previous interviews. Again, iteration is a hallmark of inductive research projects.

The observations collected for this project lead to the development of the Collaborative Enterprise Governance concept that "…helps decision makers in selecting the most appropriate governance strategy (i.e., sourcing strategy) for an inter-firm R&D relationship between a buyer and its supplier."

This type of collaboration can lead to faster development time, higher quality, and lower development costs in joint R&D projects.

Doing Well by Doing Good

Roper and Parker reported on research they conducted concerning the relationship between corporate profits and social responsibility (Roper & Parker, 2013). The researchers used a quasi-experimental method to investigate the problem caused by discarded packaging that included the corporate brand. That litter creates a negative perspective in people who see it, and that can have a negative financial impact on the company.

Roper and Parker first postulated these hypotheses:

- Branded litter harms consumers' brand evaluations.
- Consumer exposure versus non-exposure to branded litter causes a lower attitude towards the brand.
- Litter will cause consumers to have a lower intention to purchase or try a brand than those not seeing litter.
- Litter causes consumers to evaluate a brand's positive personality dimensions more negatively.
- Litter causes consumers to evaluate a brand's negative personality dimensions more negatively.
- Litter versus no litter causes consumers to have a lower view of a brand's reputation.

To test these hypotheses, the researchers created a kiosk in a public park for a fictitious company named BigBurger. They hired actors to staff the kiosk and others to purchase burgers. They also hired a film crew to produce two different 90-second newsreel-type reports on this new concept in on-the-go burger sales. One report showed the kiosk area free of any litter, while the second included apparent BigBurger litter on the ground around the kiosk. Both videos featured interviews of the BigBurger CEO and customers. The videos were identical except for the presence or absence of litter.

The researchers then gathered a group of consumers and explained they were participating in evaluating a new brand. One group of consumers was shown the video with no litter, while the other group was shown the video that included the BigBurger litter. The two groups were asked questions about their attitude towards the brand, opinions of the brand personality, and willingness to purchase a BigBurger product.

Roper and Parker found that consumers viewing the video with litter had significantly lower attitudes toward the BigBurger brand and a significantly lower intention to purchase or try the brand. The researchers then asked the consumers what they were willing to pay for a BigBurger to test the potential financial impact of the litter. The mean amount for the group that saw the video without litter was £1.96, but the group that saw the litter was £1.92, a small but significant difference. The researchers conclude that the quasi-experiment supported at least a few of the proposed hypotheses. Companies that do not practice social responsibility will potentially suffer an impact on their bottom line.

Key Takeaways

Chapter 1: Introduction
• There are many different sources of knowledge, and some are more valuable than others for formulating theories and practices. • Science is the discipline of using formalized processes to create theories to explain observed phenomena. • Scientific research is a process to use reproducible methods to create a theory or validate the tenants of an existing theory. • Scientific research can be divided into three types: exploratory, descriptive, and explanatory. • Business research has specific considerations to meet the sometimes-disparate objectives of theory-building and practical application.

2: Foundations

Introduction

Why do some researchers prefer charts and graphs while others seem to like tables full of numbers? Why do some researchers work best with computer tools while others tend to sketch their ideas on paper? Why do calendars and deadlines drive some researchers while others seem to take their time? All these questions are related to the philosophical foundations that form the researcher's analytical perspective. This chapter explores how the researcher's foundations shape methodological choices and the connection between paradigms, theories, and research methods.

OBJECTIVES
• Define paradigm. • Define theory and discuss the process used to build and test a theory. • Discuss the difference between propositions, hypotheses, and models. • Define and contrast inductive and deductive research methods.

Ontology and Epistemology

Two major branches of philosophy, ontology and epistemology, provide the principles of research methods in business, sociology, and psychology. Ontology concerns the nature of reality, and the researcher's ontological position shapes the sorts of research questions posed and how those questions are researched. Ontology posits two fundamental positions:

- **Objectivism**: Things are real and exist regardless of any sort of social activity. This position is often reflected in research about societal organization. Objectivists assume that people may differ in their perception of reality, but there is only one actual reality, and a researcher's job is to discover that reality.
- **Constructivism**: Things do not just exist apart from the society that observes them. This position is often reflected in research about culture and its influence on human activities. Constructivists assume that reality is shaped individually and that a researcher's job is to understand others' views of reality.

Like ontology, epistemology has to do with knowledge. Rather than dealing with questions about what *is*, epistemology deals with questions of how we know. Four main branches of epistemology are frequently encountered in business research, and the researcher's beliefs concerning these branches will shape the research design.

- **Pragmatism** accepts both personal experiences and measured data as sources of knowledge. These researchers will usually design applied research projects that use different perspectives to help answer a question.
- **Positivism** relies only on findings gained through measurement. These researchers tend to focus on causality and try to reduce phenomena to their simplest elements.
- **Realism** relies on observations rather than precise measures to provide credible facts and data. These researchers would use tools like structured interviews to gain an understanding of a phenomenon.
- **Interpretivism** uses subjective explanations of social phenomena. These researchers use tools like ethnographic studies to attempt to understand an entire social structure.

In their seminal book Sociological Paradigms and Organizational Analysis, Burrell and Morgan (Burrell & Morgan, 2017) suggested that epistemology shapes a researcher's approach to a project while ontology shapes the researcher's interpretation of the findings. Using these two sets of assumptions, Burrell and Morgan categorized research as in Figure 7.

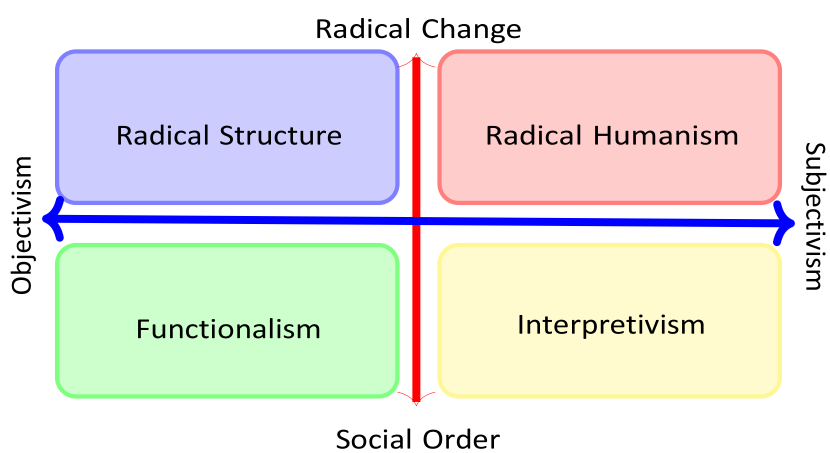


FIGURE 7: RESEARCH PARADIGM

Functionalism is the mindset adopted by researchers who

- Ontology: view the world as orderly and consisting of patterns of ordered events or behaviors.

- Epistemology: the best way to study the world is to use an objective approach independent of the person conducting the observation using standardized collection tools like surveys.

Interpretivism is the mindset adopted by researchers who

- Ontology: view the world as orderly and consisting of patterns of ordered events or behaviors.
- Epistemology: the best way to study the world is through the subjective interpretation of participants using techniques like interviewing participants and then reconciling differences using their perspectives.

Radical structure is the mindset adopted by researchers who

- Ontology: view the world as constantly changing, often radically, with few unvarying patterns or behaviors.
- Epistemology: the best way to study the world is to use an objective approach independent of the person conducting the observation using standardized collection tools like surveys.

Radical humanism is the mindset adopted by researchers who

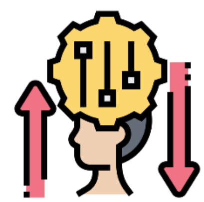
- Ontology: view the world as constantly changing, often radically, with few unvarying patterns or behaviors.
- Epistemology: the best way to study the world is through the subjective interpretation of participants using techniques like interviewing participants and then reconciling differences using their perspectives.

To date, most business research has emulated the natural sciences and adopted functionalist techniques. Thus, researchers tend to believe that social patterns can be understood in terms of their functional components. Hence, they study those components in detail using objective techniques like surveys and experimental research. However, a small but growing number of researchers adopt interpretive approaches and understand social order using interviews and ethnographic studies. Radical structuralism and radical humanism represent a negligible proportion of business research. However, social and organizational phenomena generally consist of elements of both order and change. For instance, organizational success depends on formalized business processes, work procedures, and job responsibilities while being simultaneously constrained by a constantly changing mix of competitors, competing products, suppliers, and customer base in the business environment. Therefore, obtaining a holistic understanding of phenomena like the success of some businesses and the failure of others may require a multi-modal approach.

Paradigms and Theories

The terms paradigm and theory are often used interchangeably in business and marketing research, although experts disagree about whether these are identical or distinct concepts. This text makes a slight distinction between the two ideas because thinking about each concept as analytically distinct provides a valuable framework for understanding the connections between research methods and scientific ways of thinking.

Paradigm



The researcher's frames of reference, or belief systems, form a paradigm. Thus, if a researcher is, generally, functionalist in outlook, that would be the paradigm used to design and conduct research projects. Paradigms are usually quite complex and include upbringing, family influence, educational background, societal norms, and many other factors. Paradigms are often hard to recognize because they are implicit, assumed, and taken for granted. However, identifying paradigms is key to making sense of and reconciling differences in peoples' perceptions of the same social phenomenon.

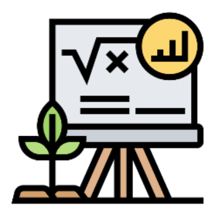
Paradigms are like "colored glasses" that govern how people structure their thoughts about the world. As one other example, imagine that a particular type of technology was successfully implemented in one organization but failed miserably in another. A researcher using a "rational lens" will look for rational explanations of the problem, such as inadequate technology or a poor fit between technology and the task context where it is being utilized. Another research looking at the same problem through a "social lens" may seek social deficiencies such as inadequate user training or management support. Another researcher seeing it through a "political lens" will look for organizational politics that may subvert the technology implementation process. Hence, subconscious paradigms often constrain the concepts that researchers attempt to measure and their subsequent interpretations of those measures. However, likely, all of the above paradigms are at least partially correct, and a fuller understanding of the problem may require an application of multiple paradigms. Two paradigms are commonly found in business research:

Positivism. This paradigm is the framework that usually comes to mind when people think about research[2]. Positivism is guided by the principles of objectivity, knowability, and deductive logic[3]. The positivist framework operates from the assumption that society can and should be studied empirically and scientifically. Positivism also calls for value-free research, where researchers attempt to abandon their own biases and values in a quest for objective, empirical, and knowable truth. Positivism is based on French philosopher Auguste Comte (1798 - 1857) and was the dominant scientific paradigm until the mid-20th century. Unfortunately, positivism eventually evolved to empiricism or blind faith in observed data and rejected any attempt to extend or reason beyond observable facts. Since human thoughts and emotions could not be directly measured, they were not considered legitimate topics for scientific research.

Postmodernism. Frustrations with the strictly empirical nature of positivist philosophy led to postmodernism during the mid-late 20th century. Postmodernism argues that one can make inferences about a phenomenon by combining empirical observations with logical reasoning. Postmodernists view science as problematic since it is based on many contingencies and often seek to explore these contingencies to understand reality better. The postmodernist camp has further fragmented into two camps. Subjectivists view the world as a subjective construction of our minds rather than as an objective reality. Simultaneously, critical realists believe that there is an external reality independent of a person's thinking, but we can never know such truth with any degree of certainty.

Theory

DEFINITION



Theories are explanations of natural or social behavior, event, or phenomenon. More formally, a scientific theory is a system of constructs (concepts) and propositions (relationships between those constructs) that collectively presents a logical, systematic, and coherent explanation of a phenomenon of interest within some assumptions and boundary conditions (Bacharach, 1989). It is important to note that people not familiar with scientific research often view a theory as speculation, a "guess," and statements like "it is only a theory" are common. However, a scientific theory is well-researched and based on repeated observations of some phenomenon. For example, plate tectonics is a theory that indicates that the continents are slowly moving across the earth's surface. This theory is well-established and based on research spanning

[2] Positivism was also discussed as one of the main branches of epistemology but since it is so common in the research community it is also recognized as a paradigm.

[3] Deductive and inductive logic methods are discussed later in this chapter.

decades of observations, not just some sort of idle speculation. A good scientific theory should be well supported using observed facts and have practical value. Karl Marx once wrote, “Practice without theory is blind. Theory without practice is sterile. Theory becomes a material force as soon as it is absorbed by the masses.” (Marx, 1971) Hence, both theory and practice are essential aspects of research.

Theories should explain *why* things happen rather than describe or predict. Note that events can be predicted without necessarily explaining why such events occur. For instance, market analysts predict fluctuations in the stock market based on market announcements, the earnings report of significant companies, and new data from the Federal Reserve and other agencies based on previously observed correlations. Prediction only requires correlation while explanation requires causation, and causation entails three conditions:

- Correlations between two constructs
- Temporal precedence (the cause must precede the effect in time)
- Rejection of alternative hypotheses (through testing)

It is also essential to understand what theory is not. It is not data, facts, typologies, taxonomies, or empirical findings. A collection of facts is not a theory, just as a pile of stones is not a house. Likewise, a group of constructs (e.g., a typology of constructs) is not a theory because theories must go well beyond constructs to include propositions, explanations, and boundary conditions. Data, facts, and findings operate at the empirical or observational level, while theories operate at a conceptual level and are based on logic rather than observations.

There are many benefits to using theories in research. First, theories provide the underlying logic explaining phenomena by describing the key drivers, outcomes, and underlying processes responsible for that phenomenon. Second, theories aid in sense-making by synthesizing prior findings within a framework. Third, theories guide future research by helping identify constructs and relationships that are worthy of further investigation. Fourth, theories contribute to the cumulative body of knowledge and bridge gaps between other theories by reevaluating those theories in a new light.

However, theories can also have their share of limitations. As simplified explanations of reality, theories may not always provide adequate explanations of the phenomena of interest. Theories are designed to be economical and straightforward explanations, while reality is usually significantly more complex. Furthermore, theories may impose blinders or limit researchers’ “range of vision,” causing them to miss out on essential concepts that are not defined by the theory.

BUILDING BLOCKS OF A THEORY

David Whetten (Whetten, 1989) suggests that there are four building blocks of a theory.

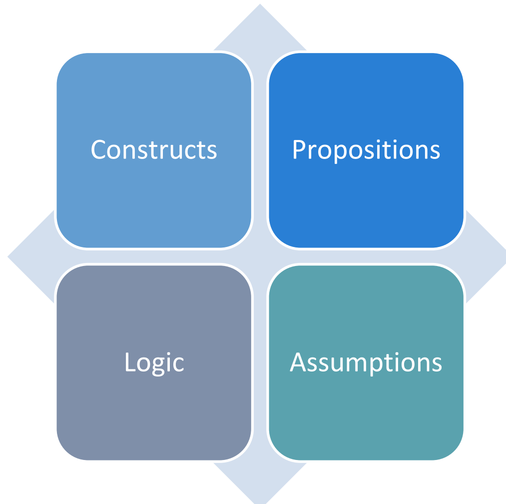


FIGURE 8: BUILDING BLOCKS OF A THEORY

1. **Constructs** capture the “what” of theories (i.e., what concepts are essential for explaining a phenomenon). They are abstract concepts specified at a high level of abstraction explicitly chosen to explain the phenomenon of interest. Constructs may be unidimensional (i.e., embody a single concept), such as weight or age, or multi-dimensional (i.e., embody multiple underlying concepts), such as personality or culture. While some constructs, such as age, education, and firm size, are easy to understand, others, such as creativity, prejudice, and organizational agility, may be more complex and abstruse. Still, others, such as trust, attitude, and learning, may represent temporal tendencies rather than steady states. Nevertheless, all constructs must have a clear and unambiguous operational definition that should specify precisely how the construct will be measured and at what level of analysis (like individual, group, or organizational).

   Measurable representations of abstract constructs are called variables. For instance, intelligence quotient (IQ) is a variable purported to measure an abstract construct called intelligence. As noted earlier, scientific research proceeds along two planes: a theoretical plane and an empirical plane. Constructs are conceptualized at the theoretical plane, while variables are operationalized and measured at the empirical (observational) plane. Furthermore, variables may be independent, dependent, mediating, or moderating.

2. **Propositions** capture the "how" (i.e., how are these concepts related to each other). They are associations postulated between constructs based on deductive logic. Propositions are stated in declarative form and should ideally indicate a cause-effect relationship (e.g., if X occurs, then Y will follow). Note that propositions may be conjectural but must be testable and rejected if empirical observations do not support them. However, like constructs, propositions are stated at the theoretical level, and they can only be tested by examining the corresponding relationship between measurable variables of those constructs. A hypothesis is the empirical formulation of a proposition that is stated as the relationship between variables.
3. **Logic** represents the "why" (i.e., why are these concepts related). Logic provides the basis for justifying the propositions as postulated. Logic acts like a "glue" that connects the theoretical constructs and provides meaning and relevance to the relationships between these constructs. Logic also represents the "explanation" that lies at the core of a theory. Without logic, propositions will be ad hoc, arbitrary, and meaningless and cannot be tied into a cohesive "system of propositions" that is the heart of any theory.
4. **Assumptions** (sometimes called "boundary conditions") examine the "who, when, and where" (i.e., under what circumstances will these concepts and relationships work). All theories are constrained by assumptions about values, time, space, and boundary conditions that govern where the theory can and cannot be applied. For example, many economic theories assume that human beings are rational and employ utility maximization based on cost and benefit expectations as a way of understanding human behavior. In contrast, political science theories assume that people are more political than rational and try to position themselves in their professional or personal environment in a way that maximizes their power and control over others. Given the nature of their underlying assumptions, economic and political theories are not directly comparable. Researchers should not use economic theories if their objective is to understand the power structure or its evolution in an organization. Likewise, theories may have implicit cultural assumptions (whether they apply to individuals or cultures), temporal assumptions (whether they apply to early stages or later stages of human behavior), and spatial assumptions (whether they apply to certain localities but not to others). If a theory is to be used appropriately and tested, all its implicit assumptions that form the boundaries of that theory must be adequately understood. Unfortunately, theorists rarely state their implicit assumptions clearly, which leads to frequent misapplications of theories to problem situations in research.

VARIABLES

A term frequently associated with, and sometimes used interchangeably with, a construct is variable. A variable is a quantity that can vary (from low to high), in contrast to constants that do not vary (remain constant). However, in scientific research, a variable is a measurable representation of an abstract construct. As abstract entities, constructs are not directly measurable, and hence, we look for proxy measures called variables. For instance, a person's intelligence is often measured as his or her IQ (intelligence quotient) score, an index generated from an analytical and pattern-matching test administered to people. In this case, intelligence is a construct (a concept), and the IQ score is a variable that measures that construct. Whether IQ scores truly measure one's intelligence is anyone's guess, though many believe that they do. Depending on how well it measures intelligence, the IQ score may be a good or a poor measure of the intelligence construct.

Depending on their intended use, variables may be classified as:

- **Independent**: Explain other variables.
- **Dependent**: Explained by other variables.
- **Moderating**: Influence the relationship between the independent and dependent variables.
- **Mediating**: Are explained by independent variables while also explain dependent variables.
- **Control**: Not pertinent to explaining a dependent variable but must be controlled.

To understand the differences between these different variable types, consider the example shown in Figure 9.

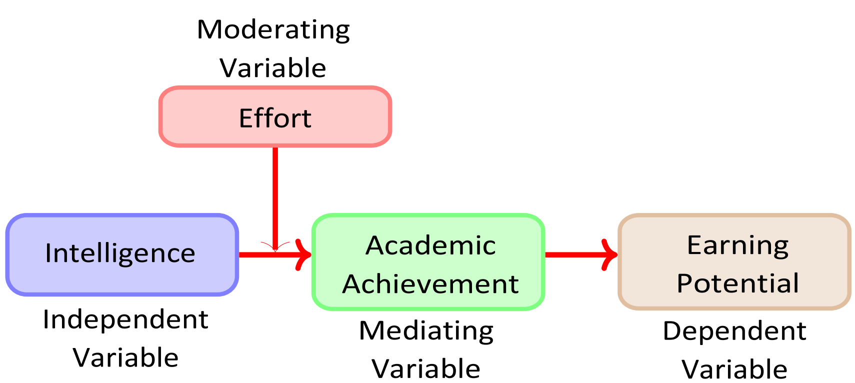


FIGURE 9: TYPES OF VARIABLES

In Figure 9, a person's intelligence is an independent variable that influences academic achievement, which, in turn, influences earning potential, making academic achievement a mediating variable. Effort independently influences academic achievement along with intelligence, so it is considered a moderating variable. To apply measures to this system, if a researcher believes that intelligence influences students' academic achievement, then a measure of intelligence, such as an IQ score, would be the independent variable. In contrast, a measure of academic success, grade point average, would be the dependent variable. Suppose the effect of intelligence on academic achievement is also dependent on the students' effort. In that case, "effort" becomes a moderating variable, and it could be measured by the number of hours spent studying. It would be reasonable to also view effort as the independent variable and intelligence as a moderating variable. The dependent variable, earning potential, could be measured by the salary offered to a graduate. In an experiment, a variable like socioeconomic status may be a control variable, so the researcher would want to ensure all participants were in a similar socioeconomic status to eliminate that variable's effect on earnings potential.

ATTRIBUTES OF A GOOD THEORY

Theories are simplified and are often only partial explanations of complex social reality. There can be good explanations or inadequate explanations, and consequently, there can be good theories or flawed theories. How can the "goodness" of a theory be evaluated? Different criteria have been proposed by different researchers, the more important of which are listed below.

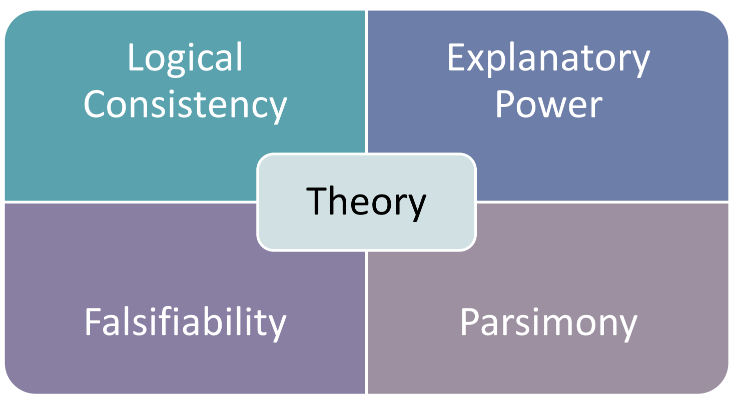


FIGURE 10: ATTRIBUTES OF A GOOD THEORY

Logical consistency: Are the theoretical constructs, propositions, boundary conditions, and assumptions logically consistent with each other? Suppose some of these "building blocks" of a theory are inconsistent with each other (e. g., a theory assumes rationality, but some constructs represent non-rational concepts). In that case, the theory is flawed.

Explanatory power: How much does a given theory explain (or predict) reality? Good theories explain the target phenomenon better than flawed theories, as often measured by the "variance explained value" (statistically known as σ^2) in regression equations.

Falsifiability: British philosopher Karl Popper stated in the 1940s that for theories to be valid, there must be a way to prove that they are false. Falsifiability ensures that the theory can be disproved if empirical data do not match theoretical propositions. This property allows for empirical testing by researchers. In other words, theories cannot be theories unless they can be empirically testable. Tautological statements, such as "a day with high temperatures is a hot day," are not empirically testable because a hot day is defined (and measured) as a day with high temperatures. Hence, such statements cannot be viewed as a theoretical proposition. Falsifiability requires the presence of rival explanations; it ensures that the constructs are adequately measurable, and so forth. However, stating that a theory is falsifiable is not the same as stating it should be falsified. If a theory is indeed falsified based on empirical evidence, then it was probably a flawed theory from the start.

Parsimony: Parsimony examines how much of a phenomenon is explained with few variables. The concept is attributed to 14th-century English logician Father William of Ockham (and is often called "Ockham's razor" or "Occam's razor"). The concept is that among competing explanations that sufficiently explain the observed evidence, the most straightforward theory (i.e., one that uses the smallest number of variables or makes the fewest assumptions) is the best. Explanation of a complex social phenomenon can consistently be increased by adding more and more constructs. However, such an approach defeats the purpose of having a theory intended to be "simplified" and generalizable explanations of reality. Parsimony relates to the degrees of freedom in each theory. Parsimonious theories have higher degrees of freedom, which allow them to be more easily generalized to other contexts, settings, and populations.

APPROACHES TO THEORIZING

How do researchers build theories? Steinfeld and Fulk (Steinfield & Fulk, 1990) recommend four approaches. The first approach is to build theories inductively based on observed patterns of events or behaviors. Such an approach is often called a grounded theory because the theory is grounded in empirical observations. This technique is heavily dependent on the observational and interpretive abilities of the researcher, and the resulting theory may be subjective and not confirmable. Furthermore, observing specific patterns of events will not necessarily make a theory unless the researcher can provide consistent explanations for the observed patterns.

The second approach to theory building is to conduct a bottom-up conceptual analysis to identify different sets of predictors relevant to the phenomenon of interest using a predefined framework. One such framework may be a simple input-process-output framework. In this case, the researcher may look for different inputs, such as individual or organizational, the output, and describe the underlying processes that link these factors to the target phenomenon. This approach is also inductive that relies heavily on the inductive abilities of the researcher, and interpretation may be biased by the researcher's prior knowledge of the phenomenon being studied.

The third approach to theorizing is extending or modifying existing theories to explain a new context, such as extending individual learning theories to explain organizational learning. While making such an extension, specific concepts, propositions, or boundary conditions of the old theory may be retained, and others modified to fit the new context. This deductive approach leverages the rich inventory of social science theories developed by prior theoreticians and is an efficient way of building new theories by building on existing ones.

The fourth approach is to apply existing theories in entirely new contexts by drawing upon the structural similarities between the two contexts. This approach relies on reasoning by analogy and is probably the most creative way of theorizing using a deductive approach. For instance, Markus (Markus, 1987) used analogical similarities between a nuclear explosion and uncontrolled growth of networks or network-based businesses to propose a critical mass theory of network growth. Just as a nuclear explosion requires a critical mass of radioactive material to sustain a nuclear explosion, Markus suggested that a network requires a critical mass of users to sustain its growth. Without such critical mass, users may leave the network, causing an eventual demise of the network.

Propositions and Hypotheses

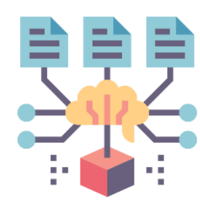
In seeking explanations to an observed phenomenon, it is not adequate to identify critical constructs underlying that phenomenon; it is also essential to state the relationships between constructs. Such patterns of relationships are called propositions. A proposition, thus, conjectures a relationship between constructs and is stated in a declarative form. An example of a proposition is: "An increase in student intelligence leads to an increase in academic achievement." (This proposition was illustrated in Figure 9: Types of Variables.) A proposition does not need to be accurate, but it must be testable to determine its truth. Propositions are generally derived from either logic (deduction) or observation (induction).

Because propositions are associations between abstract constructs, they cannot be tested directly. Instead, they are tested indirectly by examining the relationship between those constructs' corresponding measures (variables). The formulation of a proposition is called a hypothesis. Since IQ scores and grade point averages are operational measures of intelligence and academic achievement, respectively, the proposition stated above can be specified in the form of a hypothesis: "An increase in students' IQ score leads to an increase in their grade point average." Propositions are generated from theory, while hypotheses are generated from empirical evidence. Hence, hypotheses are testable using observed data and may be rejected if the data do not support them.

Hypotheses are said to be either strong or weak. "Students' IQ scores are related to their academic achievement" is an example of a weak hypothesis since it indicates neither the direction of the hypothesis (positive or negative) nor its causality (whether intelligence causes academic achievement vice versa). A more robust hypothesis is that "students' IQ scores are positively related to their academic achievement," which indicates the direction but not the causality. A still better hypothesis is "students' IQ scores have positive effects on their academic achievement," which specifies both the directionality and the causality (i.e., intelligence causes academic achievement, and not the reverse).

Also, note that hypotheses should specify independent and dependent variables. In the hypothesis, "students' IQ scores have positive effects on their academic achievement," it is clear that intelligence is the independent variable (the "cause") and academic achievement is the dependent variable (the "effect"). Further, it is also clear that this hypothesis can be evaluated as either true (if higher intelligence leads to higher academic achievement) or false (if higher intelligence has no effect on or leads to lower academic achievement). A statement such as "all students can achieve academic success" is not a hypothesis because it does not specify independent and dependent variables, nor does it specify a directional relationship.

Models



A term often used in conjunction with theory is model. A model represents all or part of a system constructed to study that system (e.g., how the system works or is triggered). While a theory tries to explain a phenomenon, a model tries to represent a phenomenon understandably. Models are often used to make crucial decisions that are based on a given set of inputs. For instance, marketing managers may use models to decide how much money to spend on advertising for different product lines based on the prior year's advertising expenses, sales, market growth, and competing products. Likewise, weather forecasters use models to predict future weather patterns based on wind speeds, air pressure, temperature, and humidity. While these models are helpful, they do not explain the theory behind advertising budgets or weather forecasting.

Models may be of different kinds, such as mathematical models, network models, and path models. Models can also be descriptive, predictive, or normative. Descriptive models are frequently used for representing complex systems, for visualizing variables and relationships in such systems. Predictive models (e.g., a regression model) allow the forecast of future events. Normative models are used to guide activities along commonly accepted norms or practices. Models may also be static if it represents the state of a system at one point in time or dynamic if it represents a system's evolution over time.

The process of theory or model development may involve inductive and deductive reasoning, as shown in Figure 11: Model Building. Induction occurs when we observe a fact and ask, "Why is this happening?" while deduction occurs when we have a theory and ask, "Do observable facts support this?" Both induction and deduction lead to preliminary conclusions that are then tested to develop a final model of the phenomenon. Researchers must move back and forth between inductive and deductive reasoning if they post extensions or modifications to a given model or theory, which is the essence of scientific research. This process is illustrated in Figure 11.

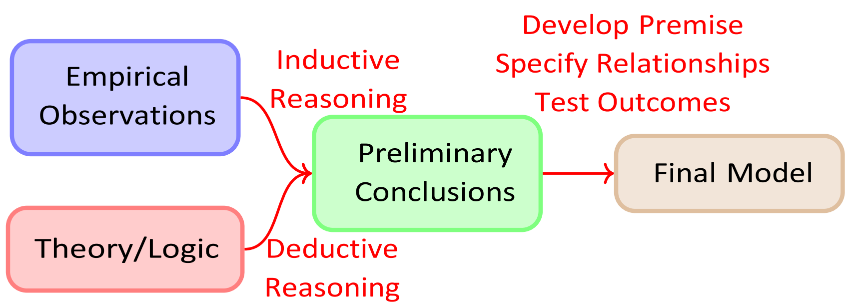


FIGURE 11: MODEL BUILDING

Inductive or Deductive Approaches

Theories are used to structure research while the research structures theory. The reciprocal relationship between theory and research becomes more evident to researchers as they determine whether an inductive or deductive approach is best. Often, researchers find that a single approach is not ideal, and projects iterate over many cycles of inductive/deductive approaches. It is common for a researcher to start with an inductive research approach, postulate a new theory, and then switch to a deductive research approach to test that theory. Later the researcher may return to an inductive approach to expand and refine the theory, followed by another deductive method probe to test the new theory.

Inductive Research

In inductive research, a researcher begins by collecting data that are relevant to the topic of interest. Once data have been collected, the researcher will start to look for trends or correlations and then develop the theory that explains those patterns. Thus, an inductive approach moves from data to theory or specific instances to general explanations and is often referred to as a “bottom-up” approach. Figure 12 broadly outlines the process used with an inductive research approach.

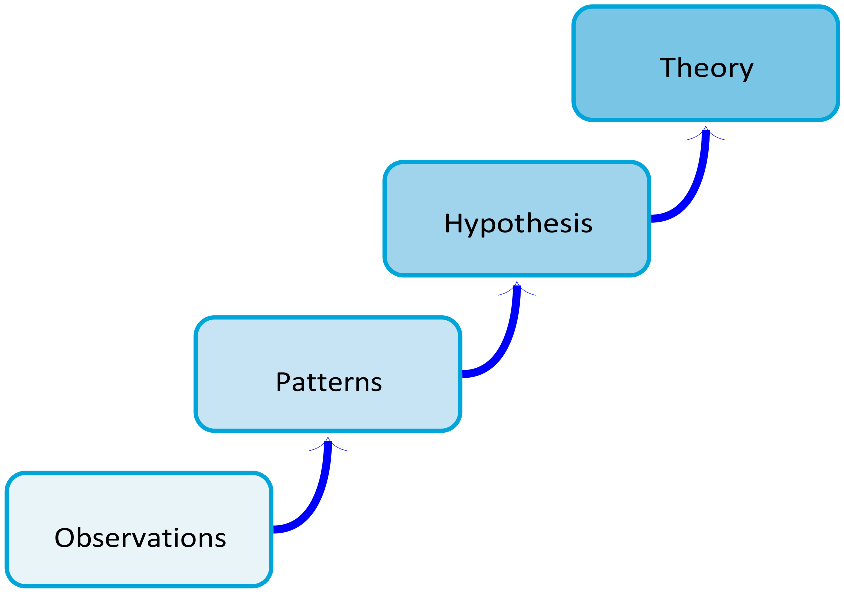


FIGURE 12: INDUCTIVE REASONING

Inductive methods are commonly applied to qualitative research projects and are frequently criticized for being too subjective. Generally, the goal is to understand the dynamics of business practices and use that understanding to draw general conclusions applicable to other businesses. Many qualitative research projects generate grounded theory,[4] where the researcher starts with no preconceived notions and generates a new theory from the data analysis.

Following are three examples of inductive methods research.

- Bansal, Pratima, and Roth (Bansal & Roth, 2000) conducted a study concerning why corporations “go green.” They collected data from 53 firms in the United Kingdom and Japan and analyzed them to formulate a theory.
- Sharma (Sharma, 2000) sent surveys to 3−5 senior managers of 110 Canadian oil and gas companies with annual sales of more than $20 million. The surveys were analyzed. The

[4] Grounded theory was first discussed in many journal articles and books in the 1960s. See, for example, Strauss, Anselm, and Juliet Corbin (Strauss & Corbin, Grounded theory methodology, 1994). Grounded theory is more thoroughly covered in Chapter 14.

researcher concluded that managers of these companies must be influenced to embrace environmental issues as a corporate goal, but that must be done within the corporate structure.

- Sia and Gopa (Sia & Bhardwaj, 2009) used an inductive method to analyze how the "psychological contract" between a corporation and its employees affects diversity. A psychological contract is what the employee "…believes he or she has agreed to…" rather than what is in the employment contract. They administered two different surveys to 207 managers of public sector units in Orrisa, India. They found that certain minority groups tended to "…protect each other when required, particularly during the time of crisis." However, members of the dominant group did not engage in that type of behavior, leading to "…a feeling of non-inclusiveness."

In addition to the research studies discussed above, several papers have been published by various journals encouraging inductive research methods, especially in analyzing case studies. For example, Eisenhardt and Graebner (Eisenhardt & Graebner, Theory building from cases: Opportunities and challenges, 2007) published an article that suggested generating theory from multiple case studies and encouraged management researchers to consider the role of theory-generation in their case studies.

Deductive Research

In deductive research, a researcher begins with a theory of interest and then collects data to test that theory. Thus, a deductive approach moves from a general explanation of some phenomenon to specific instances that prove or disprove the phenomenon and is often referred to as a "top-down" approach. Figure 13 broadly outlines the process used with a deductive research approach.

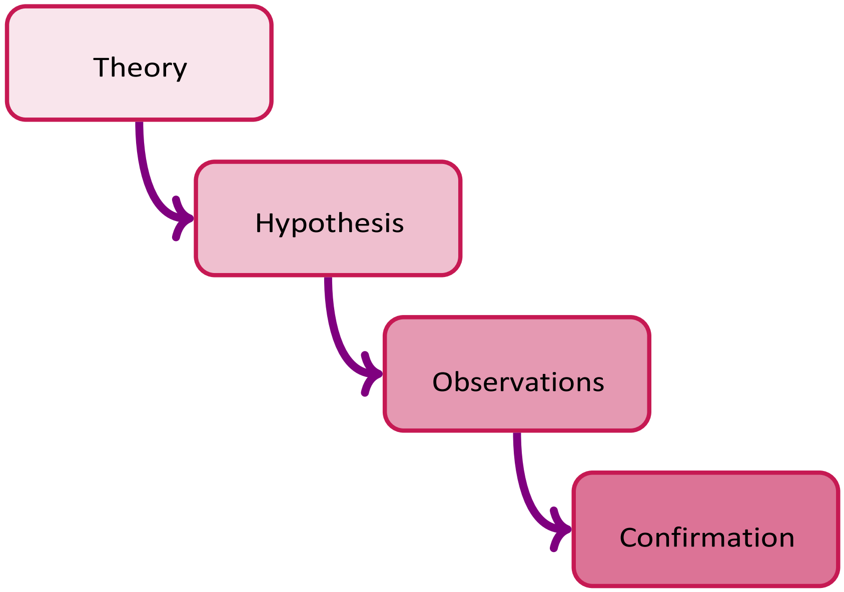


Figure 13: Deductive Reasoning

Deductive methods are commonly applied to quantitative research projects and are often considered the "gold standard" of methods, especially among researchers in the natural sciences. Generally, the goal is to test existing theories to see if they are valid in cases that have not been previously considered. Following are a few example studies that use a deductive approach.

- Parboteeah, Paik, and Cullen (Parboteeah, Paik, & Cullen, 2009) studied the influence of religion on the workplace. They used data from more than 44 thousand individuals in 39 countries to determine if Buddhism, Christianity, Hinduism, and Islam influenced extrinsic and intrinsic work values. They found that the results "…generally support the posited hypotheses, confirming that religion is positively related to work values." Because the study began with hypotheses and tested those hypotheses against gathered data, this is a deductive methodology.
- Hackman and Oldham (Hackman & Oldham, 1976) used existing theory to develop a model to predict the conditions that will motivate employees to perform effectively on their jobs. They tested 658 employees who worked at 62 different jobs in seven organizations and found that the results support the validity of their model.
- Delaney and Huselid (Delaney & Huselid, 1996) investigated the relationship between human resource management and perceptions of organizational performance. They came up with two hypotheses and then gathered data to test those hypotheses. The result of their study is that

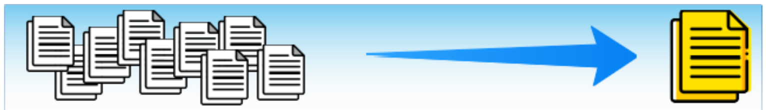
positive human resources practices (like training programs) have a positive correlation with the perception of organizational performance.

Complementary Approaches

While inductive and deductive approaches to research seem pretty different, they are complementary because one approach creates theories, and the other tests theories. In some cases, researchers plan for their research to include multiple phases, one inductive and the other deductive. In other cases, a researcher might begin a study with the plan to conduct either inductive or deductive research but then discover that the other approach is needed to develop a complete picture.

One such example is a research project completed by Lawrence Sherman and Richard Berk (Sherman & Berk, 1984). They experimented with testing two competing theories of the effects of punishment on deterring domestic violence. Specifically, Sherman and Berk hypothesized that deterrence theory would better explain the effects of arresting accused batterers than labeling theory. Deterrence theory predicts that arresting an accused spouse batterer will reduce future incidents of violence, while labeling theory predicts that arresting accused spouse batterers will increase future incidents.

Sherman and Berk found, after experimenting with the help of local police in one city, that arrest did deter future incidents of violence, thus supporting their hypothesis that deterrence theory would better predict the effect of the arrest. After conducting this research, they and other researchers conducted similar experiments in six additional cities, but the results from these follow-up studies were mixed. In some cases, arrest deterred future incidents of violence, but in other cases, it did not. These additional studies left the researchers with new data that they needed to explain. The researchers next took an inductive approach to make sense of their latest empirical observations. The new studies revealed that arrest seemed to have a deterrent effect for married and employed people but led to increased offenses for unmarried and unemployed people. In the end, the researchers turned to control theory and predicted that having some stake in conformity through the social ties provided by marriage and employment would deter future violence. This research project is an excellent example of one that evolves through several iterations of induction/deduction.



A Sampling of Research

Thriving for Small Business Owners

Jessica Macera investigated small businesses and used grounded theory techniques to determine what elements help them thrive (Macera, 2016). Macera conducted open-ended conversational interviews with 15 small business owners. After the first three interviews, she developed categories of elements that the owners identified as necessary for thriving. After eight interviews, "…ideas were strongly developed," and saturation, where no new information was formulated, was reached after 12 interviews. Her interviews were recorded, and she relied on both the recordings and her field notes to code the collected data into coherent results. She found the single most crucial factor to explain small business owners thriving is a personal investment. However, she divided this factor into 21 characteristics like constantly moving to the next step, celebration, confidence, and goal setting. She next grouped these characteristics into four categories: mastery, motivators, personal characteristics, and the role of others. She eventually developed a theory of thriving which includes "…alternating periods of personal investment and capitalization on opportunities, leading to brief periods of intense thriving and subsequently higher levels of sustained thriving."

Key Takeaways

CHAPTER 2: FOUNDATIONS

- A paradigm is the researcher's frames of reference or belief systems. Researchers should be aware of their paradigms and ensure that they will not introduce bias into the research project.
- Theories are explanations of natural or social behavior, event, or phenomenon. It is a system of constructs and propositions that describe the relationships between those constructs. Theories are built by defining the constructs being evaluated and then determining the propositions that link the constructs.
- A proposition is a relationship between constructs. A proposition does not need to be accurate, but it must be testable to determine its truth. The formulation of a proposition is called a hypothesis. A model represents all or part of a system that is constructed to study that system.
- Inductive research begins with observations and attempts to build a theory from those observations. Deductive research begins with a theory and tests observations to determine if the theory is valid

3: Research Ethics

Introduction

The Oxford dictionary defines ethics as "moral principles that govern a person's behavior or the conducting of an activity." (Dictionaries, 2018) Such principles are often defined at a disciplinary level through a professional code of conduct and sometimes enforced by a university committee called an Institutional Review Board (IRB). Often, codes of conduct are codified in written form. However, even if they are not explicitly specified, researchers are still expected to be aware of and abide by general agreements shared by the research community on what constitutes acceptable and non-acceptable behaviors in the professional conduct of their disciplines. For instance, researchers should not manipulate their data collection, analysis, and interpretation procedures in a way that contradicts the principles of science or the scientific method or advances their agenda.

OBJECTIVES

- Define ethics.
- Discuss five primary ethical principles: voluntary participation, informed consent, confidentiality, disclosure, and reporting.
- Discuss the unique ethical considerations surrounding research on humans.
- Define Institutional Review Board and discuss the types of issues that may accompany those boards.
- Discuss various professional codes of ethics.

Why is research ethics important? Because science has often been manipulated in unethical ways by people and organizations to advance their private agenda and engage in activities contrary to the norms of scientific conduct. A classic example is pharmaceutical giant Merck's drug trials of Vioxx. The company hid the fatal side-effects of the drug from the scientific community, resulting in 3468 deaths of Vioxx recipients, mostly from cardiac arrest. In 2010, the company agreed to a $4.85 billion settlement and appointed two independent committees and a chief medical officer to monitor the safety of its drug development process. Merck's conduct was unethical and violated the scientific principles of data collection, analysis, and interpretation. This incident was reported by Ronald Green (Green, 2006).

Ethics is the moral distinction between right and wrong, and what is unethical may not necessarily be illegal. A researcher's conduct may not be culpable in the eyes of the law. However, it may still lead to disciplinary hearings, professional notoriety, and even job loss on the grounds of professional misconduct. These ethical norms may vary from one society to another, but this book uses ethical standards for research in Western countries.

Background

Researchers, in general, attempt to ensure their work is valid and above reproach. Indeed, mistakes are made, but these are typically discovered by peer review and are considered a simple human error. There have been, though, a few cases of intentional fraud.

- Van der Hayden (Van der Heyden, van de Derks Ven, & Opthof, 2009) reported an outright fraud perpetrated by Woo-Suk Hwang, who published a paper in 2004 about a breakthrough in human stem cell research. It was later admitted that the entire paper was wholly fabricated.
- Sheehan reported on several cases of research fraud (Sheehan, 2007), including, for example, "Dr. Eric Poehlman of the University of Vermont was sentenced in June 2006 to 1 year in jail for falsifying and fabricating research data related to menopausal changes and metabolism."
- Evans, Smith, and Willen (Evans, Smith, & Willen, 2005) reported on drug research where human trials are ". . . poorly regulated, riddled with conflicts of interest—and sometimes deadly."
- Meier (Meier, 2005) wrote that the manufacturers of heart devices face a "huge conflict of interest" when attempting to balance notifying physicians of potential flaws with the financial harm that admission would create.

Chubin (Chubin, 1985) pointed out that several behaviors—ranging from the grave (plagiarism and fabrication) to the not-so-serious (improper acknowledgment of collaborators)—slow scientific progress, undermine trust in the research process, waste public funds, and increase external regulation of science.

While the "big three" ethics problems—falsification, fabrication, and plagiarism—are a concern, most researchers report that these are not the main issue. Instead, they worry about the mundane, everyday problems that can easily creep into a research project. Two main areas deal with data and production pressure (De Vries, Anderson, & Martinson, 2006).

When working with data, it is easy to want to "cut corners." Some researchers do not get the exact data they want, so they eliminate outliers or otherwise "shape" the data to a more desirable form while cleaning it. Even highly respected scientists have been guilty of data manipulation. Sheehan reports that

Sigmund Freud fabricated case studies, and Isaac Newton altered the lunar and solar sightings record to fit his theories (De Vries, Anderson, & Martinson, 2006).

The second area is the pressure to produce published work. The old saying in research circles is "publish or perish," and there is more than a grain of truth to the fact that researchers must continue to produce published reports to attract grants and continue their research. This pressure makes it tempting to produce more studies, some of which may be of dubious value.

Research on Humans

The U.S. Department of Health and Human Services defines a human subject as "…a living individual about whom an investigator (whether professional or student) conducting research: (i) Obtains information or biospecimens through intervention or interaction with the individual, and uses, studies, or analyzes the information or biospecimens; or (ii) Obtains, uses, studies, analyzes, or generates identifiable private information or identifiable biospecimens." (Human & Services, Protection of Human Subjects, 2018). In some states, human subjects also include deceased individuals and human fetal materials.

On the other hand, investigators may manipulate or analyze nonhuman research objects or entities. In business research, nonhuman subjects typically include newspapers, historical documents, advertisements, television shows, buildings, and even garbage.

Unsurprisingly, research on human subjects is regulated much more heavily than research on nonhuman subjects. However, there are ethical considerations that all researchers must consider regardless of their research subject.

History of Research on Humans

Research on humans has not always been regulated in the way that it is today. The earliest documented cases of research using human subjects are of medical vaccination trials. One such case occurred in the late 1700s when scientist Edward Jenner exposed an eight-year-old boy to smallpox to identify a vaccine for the devastating disease, as reported by Stefan Riedel (Riedel, 2005). Medical research on human subjects continued without much law or policy intervention until the mid-1900s. At the end of World War II, several Nazi doctors and scientists were put on trial for conducting human experimentation. They tortured and murdered many concentration camp inmates, as reported by Ruth Faden (Faden &

Beauchamp, 1986). The trials, conducted in Nuremberg, Germany, created the Nuremberg Code, a ten-point set of research principles designed to guide doctors and scientists who research human subjects. Today, the Nuremberg Code guides medical and other research on human subjects, including social scientific research.

Medical scientists are not the only researchers who have conducted questionable research on humans. In the 1960s, psychologist Stanley Milgram (pictured on the right) conducted experiments designed to understand obedience to authority. He tricked subjects into believing they were administering an electric shock to other subjects (Milgram, 1963). There were no actual shocks, but some participants experienced extreme emotional distress after the experiment. The realization that one is willing to administer painful shocks to another human being just because someone who looks authoritative requires it might indeed be traumatizing even if subjects later learn that the shocks were not genuine.

Around the same time that Milgram conducted his experiments, sociology graduate student Laud Humphreys (pictured on the right) collected data for his dissertation research on the tearoom trade, the practice of men engaging in anonymous sexual encounters in public restrooms (Humphreys, 1970). Humphreys wished to understand who these men were and why they participated in the trade. To conduct his research, Humphreys offered to serve as a “watch queen,” the person who keeps an eye out for police, in a local park restroom where the tearoom trade was known to occur. What Humphreys did not do was identify himself as a researcher to his research subjects. Instead, he watched his subjects for several months, getting to know several of them, learning more about the tearoom trade practice, and surreptitiously jotting down their license plate numbers as they pulled into or out of the parking lot near the restroom. After participating as a watch queen for some time, Humphreys used the help of several insiders with access to motor vehicle registration information to obtain the names and home addresses of his research subjects. Then, disguised as a public health researcher, Humphreys visited his subjects in their homes and interviewed them about their lives and health. Humphreys’ research dispelled a good number of myths and stereotypes about the tearoom trade and its participants. He learned, for example, that over half of his subjects were married to women, and many of them did not identify as gay or bisexual. However, once Humphreys’ work became public, it created a significant controversy at his home university, among both sociologists and the public, since it raised concerns about the purpose and conduct of sociological research.

These and other studies[5] led to increasing public awareness of and concern about research on human subjects. In 1974, the U.S. Congress enacted the *National Research Act*, which created the *National Commission for the Protection of Human Subjects in Biomedical and Behavioral Research*. The commission produced *The Belmont Report* (Biomedical & Behavioral Research, 1978), a document outlining basic ethical principles for research on human subjects. The *National Research Act* also required that all institutions receiving federal support establish an Institutional Review Board (IRB) to protect the rights of human research subjects. Since that time, many organizations that do not receive federal support but conduct research have also established review boards to evaluate the research's ethics.

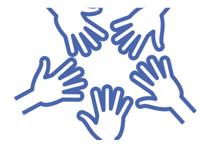
Ethical Principles

Over the past half-century or so, several ethical principles have become widely accepted. Various disciplines, like medicine, have many ethical principles that do not apply to other fields. Other ethical principles have been practiced in the past but are no longer considered relevant. The following five principles, though, are generally accepted in all research fields.

FIGURE 14: FIVE ETHICAL PRINCIPLES

[5] http://www.journalnma.org/article/S0027-9684(15)30517-4/pdf

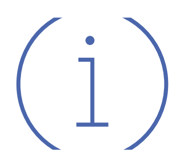
Voluntary Participation



Subjects in a research project must be aware that their participation in the study is voluntary. They have the freedom to withdraw from the study at any time without unfavorable consequences, and they are not harmed because of non-participation. The most flagrant violations of the voluntary participation principle are, arguably, the forced medical experiments conducted by Nazi researchers on prisoners of war during World War II, as documented in the post-War Nuremberg Trials (these experiments also originated the term “crimes against humanity”). Less known violations include the Tuskegee syphilis experiments (Reverby, 2009) conducted by the U.S. Public Health Service between 1932−19. Nearly 400 impoverished African-American men suffering from syphilis were denied treatment even after penicillin was accepted as an effective treatment of syphilis. Subjects were presented with false treatments such as spinal taps as cures for syphilis. Even if subjects face no mortal threat, they should not be subjected to personal agony because of their participation. In 1971, psychologist Philip Zimbardo created the Stanford Prison Experiment, where students were recruited as subjects were randomly assigned to roles such as prisoners or guards. When it became evident that student prisoners were suffering psychological damage due to their mock incarceration and student guards were exhibiting sadism that would later challenge their self-image, the experiment was terminated (Zimbardo, Maslach, & Haney, 2000).

As a less egregious example, instructors often ask students to fill out a questionnaire of some sort and inform them that their participation is voluntary. This activity must be designed so that students do not fear that their non-participation will hurt their grades in any way. For instance, it is unethical to provide bonus points for participation and no bonus points for non-participation because it places non-participants at a distinct disadvantage. To avoid such circumstances, instructors may provide an alternate task for non-participants to earn the same number of bonus points without participating in the research study or by providing bonus points to everyone irrespective of their participation.

Informed Consent



All participants in a study must receive and sign an Informed Consent form that clearly describes their right not to participate and right to withdraw before their responses in the study can be recorded. In a medical study, this form must also specify any possible risks to subjects from their participation. For subjects under 18, this form must be signed by their parent or legal guardian. Researchers must retain these informed consent forms for a period (often three years) after completing the data collection process to comply with the norms of scientific conduct in their discipline or workplace.

The consent form itself must not waive, or even appear to waive, any of the subject's legal rights. Subjects cannot release a researcher or institution from any legal liability should something go wrong during their participation in the research. Because sociological research does not typically involve asking subjects to place themselves in harm's way by, for example, taking untested drugs, sociological researchers do not often worry about the potential liability associated with their research projects. However, their research may involve other types of risks. For example, if a researcher intentionally or accidentally reveals the identity of subjects who admit unusual sexual behavior, the subject's social standing, marriage, custody rights, or employment could be jeopardized.

In some cases, subjects are asked to sign a physical consent form indicating that they have read it and fully understand its contents. In other cases, subjects are provided a copy of the consent form. Researchers are responsible for ensuring that subjects have read and understood the form before proceeding with any data collection.

One last point to consider when preparing to obtain informed consent is that not all potential research subjects are considered equally competent or legally allowed to consent to participate in research. These subjects are sometimes referred to as members of vulnerable populations, people who may be at risk of experiencing undue influence or coercion, and the consent rules are more stringent for vulnerable populations. For example, minors must have the consent of a legal guardian to participate in research. In some cases, the minors themselves are also asked to participate in the consent process by signing unique, age-appropriate consent forms explicitly designed for them. Prisoners and parolees also qualify as vulnerable populations. Concern about the vulnerability of these subjects comes from the genuine possibility that prisoners and parolees could perceive that they will receive some highly desired reward, such as early release if they participate in research. While gaining consent from vulnerable populations can be challenging, failing to work with them ensures that their stories are never told. While there is no easy solution to this double-edged sword, an awareness of the potential concerns of research on vulnerable populations is essential for identifying whatever solution is most appropriate for a specific case.

Confidentiality

Research subjects' identities must be protected in a scientific study using the dual principles of anonymity and confidentiality. Anonymity implies that the researcher or readers of the final research report or paper cannot identify a given response with a specific respondent. An example of anonymity in scientific research is a mail survey in which no identification numbers are used to track who is or is not responding to the survey. In studies of deviant or undesirable behaviors, such as drug use or illegal music downloading by students, truthful responses may not be obtained if subjects are not assured of anonymity. Further, anonymity assures that subjects are insulated from law enforcement or other authorities who may be interested in identifying and tracking such subjects in the future.

In some research designs, such as face-to-face interviews, anonymity is not possible. In other designs, such as a longitudinal field survey, anonymity is not desirable. It prevents the researcher from matching responses from the same subject at different periods for longitudinal analysis. Under such circumstances, subjects should be guaranteed confidentiality. The researcher can identify a person's responses but promises not to divulge that identity in any report, paper, or public forum. Confidentiality is a weaker form of protection than anonymity. Social research data do not enjoy the "privileged communication" status in United States courts, as do communication with priests or lawyers. For instance, two years after the Exxon Valdez spilled ten million barrels of crude oil near the port of Valdez in Alaska, communities commissioned a San Diego research firm to survey the affected households about details concerning increased psychological problems in their family. Because the cultural norms of many Native Americans made such public revelations particularly painful and challenging, respondents were assured their responses were confidential. When this evidence was presented to a court, Exxon petitioned the court to subpoena the original survey questionnaires (with identifying information) to cross-examine respondents regarding their answers given to interviewers under the protection of confidentiality was granted that request. Luckily, the Exxon Valdez case was settled before the victims were forced to testify in open court, but the potential for similar violations of confidentiality remains. Ann Cummings (Cummings, 1992) wrote an excellent review of this incident.

In one extreme case, Rik Scarce, a graduate student at Washington State University, conducted participant observation studies of animal rights activists and chronicled his findings in *Ecowarriors: Understanding the Radical Environmental Movement* (Scarce, 2016). In 1993, Scarce was called before a grand jury to identify the activists he studied. The researcher refused to answer grand jury questions in keeping with his ethical obligations as a member of the American Sociological Association and was forced to spend 159 days at Spokane County Jail. To protect themselves from travails like Rik Scarce, researchers should remove any identifying information from documents and data files as soon as they are no longer necessary. In 2002, the United States Department of Health and Human Services issued a "Certificate of Confidentiality" to protect participants in a research project from police and other authorities. Not all research projects qualify for this protection, but this can provide vital support for protecting participant confidentiality in many cases.

Disclosure

Usually, researchers must provide certain information about their study to potential subjects before data collection begins to help them decide whether to participate. For instance, researchers should answer questions about who is conducting the study, its purpose, what outcomes are expected, and who will benefit from the results. However, in some cases, disclosing such information may potentially bias subjects' responses. For instance, imagine that the purpose of a study is to examine to what extent subjects will abandon their views to conform with "groupthink." They participate in an experiment where they listen to others' opinions on a topic before voicing their own—then disclosing the study's purpose before the experiment will likely sensitize subjects to the treatment. Under such circumstances, even if the study's purpose cannot be revealed before the study, it should be revealed in a debriefing session immediately following the data collection process, with a list of potential risks or harm borne by the participant.

Reporting

Researchers also have ethical obligations to the scientific community on how data are analyzed and reported in their study. Unexpected or negative findings should be fully disclosed, even if they cast doubt on the research design or the findings. Similarly, many interesting relationships are discovered after a study is completed by chance or intentional data mining. Presenting such findings as the product of deliberate design is unethical. In other words, when using positivist research techniques, hypotheses should not be designed after the fact based on the results of data analysis because the role of data in such research is to test hypotheses and not build them. It is also unethical to "carve" data into different segments to prove or disprove hypotheses of interest or generate multiple papers claiming different data sets. Misrepresenting questionable claims as valid based on partial, incomplete, or improper data analysis is also dishonest. Science progresses through openness and honesty. Researchers can best serve science and the scientific community by fully disclosing the problems with their research to save other researchers from similar problems.

Institutional Review Board (IRB)

The IRB is tasked with ensuring that the rights and welfare of human research subjects will be protected at all institutions, including universities, hospitals, nonprofit research institutions, and other organizations that receive federal research support. IRBs typically consist of various disciplines, such as sociology, economics, education, social work, and communications. Most IRBs also include representatives from the community in which they reside. For example, representatives from prisons, hospitals, or treatment centers might sit on the IRB of nearby universities. The diversity of membership helps ensure that a knowledgeable and experienced panel will consider the complex ethical issues that may arise from research with human subjects. Investigators researching human subjects must submit proposals to the IRB for review and approval before beginning the project.

The IRB approval process may differ slightly from place to place but usually requires completing an application providing complete information about the research project, the researchers (principal investigators), and details on how the subjects' rights will be protected. Additional documentation such as an informed consent form, research questionnaire, and interview protocol may also be needed. The researchers must also demonstrate that they are familiar with the principles of ethical research by providing certification of their participation in a research ethics course. Data collection can only commence after the IRB review committee clears the project.

Even students who conduct research that involves human subjects must have their proposed work reviewed and approved by an IRB before beginning any research. Some universities permit exceptions for student projects like in-class surveys with no danger to the participants and will not be shared outside the classroom.

It may be surprising to discover that IRBs are not always popular or appreciated by researchers. In some cases, the IRB may have expertise in biomedical or experimental research but not in business or social projects. Unfortunately, business research is often open-ended, and that can be problematic for an IRB. The members of IRBs often want to know in advance precisely who will be observed, where, when, and for how long, whether and how they will be approached, precisely what questions they will be asked, and what predictions the researchers have for their findings. Providing this level of detail for a year-long participant observation within an activist group of 200-plus members, for example, would be extraordinarily frustrating for the researcher in the best of cases and most likely would prove to be impossible. Of course, IRBs do not intend to have researchers avoid studying controversial topics or avoid using specific methodologically sound data collection techniques, but, unfortunately, that is sometimes the result. The solution is not to do away with review boards, which serve a necessary and essential function, but instead to help educate IRB members about various research methods and topics covered by business, sociology, and other social scientists.

Professional Codes of Ethics

Most professional associations have established and published formal codes of conduct describing what constitutes acceptable and unacceptable professional behavior of their member researchers. The following codes are examples for researchers engaging in business and marketing research.

- Academy of Management https://aom.org/about-aom/governance/ethics/code-of-ethics
- Insights Association https://www.insightsassociation.org/issues-policies/insights-association-code-standards-and-ethics-market-research-and-data-analytics-0
- Association of Business Schools https://charteredabs.org/wp-content/uploads/2015/02/abs_ethics_guide_-_2012.pdf
- Market Research Society https://www.mrs.org.uk/standards/code%20of%20conduct

It may also be helpful to consider the codes developed by social science researchers.

- Social Research Association (SRA) https://the-sra.org.uk/research-ethics/ethics-guidelines/
- American Sociological Association (ASA) https://www.asanet.org/about/governance-and-leadership/code-ethics
- American Psychological Association (APA) https://www.apa.org/ethics/code/index

As an example, the following is the summarized Marketing Research Association's (MRA) "Code of Marketing Research Standards" (Association, 2018). The code is a 20-page document that includes 42 principles divided into three articles, an enforcement FAQ, and two appendices.

1. Article I: Responsibility to Respondents and Prospective Respondents.
    a. General Conduct. This article focuses on how to treat the respondents in a research project. It includes requirements like protect their right to drop out of a research project and their right to privacy.
    b. Purpose of Use. This article requires researchers to obtain a consent form and protect respondent information from improper use, like solicitations.

c. Transparency. This article requires researchers to be honest with respondents and make the research method transparent. It includes not collecting information without the respondent's knowledge and keeping an internal "do not call" list so respondents can opt-out of future contacts.
d. Technical Compliance. This article focuses on legal and other matters, like adhering to all state laws for projects that cross state borders and being especially careful with vulnerable populations, like children.

2. Article II: Responsibilities to Clients and Vendors. This article requires researchers to maintain a trusted relationship with clients and vendors and refrain from engaging in unacceptable practices with any research partner.
3. Article III: Professional Responsibilities. Researchers are required to report research results accurately and honestly and not falsify or omit data.

Ethical Issues in Research

Introduction

When applied to social research, ethics is concerned with creating a trusting relationship between the researcher and those researched. To ensure that trust is established, communication must be carefully planned and managed, risks minimized, and benefits maximized.[6]

Researchers adhere to several ethical principles in developing a trusting relationship, as illustrated in Figure 15: Research Ethical Principles.

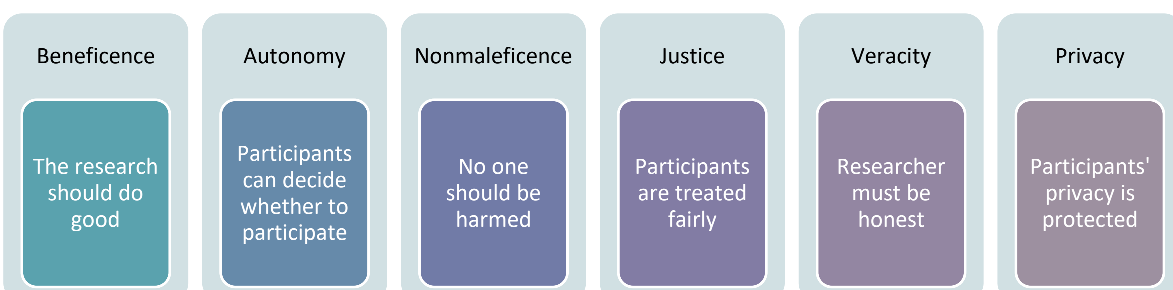


FIGURE 15: RESEARCH ETHICAL PRINCIPLES

Beneficence (Doing good)

Research should only be carried out if a benefit can be derived, i.e., contribution to the general body of knowledge or improved service/treatment. Therefore, whether a research project is worth undertaking

[6] This section of the book was adapted from information found at the *Kirklees Council* website (Council, Ethical Issues in Research, 2019).

should always be uppermost in the researcher's mind. If no benefit can be derived, then the project is unethical.

Autonomy (Self-rule)

Researchers should disclose information at a level that participants can understand to either intelligently agree or refuse to participate. In essence, autonomy is concerned with the concept of informed consent whereby people who agree to take part in a study know what they are agreeing to and authorize the researcher to collect information without any form of coercion.

Nonmaleficence (No harm)

The principle of nonmaleficence places an obligation on researchers to not harm others or expose people to unnecessary risks. Harm can come in many forms, from blows to self-esteem to "looking bad" to others, losing funding or earnings, boredom, frustration, or time-wasting. In extreme cases, research projects may even lead to physical harm. It is good practice to assume that every research project will potentially involve some form of harm and consider how best to deal with it.

Justice (Fairness)

This principle implies that everyone should be treated fairly and equally. Researchers should be careful to treat all subjects impartially and without favoritism. Of course, some research projects may be intentionally designed to offer some sort of treatment to one group and not the other so its effectiveness can be measured, but as much as possible, do not discriminate among subjects.

Veracity (Truth-telling)

This principle is about telling the truth whereby the researcher must provide comprehensive and accurate information that enhances understanding. For example, if the researcher says a questionnaire will take ten minutes to complete, the questionnaire should take ten minutes and not twenty. Researchers should always be honest with participants and keep all promises made.

Privacy

Privacy concerns the respect for limited access to another person, be it physically, emotionally, or cognitively. For example, participants may grant access to their thoughts when they agree to participate in an interview, but they do not agree to unlimited access. Therefore, they always have the right to decline to talk about certain issues or answer specific questions.

Confidentiality is an extension of privacy that relates specifically to the agreement between the researcher and participants about what can and cannot be done with the information collected throughout a project. In many cases, confidentiality will be determined by various legal constraints, but even in the absence of the law, information gathered should be protected.

Frequently asked questions

What is meant by informed consent?

They are being informed means that participants are told everything that might occur during a study in a way that they can understand. Giving consent implies that a) the agreement to participate is voluntary, free from coercion and undue influence, and b) that the person providing the consent is competent to make a rational and mature judgment about taking part. If the criteria of being informed and giving consent are met, informed consent is given.

Does consent have to be in writing?

It is an excellent practice to have consent in writing, and many IRBs will require written consent forms. However, this is not always possible in practice, especially when undertaking focus groups or field observations. The convention here is to go through the consent procedure with the group and video record any objections. For field research, informed consent should be the goal as much as that is practical. Consent forms with personally identifiable information attached should be stored in a locked container away from other information about the project to prevent breaches of confidentiality.

What information should be included on a consent form?

There are no hard and fast rules; however, as a rough guide, the following sorts of things should be included.

- A heading with the title, organization carrying out the research, and name of the researcher.
- A statement of agreement to participate.
- A statement that indicates the length of time an activity is likely to take.
- A statement that indicates what will happen to the information collected.
- A statement about confidentiality and anonymity.
- Confirmation that there is no obligation to take part and that participants have the right to withdraw at any time or not answer questions.
- Signatures and date.

The following optional statements may be included.

- A statement that the use of recording equipment has been explained.
- A statement that a leaflet has been provided and that the information has been read and understood.
- A statement that permission has been granted to contact participants in the future if that is necessary.
- A statement indicating whether permission has been granted for participants' names and contact information to be added to a database.

Research with Children and Young People

Introduction

Suppose a research project involves children and young people. In that case, more time, consideration, and planning are required when compared to adult respondents. This section provides a brief overview of some of the specific issues that need to be considered.[7]

Defining children and young people

The Office for Human Research Protections of the US Department of Health & Human Services defines children as "persons who have not attained the legal age for consent to treatments or procedures involved in the research, under the applicable law of the jurisdiction in which the research will be conducted. Generally, the law considers any person under 18 years old to be a child." (Human & Services, Special Protections for Children as Research Subjects, 2018)

Is there a minimum age for conducting research?

While the US Department of Health & Human Services does not specify a minimum age, researching with very young children should be avoided and indeed should not be undertaken by a non-specialist.

How are children and young people recruited for a research project?

Recruitment of children and young people to participate in research almost always needs to be done via a "gatekeeper." The gatekeeper will usually be a responsible adult, i.e., the person responsible for protecting the child/young person's safety and welfare at the research time. Gatekeepers will vary in different contexts, but examples might include a parent, a teacher, a caregiver, or a youth worker. It is essential to consult and involve gatekeepers during the planning stages of any research project since

[7] This section of the book was adapted from information found at the *Kirklees Council* website (Council, Research with Children and Young People, 2019).

they will usually be the ones who provide the initial consent to approach children/young people to take part.

Gaining consent

In general, children under 18 must not be consulted without a parent, guardian, or responsible adult. If the research is being undertaken within a school environment, it is suggested that consent is sought from parents/guardians and the teacher or other responsible adult at the school.

Consent is a two-stage process since it must always be obtained from the child/young person themselves and the responsible adult.

- *Stage 1*. The responsible adult must consent to approach potential participants.
- *Stage 2*. The child/young person must consent to participate in the research and decline if they wish.

Informed consent

It is essential to introduce the purpose and aims of the research clearly to ensure that both responsible adults and children/young people can give their informed consent to take part. This introductory information should be in writing wherever possible, and contact details for the person undertaking the research should always be provided.

It is not essential for consent to be provided in writing unless the subject matter is potentially sensitive, but it is often advisable to create an audit trail. The name, relationship, and role (e.g., parent) of the responsible adult giving consent should always be recorded in writing.

Different scenarios

- *Postal questionnaires*. These should be sent to the responsible adult in the first instance and not the child. Space should be provided for the responsible adult to sign that they have given their consent for the child to complete the questionnaire.
- *Telephone interviews*. The responsible adult's consent may be obtained verbally, but a written record should be retained, and that record sent to the responsible adult on request.
- *Qualitative research*. Written consent forms should be issued to parents/guardians at the recruitment stage, asking permission to invite their children to participate.
- *Online research*. A notice explaining that consent is required must be posted along with explaining the procedure for obtaining consent. Consent should be verified by letter or phone if it is provided via email. Respondents should be asked to give their age before providing any

other personal information. If the age given is under 18, they must be excluded from providing further information until the appropriate consent has been obtained.

What else should be considered?

- *Subject matter*. Extra care must be taken when consulting over sensitive or potentially contentious topics, for example, race, religion, or alcohol/drug use.
- *Questionnaire design*. The content should be appropriate to the age of respondents and relevant to their experience. The language used should be easy to understand but not patronizing.
- *Qualitative methods*. Group-based activities can encourage participation and promote discussion, and the presence of peers may put people at ease. One-to-one interviews are not recommended for young children but can work well with teenagers. “Friendship pair” interviews can be another helpful technique.
- *Venue*. Research should only be conducted in safe and appropriate environments where children/young people feel safe and comfortable.
- *Personal safety*. Precautions must be taken to ensure that research does not harm or adversely affect participants. To protect the children, interviewers who will contact them should be checked against information held by law authorities.
- *Incentives*. Any incentives used should be suitable for the age of the child/young person and appropriate to the task required.
- *Feedback*. As with all research, the results must be fed back to participants. Asking for feedback on the findings and their experience of being consulted might also help improve engaging children/young people in the future.

Commissioning a specialist

Since research with children and young people requires specialist approaches, it is often advisable to commission someone with the appropriate expertise in this area to carry out the work.

Involving Children/Young People

Experts on involving children/young people in research projects offer these recommendations.

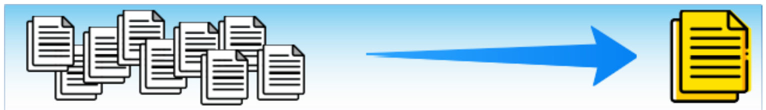
- Work with a wide range of approaches.
- Avoid creating an “elite” group who are assumed to represent children/young people on every issue.

- Meet children/young people on their territory, at times they choose and in ways that make sense to them.
- Give children/young people the chance to influence both the answers and the questions.
- Understand that children/young people are the experts on involving themselves and taking their evaluation seriously.
- Provide opportunities for enjoyment and the chance to build relationships.
- Guarantee a feedback loop.

Best Practices

The following list of “best practices” in ethics can be derived from the information in this chapter.

- *Be honest*. Nearly every code of ethical conduct boils down to only a few simple principles, and one of the most common is to research with honesty and integrity. Honest researchers will rarely make the wrong ethical decision.
- *Care*. Researchers must be careful in every aspect of the research project. Many ethical problems arise when a researcher “takes shortcuts” when gathering and analyzing data.
- *Be respectful*. Researchers must respect the intellectual property of others. If colleagues assist with a project, then they should be acknowledged. Closely aligned with respect is to avoid plagiarism.
- *Maintain confidentiality*. If participants expect their participation to be confidential, researchers must take great pains to protect that anonymity.
- *Know ethical principles*. Researchers owe it to themselves, their colleagues, and participants to be knowledgeable about ethical principles. More importantly, they need to know where their research project fits within an ethical framework and their responsibilities to the research community and the participants.
- *Disclosure*. Be open with participants about all facets of the research project. They should know the goals of the project, how the data will be protected and analyzed, and how the results will be shared. This open approach will also lead to informed consent to participate. Finally, every research project should include some sort of debriefing plan so participants can achieve a sense of closure when the project is finished.



A Sampling of Research

Zimbardo's Prison

In August 1971, Dr.Philip Zimbardo[8] started a social science experiment that has been condemned for ethical violations (Zimbardo, Maslach, & Haney, 2000). The experiment involved selecting 24 students from 70 who had volunteered to study prison life. Those students were randomly assigned to one of two groups: prisoners and guards. The guards helped build a mock prison in the basement of the Stanford University psychology department then nine prisoners were assigned to three cells in that prison. All students signed informed consent forms that indicated some of their fundamental rights would be violated if they were selected to be prisoners and that only minimally adequate diet and health care would be provided.

The first day was relatively uneventful, but after that day, the guards steadily increased their coercive and aggressive tactics and resorted to humiliation and dehumanization of the prisoners. Within 36 hours of the initial "arrest," the first prisoner had to be released because of extreme stress reactions like crying and cursing. Over the next three days, three other prisoners had to be released due to acting "crazy." The guards began to execute several "controlling" practices like waking the prisoners during the night for "counts," basically depriving them of REM sleep. They became brutal and locked misbehaving prisoners in "solitary confinement" (a closet), made them perform meaningless physical activities (like jumping jacks), and even sprayed them with high-pressure fire extinguishers.

The experiment was ended after only six days rather than the planned two-week study because, in the words of Zimbardo, ". . . too many normal young men were behaving pathologically as powerless prisoners or as sadistic, all-powerful guards." The tipping point for the experiment came after a recently graduated Ph.D. came to the prison to assist with interviews. She was not part of the experiment from the start and became emotionally upset and angry over the madness that she witnessed. She convinced Zimbardo to end the experiment for the "well-being of the young men entrusted to our care as research participants."

[8] Interestingly, Philip Zimbardo and Stanley Milgram, mentioned earlier in this chapter, were classmates at James Monroe High School in the Bronx. Zimbardo recalled the Milgram was ". . . considered the smartest kid and I voted the most popular." (Zimbardo, Maslach, & Haney, 2000)

The ethics of this experiment has been debated for decades. On the one hand, it was conducted using the guidelines promulgated by the Human Subjects Review Board. That board required fire extinguishers to be added to the prison since there were limited emergency escape routes. Additionally, the participants were told in advance to expect their rights to be suspended. Finally, the prisoners were "visited" regularly by their parents, a priest, friends, a public defender, and many graduate students and staff of the psychology department. None of those people raised any alarms.

On the other hand, it seems problematic that one group of humans was permitted to inflict pain and humiliation on another group for an extended period. The prisoners experienced social and psychological pain, but even the guards had to live with the pain of knowing that they were inflicting suffering on a peer who had done nothing wrong.

Target Stores

Target stores received considerable negative publicity for using data mining and market basket analysis to identify pregnant women[9].

Market Basket Analysis attempts to determine the types of products typically purchased together; they are in the same "market basket." For example, beer and potato chips are often purchased together, so if a store were running a special sale on beer, it might also promote potato chips. Of course, formal market basket analysis is much more in-depth and can find odd relationships between numerous products. When market basket analysis gets personal, the results can be ethically interesting.

For example, Target, Inc., like many large chains, has customer loyalty programs where customers can use a card or phone number to get special discounts on certain items. Of course, the entire shopping trip is then recorded in the company's database, and market basket analysis can determine what this one specific customer is likely to purchase.

Target's problem began when a father in Minneapolis complained that his teenage daughter had received pregnancy-related coupons. He felt that the coupons were inappropriate and promoted teen pregnancy. He later found out that his daughter was, in fact, pregnant and apologized to the store manager. Target had been able to use market basket research to determine that the girl was purchasing the types of items that pregnant women purchase, so they sent her targeted ads for pregnant women's needs.

[9] This incident was widely reported in the popular press, including Forbes (Hill, 2012) and the New York Times (Duhigg, 2012).

To build its predictive models, Target focused on women who had signed up for the baby registry. They then compared those women's purchases with all customers. Twenty-five variables were found to identify pregnant women and even when their babies were due to an astonishing degree of accuracy. The variables included buying large quantities of unscented lotions, washcloths, and supplements like calcium, magnesium, and zinc. The analytics were good enough that Target found that pregnant women tend to buy more hand sanitizers and washcloths as they get close to their delivery date. Target used these predictions to identify which women should receive specific coupons.

After that controversy settled down, Target used the same analytics to predict when people were getting married. They would send out invitations to join the bridal registry before the marrying couple had a chance to tell their parents.

Target no longer sends out ads for only one specific item in response to the negative press, but they have become much more devious. If they know, for example, that a woman is pregnant, then the circular going to that house will have ads for garden implements and coffee, but there will be several targeted ads for items that a pregnant woman would need. The house next door would get a different circular with targeted ads for, maybe, party items.

The market basket analysis being done by Target, and other stores, is perfectly legal; however, it does raise ethical questions concerning customer privacy and informed consent.

Facebook

In June 2014, Kramer, Guillory, and Hancock (Facebook employees) published a paper describing their experiment with Facebook data (Kramer, Guillory, & Hancock, 2014). The purpose of the experiment was to determine if emotional contagion occurs in social media. For the experiment, they manipulated the content in News Feeds for a select group of Facebook customers. They reduced the number of positive posts or negative posts for these specific customers. They found that when positive posts were reduced, people produced fewer positive posts, and when negative posts were reduced, the opposite happened. They suggested that emotions expressed by others on Facebook influence our own emotions, "constituting experimental evidence for massive-scale contagion via social networks."

This study was widely reported in the popular press and was the impetus for several investigations into privacy and how Facebook controls data. Specifically, there was widespread criticism about the lack of informed consent and the opportunity for users to opt out of the experiment. However, the study's authors noted that the experiment was "consistent with Facebook's Data Use Policy, to which all users agree before creating an account on Facebook, constituting informed consent for this research." There was also a discussion about oversight for the research project. However, this research was conducted by Facebook, Inc. for its internal use and fell outside the oversight of a university research department. Moreover, as a private company, Facebook is under no obligation to conform to the U.S. Department of Health and Human Services Policy provisions for the Protection of Human Research Subjects. In short, while the company could choose to follow ethical best practices concerning informed consent and participant opt-out, they are not required to do so[10].

Key Takeaways

Students who are embarking on a research project within a university setting should follow the code of ethics from the IRB of their institution. Researchers who are independently working should join an appropriate professional organization (depending on the type of research they are conducting) and adopt the code of ethics from that organization. Following are the major topics covered in this chapter.

Chapter 3: Research Ethics
• Define ethics as the "…moral principles that govern a person's behavior or the conducting of an activity." • Discuss five primary ethical principles: voluntary participation, informed consent, confidentiality, disclosure, and reporting. • Discuss the unique ethical considerations surrounding research on humans. • Define Institutional Review Board and discuss the types of issues that may accompany those boards. • Discuss various professional codes of ethics.

[10] It is understood that Facebook did not break any laws or physically harm any customers with this research project.

4: Research Design

Introduction

A research project always starts with someone curious about something observed. A crafts store owner notices that yarn displayed on one side of the store seems to sell better than that on the other side of the store and wonders why. An economist notices that during certain times of the year, the motels seem to be occupied, and at other times they are not and wonders what causes that pattern. A driver on a delivery service wonders if there is a more efficient route for the daily deliveries. These types of questions are the first step in a research project.

OBJECTIVES
• Describe how to start a research project. • Discuss the design process. • Recognize mixed methods research. • Discuss how to avoid common research mistakes. • Define seven primary research methods: experimental, survey, secondary data analysis, case study, focus group, and ethnography. • Complete the steps used in selecting an appropriate design.

Researchers typically "start where they are," an idea eloquently described by Kristin Esterberg (Esterberg, 2002), who stated, "Instead of thinking of yourself as a neutral, disinterested observer, think about the connections that you bring to what you plan to study." Researchers identify an interesting question from patterns in their everyday life and then collect and analyze data that help answer that question. This chapter concerns creating a worthwhile question and planning a research project, while later chapters describe collecting and analyzing data to answer a question.

Once researchers become curious about some topic of interest, they must determine how they feel about the topic. Honest introspection is necessary to ask themselves what they may already believe about the topic and whether they believe their perspective is the only valid one. If they determine that they have a preconceived notion that they think is the wisest perspective, that could be a problem. Researchers must also consider how they would react if their research proved them wrong about some belief. If they would be comfortable examining, and perhaps changing, what may be cherished notions based on their research, then that is one thing. However, if they would deny the research, hide the outcomes, or even change the data, that would be a different problem altogether. Of course, just because a researcher feels strongly about a topic does not mean that it should be avoided; sometimes, the best topics to research are those about which someone feels strongly.

Researchers who are prepared to accept all findings, even those that may be unflattering or challenging, may want to study a topic that evokes strong feelings intentionally. Sociology professor Kathleen Blee (Blee, 2005) has taken this route in her research. She studies hate movement participants and the people whose racist ideologies she does not share. Blee's research is successful because she was willing to honestly report her findings and observations, even those with which she may have personally taken a severe issue.

One final step at this first stage is for researchers to think about what they already know about the topic of interest. There are many sources of knowledge, and some are more prone to creating bias than others. For example, researchers may know a topic from family history, a television program, or casual conversations with friends. These sources could introduce bias in the researcher's mind, and researchers must think about how they know what they know to help identify and correct biases that they may bring to the research project.

This chapter aims to outline a process that can be used to design a research project. To be sure, this is not the only possible way to design research; it may not even be the best way for a given research project, but it would work as a starting point for many investigations.

First Considerations

Before starting the research design process, it would be helpful for the researcher to consider certain philosophical aspects. While a failure to consider these items would not doom a research project, making these decisions early on may help avoid messy re-starts.

Exploration, Description, Explanation

Three general types of research are exploratory research, descriptive research, and explanatory research. These were all more thoroughly described in Chapter 1 of this book but are briefly reconsidered here.

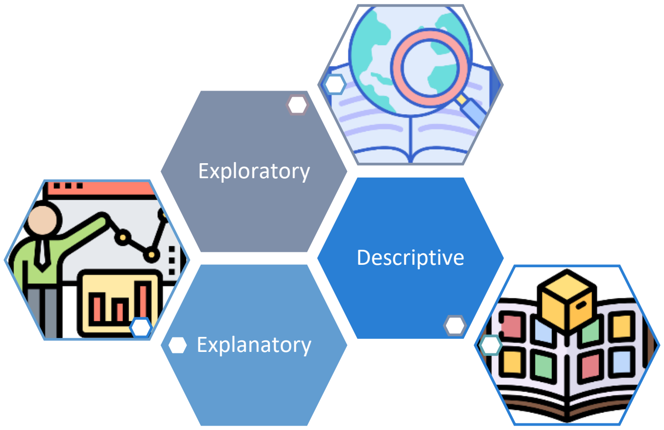


FIGURE 16: THREE TYPES OF RESEARCH

Researchers conducting exploratory research are typically at the early stages of examining a topic. Exploratory research is often designed to determine the feasibility of a more extensive study. Descriptive research describes or defines a particular phenomenon. For example, an economist publishing the gasoline prices in various parts of a city is conducting descriptive research. Finally, research that answers "why" questions are referred to as explanatory research. In this case, the researcher is trying to identify the causes and effects of an observed phenomenon.

Although research can be exploratory, descriptive, or explanatory, most business research tends to be either descriptive or explanatory. Economists frequently produce research reports that describe the state of the economy without necessarily proposing some experiment to test that description. On the other hand, business and marketing research is frequently explanatory and is designed to develop concepts and theories that explain some observed phenomenon.

Is the Topic Empirical?

An empirical topic can be investigated by observation or experience rather than one that concerns only opinions or theories. For example, if a researcher investigated healthcare costs to answer the question, "What is the best way to fund health care?" it would not make an appropriate empirical study since the definition of "best way" is nebulous. However, the question could be answered if it were re-framed a bit to describe how health care is funded, so it would be a topic that could be measured and reported.

As a second example, in 2005, the Christian group *Focus on the Family* denounced Spongebob Squarepants because they believe that he is a pro-gay activist, as reported by David Kirkpatrick (Kirkpatrick, 2005). Could a researcher determine if Spongebob is immoral? No, this is an ethical question, not empirical. A researcher could gather facts about what people think about Spongebob and even interview the program creators to see what they intended. Answering the question of morality belongs to the world of ethicists or theologians, not business researchers.

The Research Question

Once a researcher finds an empirical topic, the next step is to write the research question. Following are the qualities of a good research question.

1. *Question*. It may be rather obvious, but it must be written in the form of a question. So, for example, the topic "childfree adults" or "movies" would not be a question.
2. *Focused*. A research question must be focused on one topic of interest rather than trying to explore many areas and hope that one of them "sticks."
3. *Open-ended*. A research question should not be answered with a simple yes or no. For example, if a researcher asks, "Does location influence the price of a real estate sale," then there is nothing left to say once the "yes" or "no" answer is determined. Instead, a question like "How does location influence the price of a real estate sale" would be much better.
4. *Several answers*. A good research question should have more than one plausible answer. If the question only has one possible answer, then there is nothing to research.

Hypotheses

The purpose of positivist research[11] is to test a theory and to do that, a researcher must create a hypothesis derived from the theory. A hypothesis is a statement, sometimes causal, describing a researcher's expectation regarding the expected result of the investigation. Often, hypotheses are written to describe the expected relationship between two variables. Hypotheses are typically based on a theory and describe how an independent variable is expected to affect some dependent variable. If the theory accurately reflects the phenomenon it is designed to explain, then the researcher's hypotheses should be verified.

As an example, Social Exchange Theory postulates, among other things, that positive outcomes from social exchanges over time increase trust and commitment (Lambe, Wittmann, & Spekman, 2001). Thus, a researcher may hypothesize that brand loyalty increases due to positive outcomes from social exchanges and then design an investigation to test that hypothesis.

[11] Researchers engaged in interpretive projects, as described later in this chapter, may not start with a hypothesis, but one would be developed as the research project proceeded.

Researchers hypothesize that a relationship will take a specific direction, so an increase in one variable might increase another; the variables are correlated. For example, a researcher may study the relationship between age and consumers' preference for sustainable products. The hypothesis may be something like, "younger consumers tend to prefer sustainable products more than older consumers." The research would be designed to determine if there is a difference in product preference by age.

Note that researchers never say that they have proven a hypothesis. That statement implies that a relationship exists with absolute certainty, and there are no conditions under which the hypothesis fails. Instead, researchers tend to say that their hypotheses have been supported (or not). This more cautious way of discussing findings allows for the possibility that new evidence or new ways of examining a relationship may be discovered. Researchers may also discuss a "null hypothesis," one that predicts no relationship between the variables being studied. If a researcher "rejects the null hypothesis," then it means that the variables in question are somehow related.

Feasibility

In Chapter 3, Research Ethics, ethical considerations were discussed that may make some research projects unfeasible. Indeed, no researcher will design an experiment where a business enterprise would intentionally injure children to test some theory. However, a few practical matters are related to the feasibility of a study that researchers should consider before beginning a project.

Gaining unfettered access to a population could be problematic. For example, a project that included exploring the day-to-day experiences of maximum-security prisoners may not be feasible due to the limited access a researcher would have to that population. On a more practical level, even research about something as familiar as children's behavior concerning snacks can raise research issues. For example, Marshall, O'Donohoe, and Kline (Marshall, O'Donohoe, & Kline, 2007) conducted a study where they interviewed 8-to-11-year-old children to explore their exposure to food advertising and subsequent snack preference. While it is generally no trouble finding children that age to interview, there are questions about how honest children are with adults in a formal interview setting. While children do not necessarily lie, their responses to interview questions are almost certainly influenced by what an adult is asking. What children say to each other during play would, no doubt, be far different from what they tell an adult during an interview. It may be impossible for an adult to ever truly enter a child's world to observe what they say and do.

Another consideration would be the limits imposed by time. Suppose a researcher wants to investigate how shopping habits change in a community that is becoming gentrified. Sullivan (Sullivan, 2014) conducted surveys to determine the demographic characteristics of shoppers purchasing organic food in gentrified neighborhoods. Bridge and Dowling (Bridge & Dowling, 2001) considered gentrification from the perspective of the retail landscape in several gentrified neighborhoods. However, to understand the change that gentrification brings, a researcher may need to observe a neighborhood for many years, record the shopping families' demographics, interview them to find out what they are thinking and experiencing, and even analyze what they purchase. Unfortunately, researchers rarely have decades to devote to a single project, so this type of longitudinal study may be unfeasible.

The funding available for a study is also potentially limiting. Medical research often requires the use of costly equipment, like particle accelerators (more than $100 million), Computerized Axial Tomography (CAT) Scanners (up to $2.5 million), and Magnetic Resonance Imaging machines(about $1 million). Even surveys that use equipment no more expensive than paper and pencil require researchers to interview shoppers. If the research project involves a team of survey-takers fanning out through a wide geographic area over several weeks, the personnel cost could easily top $100 thousand. Even something as inexpensive as offering a participant a cup of coffee during an interview has a cost that must be met.

In sum, a research project’s feasibility must be considered when deciding how to complete the project, or even if it can be completed at all.

Idiographic or Nomothetic?

In general terms, research can be described as idiographic or nomothetic, as described by Joseph Ponterotto (Ponterotto, 2005). These terms derive from Kantian philosophy and are frequently found in research reports, especially psychology and sociology. However, understanding these concepts is beneficial in the planning stage for research in any field.

- *Idiographic*. This term comes from the Greek *idios*, which refers to an individual. Idiographic research concerns a single case or entity with no expectation that the research would apply to a broader application. Idiographic research sacrifices breadth of application for a deeper, richer understanding of a single case. Many case studies are idiographic because only a single individual or location is studied, and applicability beyond that case is not expected. Much of the small business research being done is idiographic.
- *Nomothetic*. This term comes from the Greek *nomos*, which refers to the traditional social norm. Nomothetic research aims to predict or explain general phenomena found in a population rather than a single case. Nomothetic research sacrifices understanding of single cases for a broader application across an entire industry. Much economic research is nomothetic in nature

since it attempts to explain broad trends in an entire population. For example, an economist may predict that the economy will improve after some significant event, but that does not guarantee that a specific business will benefit.

Applied or Basic?

Researchers hope to contribute to the body of knowledge depending on whether they are conducting applied research or basic research. Applied research can be immediately applied to a specific case. For example, applied research would help a small business owner make advertising changes to improve the number of customers entering the store. On the other hand, basic research is designed to create or validate theories and would be helpful to a legislator considering some change in the business laws of a state.

Units of Analysis

Another point to consider when designing a research project, which might differ slightly in qualitative research and quantitative research, has to do with units of observation and units of analysis. These two items concern what the researcher observes during data collection and what can later be said about those observations. A unit of observation is the item (or items) observed, measured, or collected during the research study. A unit of analysis is the entity reported at the end of the study, the “main focus” of the study. The unit of observation might be the same as the unit of analysis, but that is not always the case. However, researchers must be clear about the units of observation and analysis.

As an example, one standard unit of analysis is an individual. A research project designed to look at people's shopping habits would use the individual as the unit of analysis. Market basket research, where the content of a shopper’s basket is analyzed, uses the individual as the unit of analysis. A researcher may be interested in how some product makes a person feel or what thought process someone used to select a given product. One example of an individual unit of analysis can be found in investigating the role of social marketing on sales and services, as investigated by Alan Bright (Bright, 2000) and Philip Kotler (Kotler & Roberto, 1989).

A second standard unit of analysis is groups. Groups, of course, vary in size, and almost no group is too small or too large to be of interest to researchers. Families, friendship groups, and civic clubs (like Rotary) are a few common groups examined by researchers. For example, researchers might study how norms of workplace behavior vary across professions or how children’s sporting clubs are organized. A rich and vast body of research has been done on small businesses, and this would be using a group unit of analysis (Yusuf, 1995) (Huck & McEwen, 1991).

Organizations are yet another potential unit of analysis that researchers might wish to say something about. Organizations are large groups where the members are not necessarily as homogeneous as in a small group and include corporations, universities, and even nightclubs.

As examples, researchers might study the economic impact of globalization on the behavior of industry leadership, as researched by Diana Hechavarria (Hechavarria & Reynolds, 2009) and Randall Schuler (Schuler & Rogovsky, 1998).

Social phenomena are a potential unit of analysis. Social phenomena such as voting and even cell phone app use or misuse would be phenomena that could be researched.

Finally, researchers examine policies and principles in businesses, and those are typically contained in documents. In this case, the unit of observation would be a document, while the unit of analysis is the business. This type of project is also an excellent example of where the unit of observation and unit of analysis are different.

In sum, there are many potential units of analysis that a sociologist might examine, but some of the most common include the following.

1. Individuals
2. Groups
3. Organizations
4. Social phenomena
5. Policies and principles

Many topics could be studied from more than one level of analysis, though that would become a more complex study. For example, Kuruvilla and Ranganathan (Kuruvilla & Ranganathan, 2008) researched how micro and macro human resource policies influenced economic development strategy in India.

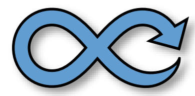
The Research Process

Research methods can be grouped into two broad approaches: positivism and interpretivism. The following table compares these two research approaches.

	Positivism	Interpretivism
Goal	Theory testing	Theory building
Methods	Laboratory experiments and surveys	Action research and ethnography
Approach	Deductive—start from theory and generate empirical data to test the theory	Inductive—start from observations and generate theory
Data	Quantitative, numeric	Qualitative, textural
Analysis	Statistical	Coding

TABLE 1: COMPARISON OF RESEARCH APPROACHES

Iterative Design



At its core, all scientific research is an iterative process of observation, rationalization, and validation. Researchers observe a natural or social phenomenon, event, or behavior of interest in the observation phase. In the rationalization phase, they try to make sense of the observed phenomenon, event, or behavior by logically connecting the different pieces of the puzzle, which, in some cases, may lead to the construction of a theory. Finally, in the validation phase, those theories are scientifically tested using data collection and analysis, leading to modifying the initial theory. However, research designs vary based on whether the researcher starts with a theory and attempts to validate it with observations (positivist research) or starts with an observation and generates a theory (interpretive research).

Most traditional research tends to be positivist, and Figure 17 provides a schematic view of such a research project. This figure depicts a series of activities to be performed, categorized into five phases: exploration, design, proposal, execution, and report. This generalized design does not fit all research, and it may be modified to fit the needs of a specific project.

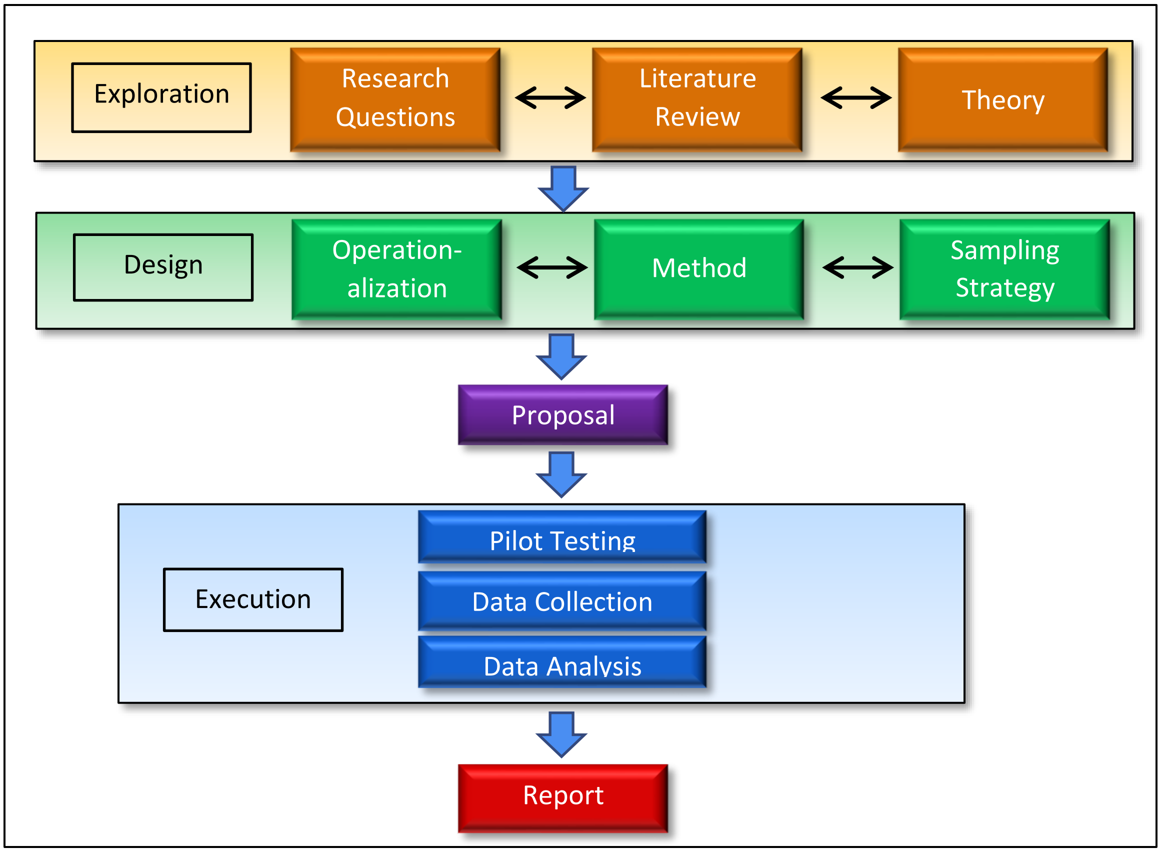


Figure 17: Positivist Research Design Process

EXPLORATION

The first phase is exploration, which includes exploring and selecting research questions for further investigation, examining the published literature to understand the current state of knowledge in the area under investigation, and identifying theories that may help answer the research questions of interest. The diagram clarifies that these three steps often run concurrently, and researchers typically shift back-and-forth between them as needed. For example, a literature review may uncover pertinent theories, but finding those theories may lead to further literature review.

Research Questions. The first step in the exploration phase is identifying one or more research questions that deal with a specific behavior, event, or phenomena of interest. Examples include what factors motivate consumers to purchase goods and services online without knowing the vendors of these goods or services, how high school students improve creativity, and why some people commit terrorist acts. More exciting research questions appeal to a broader population, address complex problems, or investigate where the answers are not obvious. Narrowly focused research questions (often with only a yes/no answer) tend to be less valuable and generally lead to unpublishable research findings.

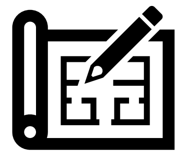
Literature Review. The second step is to conduct a literature review of the domain of interest. The purpose of a literature review is threefold.

1. To survey the current state of knowledge in the area under investigation.
2. To identify critical authors, articles, theories, and findings in that area.
3. To identify gaps in the knowledge that a research project may be able to fill.

Once a shortlist of relevant articles is generated from a search, the researcher must manually browse through each article or its abstract to determine its suitability for a detailed review. Literature reviews should be complete and not restricted to only a few journals, a few years, or a specific methodology. Reviewed articles may be summarized in tables and can be further structured using organizing frameworks such as a concept matrix. A well-conducted literature review should determine whether the research questions have already been addressed, whether there are newer or more intriguing research questions, and whether the original research questions should be modified or changed. The review can also provide some intuitions or potential answers to the questions of interest and help identify theories that have previously been used to address similar questions. Reading scholarly literature is different from reading a textbook or novel. Scholarly literature is typically divided into predictable sections. One of the easiest to find is the abstract, a short paragraph at the beginning of an article summarizing the research question, methods used to answer the question, and primary findings. The abstract often shows whether the article is relevant to the research project. Most scholarly articles contain these sections: introduction, literature review, methodology, findings, and discussion. After reading the abstract, the discussion section is usually the next most informative. Finally, the methodology section may include essential clues about the most productive way to approach a research project.

Theory. The third step is to identify one or more theories that can help address the desired research questions. While the literature review may uncover a wide range of concepts or constructs potentially related to the phenomenon of interest, a theory will help identify which of these constructs is logically relevant to the target phenomenon and how. Failing to identify related theories may result in measuring a wide range of less relevant or even irrelevant constructs while also minimizing the chances of obtaining meaningful results. In positivist research, theories can be used as the logical basis for postulating hypotheses needed in a later step. Not all theories are well-suited for studying all phenomena. Theories must be carefully selected based on their fit with the target problem and the extent to which their assumptions are consistent with that of the target problem.

RESEARCH DESIGN



The next phase in the research process is the design. This process creates a blueprint of research activities that will satisfactorily answer the questions identified in the exploration phase. The design process includes operationalizing constructs of interest, selecting a research method, and devising an appropriate sampling strategy.

Operationalization. This step of the process is to design precise measures for abstract theoretical constructs. Determining measures is a significant problem in business and marketing research, given that many of the constructs, like "average family" and "organizational culture," are hard to define and challenging to measure. Operationalization starts with specifying an "operational definition" (or "conceptualization") of the constructs of interest. Next, the researcher searches the literature to determine if existing measures can be modified to measure the constructs of interest. If such measures are not available or reflect a different conceptualization than intended by the researcher, new instruments may have to be designed. Redesigning instruments can easily be a long and laborious process, with multiple rounds of pretests and modifications before the new instrument can be accepted as "scientifically valid."

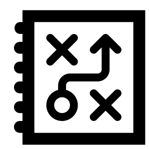
Method. Simultaneously with operationalization, the researcher must also decide what research method to employ for collecting data that will address the research question. This stage of the process is influenced by whether the research is exploratory, descriptive, or explanatory; will the approach be interpretive or positivist, will it have some direct application or contribute more generally to the field; and the unit of analysis and observation used. Research methods may include experimentation, surveys, case studies, and others, or combinations of several methods to triangulate an answer. The selected method must then be further refined; for example, surveys could be administered by mail, telephone, web, or a combination.

Sampling Strategy. Researchers must also carefully choose the target population and a sampling strategy for data collection. While selecting a sample, care should be taken to avoid a biased sample that may generate biased observations. Sampling is covered in depth in Chapter 7: Sampling.

PROPOSAL

At this stage, it is often a good idea to write a research proposal detailing all the decisions made in the preceding stages of the research process and the rationale behind each decision. This multi-part proposal should address why the research questions are being studied, the current state of knowledge, the theories or hypotheses to be tested, how the constructs will be measured, the research method to be employed, and the sampling strategy. Funding agencies typically require a detailed proposal for them to select which to fund. Even if funding is not sought for a research project, a proposal may serve as a valuable vehicle for seeking feedback from other researchers and identifying potential problems with the research project before starting data collection. This initial feedback is invaluable because it is often too late to correct critical problems after data are collected in a research study.

RESEARCH EXECUTION



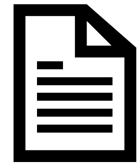
Having decided whom to study (subjects), what to measure (concepts), and how to collect data (research method), the researcher is now ready to proceed to the research execution phase. This phase includes pilot testing the measurement instruments, data collection, and data analysis.

Pilot testing. This step of the process is often overlooked but essential. It helps detect potential problems in the research design and instrumentation (e.g., whether survey questions are intelligible to the targeted sample) and ensure that the measurement instruments used in the study are reliable and valid measures of the constructs. The pilot sample is usually a small subset of the target population. After a successful pilot testing, the researcher may then proceed with data collection using the sampled population.

Data Collection. Next comes the actual collection of data. At this investigation phase, the researcher conducts surveys, visits field sites, interviews subjects, reads corporate documents, or generates other data specified by the plan.

Data Analysis. Following data collection, the data are analyzed and interpreted. Depending on the type of data collected (quantitative or qualitative), data analysis may be quantitative (e.g., employ statistical techniques such as regression or structural equation modeling) or qualitative (e.g., coding or content analysis).

RESEARCH REPORT



The final research phase involves preparing the research report documenting the entire research process and its findings in a research paper, dissertation, or monograph. The report should outline in detail why the various choices were made during the research process (e.g., the theory used, constructs selected, measures used, research methods, sampling, and similar concerns), along with the outcomes of each phase of the research process. The research process must be described in sufficient detail to allow other researchers to replicate the study, test the findings, or assess whether the inferences derived are scientifically acceptable. Research is of no value unless the process and outcomes are documented for future generations, and such documentation is essential for the progress of science.

Mixed Methods

Up to this point, the research design has been treated as if it is an either/or proposition. A research project may be positivist with numeric data, or it may be interpretative with textural data. In truth, researchers do not necessarily have to choose one approach over another. Some of the most highly regarded business and marketing investigations combine approaches to understand their topic possible. Using multiple and different research strategies is called mixed methods because the goal is to focus on "truth" from several different approaches.

Imagine that a researcher was interested in finding out how college students used electronic devices on campus. Instead of just conducting one type of research, maybe a survey, two research techniques could be used, a survey and individual interviews. Finally, add to the project a content analysis of campus policies and observations of students in their natural environments[12]. Researchers would end up with a comprehensive understanding of how students use electronic devices on campus. The drawback, of course, is that a mixed-method project requires a significant number of resources, time, and expertise to complete. Finally, along with the combined strength of each type of research, the combined weaknesses can become problematic.

Common Research Mistakes

The research process contains many pitfalls. After investing substantial amounts of time and effort, researchers may find that the questions were not answered, the findings were not interesting, or the

[12] For more information about using a mixed method type of research design, see John Brewer (Brewer & Hunter, 1989) and Charles Teddlie (Teddlie & Tashakkori, 2006). Chapter 14 of this text also details the Mixed Methods process.

research was not of "acceptable" scientific quality. Such problems typically result in research papers being rejected by journals.

- *Insufficiently motivated questions*. Often, researchers choose "pet" problems that are interesting to the individual but not to the scientific community at large, i.e., it does not generate new knowledge or insight about the phenomenon being investigated. Because the research process involves a significant investment of researchers' time and effort, they must be sure that the questions they seek to answer deal with real problems that affect a substantial portion of a population and have not been addressed in prior research.
- *Pursuing research fads*. Another common mistake is pursuing "popular" topics with limited shelf life. A typical example is studying technologies or practices that are popular today but may be obsolete in just a few years (or months). Because research takes several years to complete and publish, widespread interest in these fads may die when the research is completed and submitted for publication. A better strategy may be to study "timeless" topics that have persisted through the years.
- *Problems that cannot be researched*. Some research problems may not be answered adequately based on empirical evidence alone or currently accepted methods and procedures. Such problems are best avoided. However, ambiguously defined problems may be modified into researchable problems.
- *Favored research methods*. Many researchers tend to recast a research problem to be amenable to their favorite research method (e. g., survey research). This is an unfortunate trend. Research methods should be chosen to fit a research problem and not the other way around.
- *Blind data mining*. Some researchers tend to collect data first (using already available instruments) and figure out what to do with them. Data collection is only one step in the long process of planning, designing, and executing research. A series of other activities are needed in a research process before data collection. If researchers jump into data collection without such elaborate planning, the data collected will likely be irrelevant, imperfect, or useless, and their data collection efforts may be entirely wasted. An abundance of data cannot make up for deficits in research planning and design, particularly for the lack of exciting research questions.
- *Ecological fallacy*. This type of error occurs when claims about lower-level units of analysis are made based on data from higher-level units of analysis. In many cases, this occurs when claims are made about individuals, but only group-level data have been gathered.

- *Reductionism*. This type of error occurs when claims about some higher-level units of analysis are made based on data from some lower-level units of analysis. As an example, claims about groups are made based on individual-level data.

Research Designs

As noted on page 25, research designs can be classified into positivism and interpretivism, depending upon the researcher's background, temperament, and research goal. Positivist designs are meant for theory testing, while interpretive designs are meant for theory building. Prevalent examples of positivist designs include experimental, surveys, secondary data analysis, and case research, while interpretive designs include case research, phenomenology, and ethnography. Note that case research can be used for both theory building and theory testing, though not simultaneously. Some techniques, such as focus groups, are best suited for exploratory research, others such as ethnography are best for descriptive research, and still others such as laboratory experiments are ideal for explanatory research.

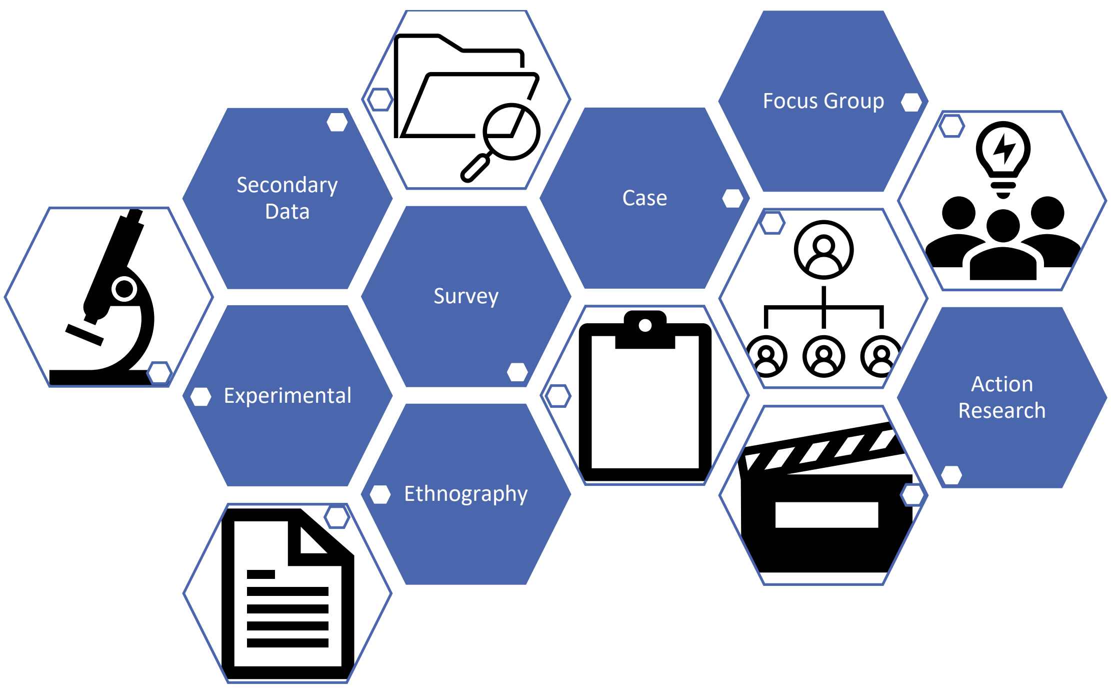


FIGURE 18: RESEARCH DESIGNS

Experimental

Experimental studies are intended to test cause-effect relationships (hypotheses) in a tightly controlled setting. This testing is accomplished by separating the cause from the effect, administering the cause to one group of subjects (the "treatment group") but not to another group ("control group"). Then, observing how the mean effects vary between subjects in these two groups. For instance, if a laboratory experiment is designed to test the efficacy of a new drug in treating a specific ailment, then a sample of people afflicted with that ailment is found. They are randomly assigned to one of two groups (treatment and control). The drug is administered to subjects in the treatment group, while a placebo is given to the control group. Finally, the two groups are monitored to see if the treatment group has a better response than the control group. More complex designs may include multiple treatment groups, such as low versus high dosage of the drug, and multiple treatments, such as combining drug administration with dietary interventions.

In an experimental design, the subjects are randomly assigned to a group. It is ideal if the researcher knows whether individuals are in the treatment or control groups. However, the treatment protocol scientists are unsure if a specific subject receives the drug under test or a placebo. This type of design is called a "double-blind" study since neither the subject nor the person administering the treatment is sure who is in the treatment group.

If random assignment is not possible for some reason, then the research design becomes "quasi-experimental."

Experiments can be conducted in a laboratory or the field where the phenomenon of interest can be studied. Laboratory experiments allow the researcher to isolate the variables of interest and control for extraneous variables, which may not be possible in field experiments. Hence, inferences drawn from laboratory experiments tend to be stronger in internal validity[13] , while field experiments tend to be stronger in external validity.

Experimental data are analyzed using quantitative statistical techniques. The primary strength of the experimental design is its compelling internal validity due to its ability to isolate, control, and intensively examine a small number of variables. In contrast, its primary weakness is limited external generalizability since real life is often more complex (i.e., involves more extraneous variables) than contrived laboratory settings. Furthermore, if the research does not identify relevant extraneous variables and controls those variables, it may decrease internal validity and lead to spurious correlations.

[13] Validity is discussed in Chapter 5.

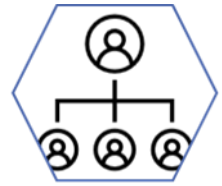
Surveys

Field surveys are non-experimental designs that do not control or manipulate independent variables or treatments but measure these variables and test their effects using statistical methods. Field surveys capture snapshots of practices, beliefs, or situations from a random sample of subjects in field settings through a survey questionnaire or, less frequently, through a structured interview. In cross-sectional field surveys, independent and dependent variables are measured at the same point in time (e.g., using a single questionnaire). In contrast, in longitudinal field surveys, dependent variables are measured later than the independent variables. The strengths of field surveys are their external validity (since data are collected in field settings), their ability to capture and control for many variables, and their ability to study a problem from multiple perspectives or using multiple theories. However, because of their non-temporal nature, internal validity (cause-effect relationships) is problematic. Surveys may also be subject to respondent biases (e.g., subjects may provide a “socially desirable” response rather than their actual response) which further decreases internal validity.

Secondary Data Analysis

Secondary data analysis scrutinizes data that have previously been collected and tabulated by other sources. Data sources may include government agencies (e.g., employment statistics from the U.S. Bureau of Labor Statistics), other researchers (e.g., dissertations), or publicly available third-party data (financial data from stock markets). This design contrasts with most other research designs where collecting primary data for research is part of the researcher’s job. Secondary data analysis may be an effective means of research where primary data collection is too costly or unfeasible and secondary data suitable for answering the research question is available. The limitations of this design are that the data may not have been collected systematically or scientifically and hence unsuitable for scientific research. Also, since the data were collected for a presumably different purpose, they may not adequately address the research questions. Finally, internal validity is problematic if the temporal precedence between cause and effect is unclear.

Case Research



Case research[14] is an in-depth investigation of a problem in one or more real-life settings (case sites) over an extended period. Data may be collected using a combination of interviews, personal observations, and internal or external documents. Case studies can be positivist in nature (for hypotheses testing) or interpretive (for theory building). The strength of this research method is its ability to discover a wide variety of social, cultural, and political factors potentially related to the phenomenon of interest that may not be known in advance. Analysis tends to be qualitative but heavily contextualized and nuanced. Weaknesses of case research include dependence on the observational and analytical ability of the researcher, lack of control which makes it difficult to establish causality, and inability to generalize findings from a single case site to other case sites. Generalizability can be improved by comparing the analysis from other case sites in a multiple case design.

Focus Groups

Focus group research is a type of research that involves bringing in a small group of subjects (typically six to ten people) to one location and having them discuss a phenomenon of interest for about two hours. The discussion is moderated by a trained facilitator who sets the agenda and poses an initial set of questions for participants, then ensures that ideas and experiences of all participants are recorded, and then attempts to build an understanding of the problem based on participants' comments. One prevalent use for a focus group is to evaluate several different advertising campaigns to determine the most viable. Internal validity cannot be established due to a lack of controls, and the findings may not be generalized to other settings because of the small sample size. Hence, focus groups are not generally used for explanatory or descriptive research but are suited for exploratory research projects.

Action Research

Action research assumes that complex social phenomena are best understood by introducing interventions, or "actions," into those phenomena and then observing the effects of those actions. In this method, the researcher is usually a consultant or a person embedded within a social context, such as an organization that initiates an action, such as new organizational procedures or new technologies, in response to a real problem, such as declining profitability operational bottlenecks. The researcher's choice of actions must be based on a

[14] It is important to keep in mind that case research is not the same as a university class discussing a case study. Case research is the process of going to a site, gathering data about that site, and analyzing that data.

theory explaining why and how such actions may cause the desired change. The researcher then observes the results of that action, modifying it as necessary, while simultaneously learning from the action and generating theoretical insights about the target problem and interventions. The initial theory is validated by the extent to which the chosen action successfully solves the target problem. Simultaneous problem solving and insight generation is the central feature that distinguishes action research from all other research methods. Hence, action research is an excellent method for bridging research and practice. This method is also suited for studying unique social problems that cannot be replicated outside that context. However, it is also subject to researcher bias, and the generalizability of findings is often restricted to the context where the study was conducted.

Ethnography

Ethnography is an interpretive research design inspired by anthropology that emphasizes that a phenomenon must be studied within its culture. The researcher is deeply immersed in a particular culture over an extended period (a few months to several years). During that period, the researcher engages, observes, and records the daily life of the studied culture. The goal is a theory about the behaviors in that culture. Data are collected primarily via observational techniques, formal and informal interaction with participants in that culture, and personal field notes, while data analysis involves “sense-making.” The advantages of this approach are its sensitiveness to the context, the rich and nuanced understanding it generates, and minimal respondent bias. However, this is also an extremely time and resource-intensive approach, and findings are specific to a given culture and less generalizable to other cultures.

Selecting the Research Design

Researchers tend to select designs that they are most comfortable with and feel most competent to handle; ideally, the choice should depend on the nature of the studied research phenomenon. In the preliminary phases of research, when the research problem is unclear, and the researcher wants to scope out the nature and extent of a specific research phenomenon, a focus group (for an individual unit of analysis) or a case study (for an organizational unit of analysis) is an ideal strategy for exploratory research. As the research project evolves, interpretive designs, such as case research or ethnography, may be appropriate. If a literature review finds competing theories, positivist designs such as experimental, survey, or secondary data analysis are more appropriate.

Regardless of the specific research design chosen, the researcher should attempt to collect quantitative and qualitative data using various techniques such as questionnaires, interviews, observations, documents, or secondary data. For example, structured surveys typically collect quantitative data, but researchers may leave room for a few open-ended questions to collect qualitative data that may generate unexpected insights. Likewise, case research may employ face-to-face interviews to collect qualitative data; administering a concurrent survey to collect quantitative data may be effective. For example, in a case study of organizational decision-making processes, a survey could record numeric quantities such as how many weeks it took to make organizational decisions to provide insights not otherwise available from narrative responses. Irrespective of the specific research design employed, the researcher's goal should be to collect as much and as diverse data as possible that can help generate the best possible insights into the phenomenon of interest.

A Sampling of Research

Does Human Capital Matter?

Any given topic will generate numerous published research reports. Consequently, one exciting research method is to examine already published reports in a meta-analysis that attempts to find patterns that may not be evident from a single study. For example, Crook et al. conducted a meta-analysis of 66 studies focused on the relationship between human capital and firm performance (Crook, Todd, Combs, Woehr, & Ketchen Jr, 2011). Their goal was to determine to what extent investment in human capital (like hiring and training) improves a firm’s performance. Performing a meta-analysis is a challenging form of research.

On the one hand, there is no reason to “reinvent the wheel” if a study has already been completed. On the other hand, prior studies have widely divergent goals, methods, and analyses that make them challenging to merge into a single outcome. In Crook’s study, the researchers found that human capital is strongly related to firm performance. However, that relationship is influenced by the market's competitiveness for the type of human capital under consideration. In other words, while human capital is essential to a firm’s performance, the cost of acquisition and development must temper importance. This result would not have been easy to determine from a single study but is more apparent from a metanalysis of other studies.

Key Takeaways

This chapter focused on designing research projects. Included was information about the following topics.

Chapter 4: Research Design
• Describe how to start a research project. • Discuss the design process. • Recognize mixed methods research. • Discuss how to avoid common research mistakes. • Define seven primary research methods: experimental, survey, secondary data analysis, case study, focus group, and ethnography. • Complete the steps used in selecting an appropriate design.

Quantitative Methods

Quantitative methods are based on the measurement of concepts and the statistical analysis of those measures. Quantitative methods include activities like sampling, surveys, and experimental research.

5: Defining and Measuring Concepts

Measurement

Measurement is important. People who have attempted to bake a cake from scratch without measuring the ingredients will find, no doubt, that measurement is the difference between a sweet desert and a disaster. Just like in baking, measurement is essential to a researcher. Measurement means the process by which key facts, attributes, concepts, and other phenomena are described. At its core, measurement is about defining the research project's terms precisely and measurable. Of course, measurement in business research is not quite as simple as using some predetermined or universally agreed-on tool, such as a measuring cup. However, there are some basic tenants on which most researchers agree when it comes to measurement.

Objectives
• Define measurement.
• Distinguish between conceptualization and operationalization.
• Define reliability and validity.
• Compare and contrast reliability and validity.
• Suggest ways to improve reliability and validity in research projects.

What Do Researchers Measure?

The question of what business researchers measure can be answered by asking what business researchers study. Researchers study a wide variety of business and marketing concepts, like corporate culture (Denison, 1990), the price elasticity of gasoline (Hughes, Knittel, & Sperling, 2006), employee turnover (Hom & Griffeth, 1995), and automobile "lemons" (Akerlof, 1978). Each of these topics required measurements of various types, and researchers had to determine the best way to do that. It may be evident that researchers measure just about anything they are interested in investigating.

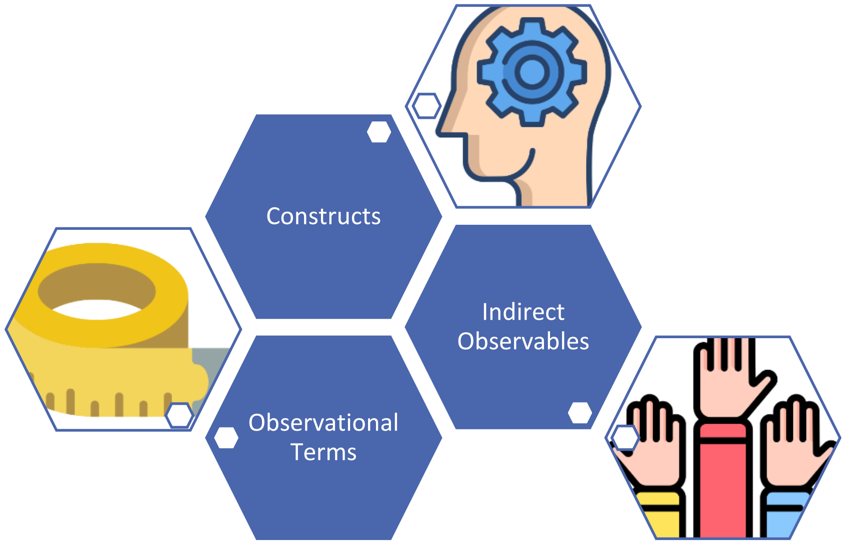


FIGURE 19: CATEGORIES OF THINGS RESEARCHERS MEASURE

In 1964, philosopher Abraham Kaplan wrote what has since become a classic work in research methodology, The Conduct of Inquiry (Kaplan, 2017). In his text, Kaplan describes different categories of things that behavioral scientists observe. One of those categories, which Kaplan called "observational terms," is probably the simplest to measure and are the sorts of things that can be seen with the naked eye simply by looking at them. Moreover, they are terms that "lend themselves to easy and confident verification." So if, for example, researchers wanted to know how the conditions of playgrounds differ across different neighborhoods, they could directly observe the variety, amount, and condition of equipment at various playgrounds.

Indirect observables, on the other hand, are less straightforward to assess. They are "terms whose application calls for relatively more subtle, complex, or indirect observations, in which inferences play an acknowledged part. Such inferences concern presumed connections, usually causal, between what is directly observed and what the term signifies." If researchers conducted a study for which they wished to know a person's income, they could simply ask in an interview or a survey. Thus, they would have observed income, even if it were only observed indirectly. Birthplace might be another indirect observable. Researchers can ask study participants where they were born, but the chances are good that they will not directly observe anyone being born in their research locations.

Sometimes the measures of interest are more complex and more abstract than either observational terms or indirect observables. For example, think about concepts like ethnocentrism, the way a person judges another person's culture, and how measuring that concept would be very challenging. In the same way, a concept like "bureaucracy" would be complicated to measure. In both cases, ethnocentrism and bureaucracy, the theoretical notions represent ideas whose meaning is known, but the concept's measurement may be nearly impossible. Kaplan referred to these more abstract things as constructs. Constructs (pronounced “CON-structs”) are a cluster of behaviors that are often seen together. As an example, anxiety could be considered a construct that includes behaviors like fidgeting and fingernail biting. Constructs are “not observational either directly or indirectly,” but they can be defined based on other observable factors, called variables.

How Do Researchers Measure?

Measurement in business research is a process. It occurs at multiple stages of a research project: in the planning stage, in the data collection stage, and sometimes even in the analysis stage.

For example, imagine that the research question is: How do new college students cope with the adjustment to college? The first problem is to define “cope” in such a way that it can be measured. After that, the data collection phase can be designed to measure whatever “cope” means. After the data are collected, then the analysis begins. Perhaps during the analysis phase, an unexpected facet of coping is discovered, and that may mean that the measures taken would need to be revisited to allow for that facet. Once the analysis is complete, then there are certain decisions concerning the report. For example, perhaps one method of coping is more effective than others, so the report may contain a recommendation that future research is conducted that measures just that one method of coping. The point is that those measurement considerations are essential throughout the research project.

The measurement process could also involve multiple stages. Starting with identifying and defining critical terms to determining how to observe and measure them to assessing the quality of the measurements, there are multiple steps involved in the measurement process.

An additional step in the measurement process involves deciding what type of data will be collected and an appropriate analysis process for those data elements.

Conceptualization

One of the first steps in the measurement process is conceptualization, which defines the terms of the project as clearly as possible. Keep in mind that terms mean only what the researcher determines, nothing more and nothing less.

A concept is a notion or image conjured up when the researcher thinks of some related observations or ideas. For example, masculinity is a concept. A researcher thinking about that concept may imagine some set of behaviors and perhaps even a particular style of self-presentation. Of course, not everyone will conjure up that same set of ideas or images: in fact, there are many possible ways to define the term. While some definitions may be more common or have more support than others, there is not one true, always-correct-in-all-settings definition for "masculine." Moreover, that definition may change over time, from culture to culture, and even from individual to individual, as explained by George Mosse (Mosse, 1996). Therefore, defining concepts is so important before any data gathering begins.

It may seem unreasonable for a researcher to define a term for which there is no single, correct definition. Unfortunately, this will be a problem for most concepts measured in a business or marketing study. William Clinton, the 42d President of the United States, famously stated, "It depends upon what the meaning of the word 'is' is."[15] Without understanding how a researcher has defined the key concepts, it would be impossible to understand the importance of the findings.

Defining concepts is an early part of conceptualization, which involves writing out clear, concise definitions for key concepts. Brainstorming may help conceptualize a topic, but it would also make sense to consult existing research and theory to see if other scholars have already defined the concepts of interest. This historical review does not necessarily mean that their definitions are correct but understanding how concepts have been defined in the past will help with a current project. Conceptualization is not as simple as applying a definition from a dictionary; it requires careful consideration and evaluating alternative concepts.

Concepts often include explicit and implicit constructs, so those constructs should be identified at this stage in the design process. Unidimensional constructs are those that have a single underlying dimension and can be measured using a single test. Examples include simple constructs such as weight, wind speed, or duration. Multidimensional concepts consist of two or more underlying constructs. For instance, if a person's academic aptitude is conceptualized as two constructs, mathematical and verbal ability, academic aptitude is a multidimensional concept. Each of the underlying dimensions, in this case, must be measured separately using different tests for mathematical and verbal ability. Then those two scores would be combined to create an overall value for the academic aptitude concept.

Before moving on to the next steps in the measurement process, it would be wise to consider one of the dangers of conceptualization. While it is essential to consult prior scholarly definitions of key concepts, it

[15] This was widely reported in the press and can be easily found on-line, including YouTube videos of him making that statement.

would be wrong to assume that those definitions are better than those generated by the researcher. Furthermore, it would also be wrong to assume that defining a concept means that it exists beyond an abstract idea, a notion known as *reification*.

To better understand reification, take a moment to think about the concept of “family.” This concept is central to sociological thinking, but it is an abstract term. If researchers were interested in studying this concept, they would consult prior research to understand how others have conceptualized the term. However, they should also question past conceptualizations. Today’s conceptualization of “family” would be very different from one used a hundred years ago or even ten years ago. The point is that terms mean nothing more and nothing less than whatever definition is assigned by the researcher. Sure, it makes sense to come to some agreement about what various concepts mean. Without that agreement, it would be challenging to navigate through everyday living. Nevertheless, at the same time, it is imperative to remember that society has assigned those definitions and that they are no more correct than any other definition a researcher might choose to assign.

Closely associated with the ideas of concepts and constructs is that of theory. People often throw around this term with a phrase like “it is just a theory.” However, scholars have thoroughly researched a theory and accepted it as the best explanation for a phenomenon. Sutton (Sutton & Staw, 1995) provides this definition:

> . . . theory is the answer to queries of why. Theory is about the connections among phenomena, a story about why acts, events, structure, and thoughts occur. Theory emphasizes the nature of causal relationships, identifying what comes first as well as the timing of such events. Strong theory, in our view, delves into underlying processes so as to understand the systematic reasons for a particular occurrence or nonoccurrence.

Theories are the start point for most quantitative research (known as theory-testing) and the endpoint for most qualitative research (known as theory-building).

Operationalization

Once a theoretical construct is defined, measuring the construct is defined in a process called operationalization. For instance, if a family’s socioeconomic status is defined as their income level, that can be operationalized by asking respondents about their family income on a survey. Given the high level of subjectivity inherent in social science constructs, most are measured using multiple indicators.

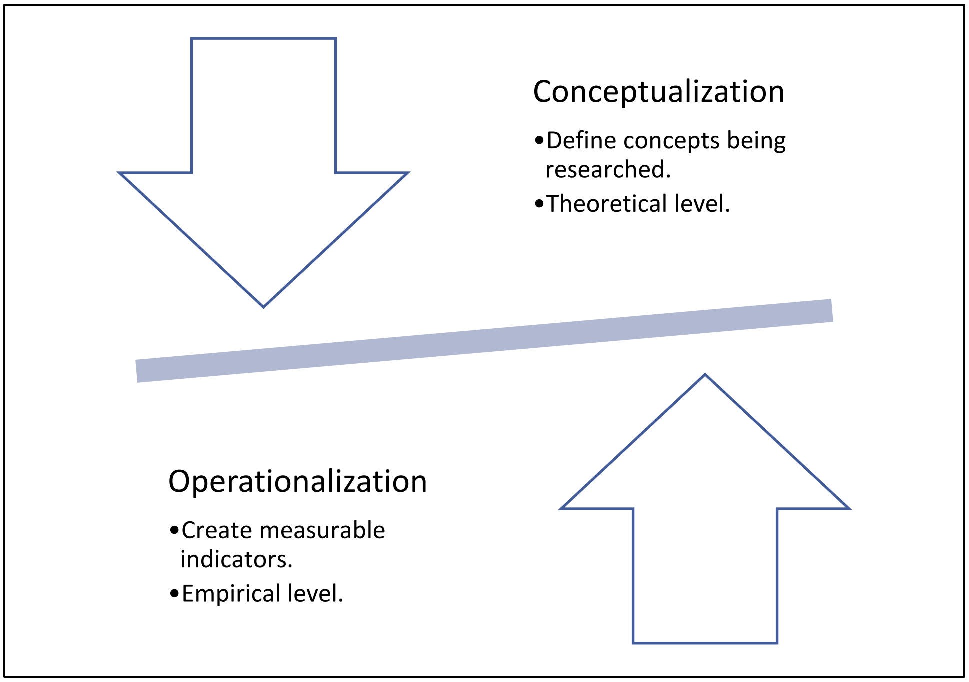


FIGURE 20: CONCEPTUALIZATION VS. OPERATIONALIZATION

Indicators operate at the empirical level in contrast to constructs, which are conceptualized at the theoretical level. The indicators representing a given construct are called variables, and up to six types are found in research projects[16]:

- *Independent variables* are those that cause an observed outcome. For example, if a researcher's concept is that older people have more traffic accidents, then the independent variable is age.
- *Dependent variables* are those that depend on the independent variable; these are the outcomes of some influence. For example, in the traffic accident concept from above, the dependent variable is the number of traffic accidents.
- *Intervening (or mediating) variables* stand between the dependent and independent variables and mediate the effects of the independent variable. For example, if the research concept is that poor people have short lifespans, there must be an intervening variable since poverty alone cannot determine lifespan. In this case, perhaps the intervening variable is lack of access to health care or inadequate nutrition.

[16] An example of the types of variables can be found in Chapter 2, page 25.

- *Moderating variables* affect the direction and strength of the relationship between the independent and dependent variables. For example, in medical experiments, the dosage level (independent variable) moderates the outcome (dependent variable).
- *Control variables* are only found in experiments. A control is a subject that has not been exposed to treatment and is used to compare the treatment outcomes. For example, if an experiment is conducted to see if tutoring improves test scores, there must be a control group that does not get tutoring.
- *Confounding variables* are "extra" unaccounted variables. If an experiment does not yield the anticipated results, researchers will often search for a confounding variable that may have influenced the experiment in unexpected ways.

Each indicator may have several attributes (or levels), and each attribute represents a value. For instance, a "homeownership" variable may have two attributes: own/rent. Likewise, a customer satisfaction scale may be constructed to represent five attributes: "strongly dissatisfied," "somewhat dissatisfied," "neutral," "somewhat satisfied," and "strongly satisfied."

Variables may be quantitative (numeric) or qualitative (textual). Quantitative data are analyzed using regression or equation modeling techniques, while qualitative data are analyzed with coding techniques. Note that many variables in business research are qualitative, even when represented with numbers. For instance, imagine a customer satisfaction indicator with five attributes: strongly dissatisfied, somewhat dissatisfied, neutral, somewhat satisfied, and strongly satisfied. If the researcher assigns the numbers 1−5 for these five attributes, then sophisticated statistical tools for quantitative data analysis can be used. However, note that the numbers are only labels associated with respondents' evaluation and the underlying variable (satisfaction) is qualitative.

Indicators may be reflective or formative. A reflective indicator is a measure that "reflects" an underlying construct. For example, if religiosity is defined as a construct that measures how religious a person is, attending religious services would reflect that construct. A formative indicator is a measure that "forms" or contributes to an underlying construct. Such indicators may represent different dimensions of the construct of interest. For instance, if the definition of religiosity includes a belief dimension, a devotional dimension, and a ritual dimension, then indicators chosen to measure each of these different dimensions will be considered formative. Thus, unidimensional constructs are measured using reflective indicators. In contrast, multidimensional constructs are measured as a formative combination of the multiple dimensions, even though each of the underlying dimensions may be measured using reflective indicators.

It is vital to keep in mind that the process of coming up with indicators cannot be arbitrary or casual. One way to avoid taking an overly casual approach in identifying indicators is to turn to prior theoretical and empirical work. Theories will point to relevant concepts and possible indicators, while empirical work will detail specific examples of how key concepts have been measured in the past. One final important detail to think about when deciding on indicators is the strategy used for data collection. A survey implies one way of measuring concepts, while field research implies a very different way. Thus, the data-collection strategy employed will play a significant role in shaping how concepts are operationalized.

Measurement Quality

The previous section examined some of the difficulties with measuring constructs. What makes the task more challenging is that sometimes these constructs are imaginary concepts (i.e., they do not exist in reality) and multi-dimensional (in which case, there is the additional problem of identifying their constituent dimensions). It is not adequate to measure constructs using any random scale; the scales must be tested to ensure that:

- they measure the construct consistently and precisely (i.e., the scales are "reliable")
- they measure the construct being investigated (i.e., the scales are "valid").

Reliability, the consistency of a measure, and validity, the efficacy of a measure, are the two yardsticks against which the accuracy of measurements is evaluated in scientific research. A measure can be reliable but not valid if it measures consistently, but it is the wrong construct. Likewise, a measure can be valid but not reliable if it inconsistently measures the construct. Using the analogy of a shooting target, as shown in Figure 21, a reliable and valid measure is like a tightly clustered group near the center of the target. A measure that is reliable but not valid is like a group that is tightly clustered but off-center. A measure that is valid but not reliable is a group that is widely scattered but centered. Finally, a measure that is neither reliable nor valid is like a widely scattered and off-center group.

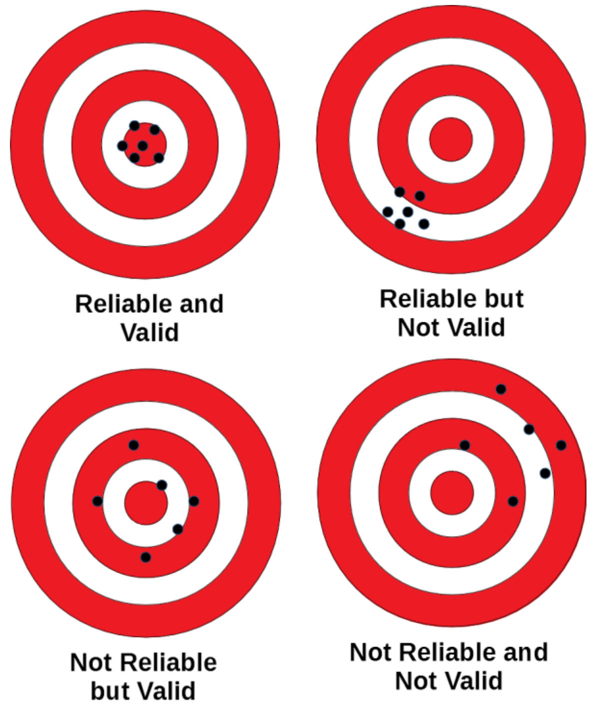


Figure 21: The Relationship Between Reliability And Validity

Reliability

Reliability "…is the extent to which measurements are repeatable – when different persons perform the measurements, on different occasions, under different conditions, with supposedly alternative instruments which measure the same thing." (Drost & others, 2011)

Any score obtained by a measuring instrument (the observed score) is composed of two parts, the "true" score, which is the score that a person would have received if the measurement were perfectly accurate, and the "error" in the measurement process. Imagine a simple example, a bathroom scale. If a person's accurate weight were 150 pounds, then, ideally, the scale would read 150 every time that person stepped on the scale. The scale's reliability is the consistency of its output from one day to the next. If a person stepped on the scale one day and it read 160, but the next day, it read 140, then the scale would not be a reliable instrument.

There are two types of reliability errors that researchers need to understand. First is systematic error, one that is caused by the system and is predictable. For example, if the bathroom scale mentioned above constantly read five pounds heavy, it would be an error, but it would be consistent and correctable in the analysis. The second type of error is a random error. If the bathroom scale were accurate, but one day it read 151 and the next 145, then that would be a random error. Random errors cannot be corrected but tend to cancel out if there are many data points due to the random nature of the error—sometimes the reading will be a bit high and other times low.

Unreliable measurements in business research could be for several reasons. One is the researcher's subjectivity. For example, imagine that employee morale in a firm is measured by watching whether they smile at each other, make jokes, or similar indicators. Different observers may infer different measures of morale if they are watching the employees on a hectic day (when they have no time to joke or chat) or a light day (when they are more jovial or chatty). Two observers may also infer different morale levels on the same day, depending on what they view as a joke.

"Observation" is a qualitative measurement technique, but reliability may be improved by using quantitative measures. Counting the number of grievances filed over one month may measure (the inverse of) morale. Of course, grievances may or may not be a valid measure of morale, but it is less dependent on human subjectivity and therefore more reliable.

Another source of unreliable observation is asking imprecise or ambiguous questions. For instance, if shoppers were asked where they like to shop, it is unclear if that means shopping for groceries, clothes, or another product type. Thus, the resulting information will be divergent and unreliable.

A final source of unreliability is asking questions about issues that respondents neither understand nor care about, such as asking an American college graduate about Canada's relationship with Slovenia.

To improve reliability, researchers can replace subjective data collection techniques (observation) with objective (questionnaire), ask questions that respondents care about, avoid ambiguous items, and simplify the wording in indicators. While these strategies can improve the reliability of measurements, instruments must still be tested for reliability using techniques like the following.

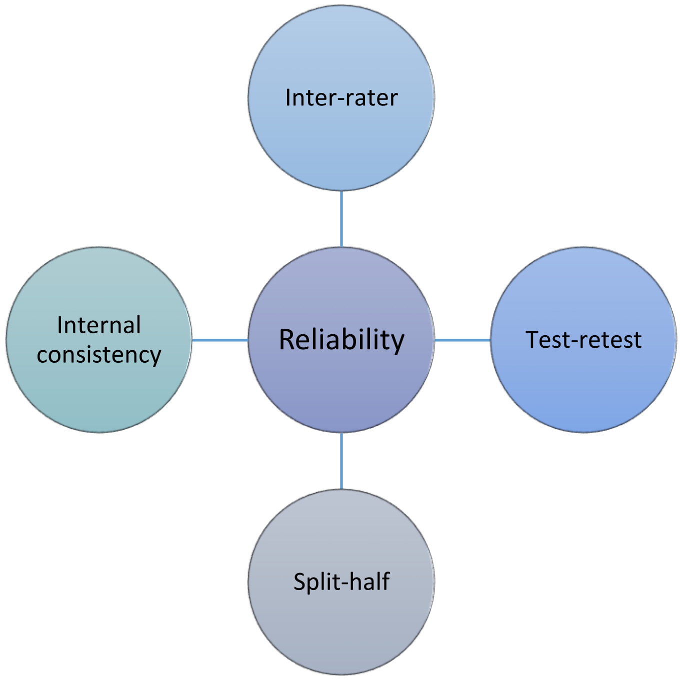


FIGURE 22: RELIABILTY

- *Inter-rater reliability*. Inter-rater reliability, also called inter-observer reliability, is a measure of consistency between two or more independent raters (observers) of the same construct. Usually, a pilot study can assess reliability in two ways, depending on the level of measurement being used. If the measure is categorical, all categories are defined, and raters check which category each observation falls in. The percentage of agreement between the raters is used to estimate inter-rater reliability. For instance, if two raters are rating 100 observations into one of three possible categories, and their ratings match 75% of the observations, then inter-rater reliability is 0.75. Suppose the measure is interval or ratio scaled (e.g., classroom activity is measured once every five minutes by two raters on a one-to-seven scale). In that case, a simple correlation between measures from the two raters can also serve as an estimate of inter-rater reliability.
- *Test-retest reliability*. Test-retest reliability measures consistency between two measurements (tests) of the same construct administered to the same sample at two different points in time. If the observations have not changed substantially between the two tests, then the measure is reliable. The correlation in observations between the two tests is an estimate of test-retest reliability. Note here that the time interval between the two tests is critical. Generally, the

longer the time gap, the greater the chance that the two observations may change due to random error and the lower the test-retest reliability.

- *Split-half reliability*. Split-half reliability is a measure of consistency between two halves of a measure. For instance, if a ten-item test of a given construct is administered to a group of subjects, then the ten items are randomly split into two sets of five, but all ten items are kept in the test. That test is next administered to a sample of respondents. Finally, the score for each group of five questions is calculated for each respondent, and the correlation between the two half scores is the split-half reliability. The longer the instrument, the more likely the two halves of the measure will be similar (since random errors are minimized as more items are added). Hence, this technique tends to overestimate the reliability of longer instruments systematically.
- *Internal consistency reliability*. Internal consistency reliability is a measure of consistency between different items of the same construct. If a multiple-item construct measure is administered to respondents, the extent to which respondents similarly rate those items reflects internal consistency. This reliability can be estimated in terms of average inter-item correlation, average item-to-total correlation, or more commonly, *Cronbach's alpha*.

Validity

Validity is concerned with the meaningfulness of research results. In brief, does the research measure what it was purported to measure? For example, does the Scholastic Aptitude Test (SAT) predict the likelihood of a high school student completing college? (Drost & others, 2011) There are numerous types of validity found in the literature. However, they generally form two large groups: Measurement Validity (the measurement should accurately reflect the construct) and Hypothesis Validity (the hypotheses should accurately reflect the construct). Figure 23 shows the relationship between the various types of validity covered in this section of the chapter.

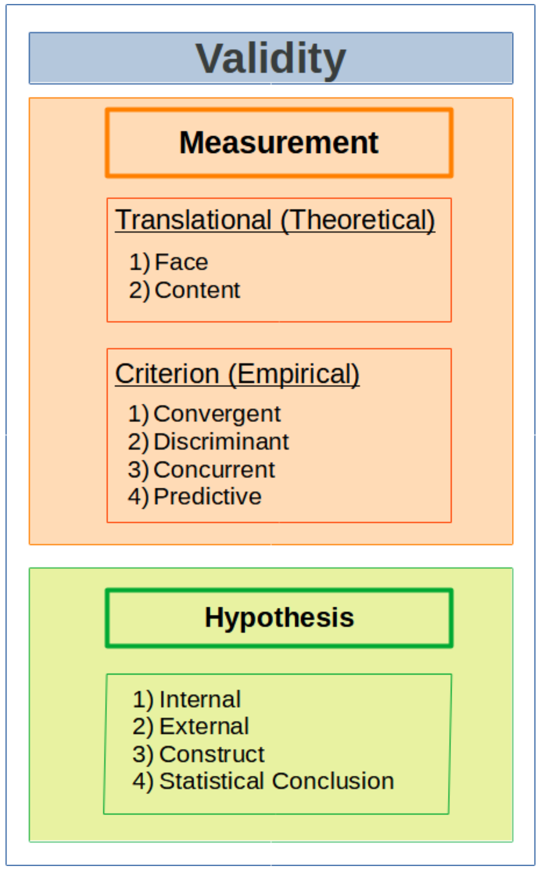


FIGURE 23: TYPES OF VALIDITY

MEASUREMENT VALIDITY

The theoretical assessment of validity focuses on how well an abstract construct is translated into an operational measure. This type of validity is called translational validity and is divided into two sub-types: face and content validity. Translational validity is typically assessed using a panel of expert judges who rate each item (indicator) on how well it fits the conceptual definition of that construct along with a qualitative technique called Q-method, as explained by Pnina Shinebourne (Shinebourne, 2009).

The empirical assessment of validity examines how well a given measure relates to one or more external criteria based on empirical observations. This type of validity is called criterion validity, divided into four sub-types: convergent, discriminant, concurrent, and predictive. While translation validity examines whether a measure is a good reflection of its underlying construct, criterion validity examines whether a given measure behaves the way it should, given the theory behind that construct.

Translational Validity

Face validity refers to whether an indicator seems to be a reasonable measure of its underlying construct "on its face." For instance, the frequency of attendance at religious services seems to make sense as an indication of a person's religiosity without much explanation. Hence this indicator has face validity. However, if it were suggested that the number of books checked out of an office library is a measure of employee morale, such a measure would probably lack face validity because it does not seem to be related to morale. Interestingly, some of the famous measures used in organizational research appear to lack face validity, though they may have other types of validity. For instance, the absorptive capacity of an organization (how much new knowledge can it assimilate for improving organizational processes) has often been measured by research and development intensity (i.e., R&D expenses divided by gross revenues). Research that includes highly abstract constructs or is hard to conceptually separate from each other (e.g., compassion and empathy) may benefit from a panel of experts who can evaluate the face validity of the measures.

Content validity is an assessment of how well a measure matches the content domain of the construct being measured. For instance, to measure the construct "satisfaction with restaurant service," then the content domain should include variables like the quality of food, courtesy of wait staff, duration of the wait, and the restaurant's overall ambiance (factors like the noise level). Of course, this approach assumes the researcher can create a detailed description of the content domain, which may be difficult for complex constructs such as self-esteem or intelligence. As with face validity, an expert panel of judges may be employed to examine the content validity of constructs.

Criterion Validity

Convergent validity is how close a measure relates to (or converges on) the construct it purports to measure. Convergent validity can be established by comparing the observed values of one indicator with other indicators to find a high correlation between those indicators. Compare this to discriminant validity.

Discriminant validity refers to how a measure does not measure (or discriminates from) other constructs that it is not supposed to measure. Usually, convergent validity and discriminant validity are assessed jointly for a set of related constructs. For example, if an organization's knowledge is related to its performance, organizational knowledge must be measured (convergent validity), not organizational performance (discriminant validity). Discriminant validity is established by demonstrating that indicators of one construct are dissimilar from (i.e., have low correlation with) other constructs.

Concurrent validity examines how well a measure of one outcome relates to another outcome presumed to coincide. For instance, do students' scores in a calculus class correlate well with their scores in a linear algebra class? Since both are mathematics classes, it would be presumed that there is high concurrent validity between scores in those classes.

Predictive validity is the degree to which a measure successfully predicts a future outcome. For example, a standardized test score (e. g., Scholastic Aptitude Test) can be used to predict a student's academic success in college. Concurrent and predictive validity are not often considered in empirical business research.

HYPOTHESIS VALIDITY

In general, four types of hypothesis validity are referred to in the literature.

Internal validity examines whether the observed change in a dependent variable is caused by a corresponding change in the hypothesized independent variable and not extraneous variables to the research context. This type of validity is sometimes called "causality," and it requires three conditions.

1. Covariation of cause and effect (if the cause happens, then the effect also happens; and if the cause does not happen, the effect does not happen)
2. Temporal precedence (the cause must precede the effect in time)
3. Lack of plausible alternative explanation (or spurious correlation).

Specific research designs, such as laboratory experiments, have compelling internal validity since researchers can manipulate the independent variable (cause) using a specific treatment and observe the effect (dependent variable) of that treatment. Other designs, such as field surveys, have poor internal validity because researchers cannot manipulate the independent variable (cause). If cause and effect are measured at the same time, temporal precedence is defeated.

External validity refers to whether the observed associations can be generalized from the sample to the population (population validity) or entities outside the population (ecological validity). For example, if results drawn from a sample of financial firms in the United States can be generalized to all financial firms, it would have strong population validity. If those results can be generalized to other types of firms, it would have strong ecological validity. Survey research, where data are sourced from various individuals or firms, tends to be more generalizable than laboratory experiments, where extraneous variables are controlled, so the findings are tightly focused.[17]

Construct validity examines how well a given measurement scale measures the theoretical construct it is designed to measure. One frequent problem with construct validity is simply defining the construct so that it is measurable. For example, “property ownership” is a construct of a market economy explained by Robert Reich (Reich, 2016). That is, the fact that people can own property drives a local economy. Nevertheless, this construct relies on several external forces that cannot be controlled, such as local politics (a city’s eminent domain can take a person’s property) and the property's value on the open market. Measuring the effect of property ownership in a local economy (the construct) would be very difficult since there are many confounding variables.

Statistical conclusion validity examines the extent to which conclusions derived from a statistical procedure are valid. For example, it examines whether a suitable statistical method was used and whether the variables meet the statistical test's assumptions (such as sample size or distributional requirements).

Improving Internal and External Validity

The best research designs are those that can assure high levels of internal and external validity. Such designs would guard against spurious correlations, inspire greater faith in the hypotheses testing, and ensure that the results drawn from a small sample are generalizable to the population at large. The internal validity of research designs and can be improved using four methods.

1. *Manipulation* involves the researcher manipulating the independent variables in one or more ways (called “treatments”) and comparing the effects of the treatments against a control group where subjects do not receive the treatment. Treatments may include a new drug or a different dosage of the drug (for treating a medical condition), a new teaching style (for education), and so forth. This type of control can be achieved in experimental or quasi-experimental designs but not in non-experimental designs such as surveys.
2. *Elimination* relies on eliminating extraneous variables by holding them constant across treatments, like restricting the study to a single gender or socioeconomic status.

[17] Some researchers claim that increased external validity leads to decreased internal validity and vice-versa, but this is not always true. Some research designs, such as multiple case studies, have high degrees of both internal and external validities.

3. *Inclusion* is the process of separately estimating the effects of spurious variables on the dependent variable. As an example, consider the process of estimating the effect of gender on a marketing study. Inclusion techniques allow for greater generalizability of the study but also require substantially larger samples.
4. *Randomization* is aimed at canceling out the effects of extraneous variables through a process of random sampling. Two types of randomization are 1) random selection, where a sample is selected randomly from a population, and 2) random assignment, where subjects selected in a non-random manner are randomly assigned to treatment groups. Randomization improves external validity since inferences can be generalized to the population from which the sample is drawn. However, generalizability across different populations is harder to ascertain since populations may differ on multiple dimensions.[18]

Validity in Qualitative Research

The validity measures in this chapter have been designed primarily for quantitative research. For many years researchers have had concerns with the validity of qualitative studies. Researchers engaged in qualitative research projects often assure "validity" using some tortured definition of validity used in a quantitative project. Cho and Trent revisited the concept of validity in qualitative research and proposed a recursive, process-oriented view of validity (Cho & Trent, 2006).

The researchers first define two current qualitative validity standards.

1. *Transactional*. This type of validity is an interactive process between the researcher, the researched, and the collected data to achieve a relatively higher level of accuracy and consensus. This approach assumes that qualitative research can be more credible if specific techniques, methods, or strategies are employed during the inquiry. The significant ways this validity is performed is with "member checking" (reassuring the credibility of the participants) and triangulation (verifying facts using multiple data sources).
2. *Transformational*. This type of validity is a progressive process leading toward social change achieved by the research endeavor itself. It acknowledges that qualitative researchers emphasize the value-laden nature of social interactions, and multiple perspectives on a topic would yield multiple meanings. Qualitative researchers assert that a positivist inquiry is not absolute truth but merely one facet of the truth. The validity of this type of study is found

[18] The concept of random sampling is covered more thoroughly in Chapter 7: Sampling.

through self-reflection, where the researchers deconstruct/reconstruct the meanings attached to their findings to make them more fruitful.

Cho and Trent then argue that the traditional quantitative approaches are not adequate measures of validity for qualitative research projects and propose, instead, to extend Donmoyer's framework of "five overarching purposes undergirding contemporary qualitative research." (Donmoyer, 2001)

- **Truth-seeking**. Finding "truth" is perhaps the goal for any research project, but Cho and Trent maintain that there is no single measure that will validate the truthfulness of findings. Instead, they propose that the "...'truth' seeking purpose is progressive induction through which data need to be collected, analyzed, interpreted, triangulated, and thus represent 'what is' through a credible, corresponding account." They believe that member checking and triangulation, as described above, fill this purpose.
- **Thick description**. Qualitative research often does not intend to find "truth" as in a quantitative project; instead, the goal is to explain the unique perspectives constructed by individuals and groups. Thus, the outcome for many qualitative research projects is a "thick description" of some phenomenon in the form of a case study or ethnography. Validity is achieved through holistic and prolonged engagement between the researcher and the subject being studied. A study's validity is assured from the extent to which data are presented ("let readers see for themselves") and the researcher's competence in interpreting what was found.
- **Developmental**. Often, a qualitative research project concerns observing the shared interests of individuals or groups as those interests develop over time. These types of projects may be little more than a comprehensive investigation of what happens over time or as involved as explaining how things make progress stage by stage. Validity is an ongoing process as the project unfolds rather than a single event completed at the end of the project.
- **Personal essay**. This type of qualitative research project is like a thick description but is intentionally and openly subjective. The researcher tells a personal story and reflects on lessons learned in that process. Validity is gained through the expertise of the researcher in the field being explored. This type of research is sometimes called "autoethnography." For example, a research project where professional actors reflect on the art of acting would have a high degree of validity.
- **Praxis/social change**. Some qualitative research projects include as a goal the desire to bring about change. This goal is especially true in "action research" projects, as covered in Chapter 13.

> Action research is more frequently found in educational than business settings. Validity is assured through three methods: member checks, researchers thinking critically about their involvement in the project, and whether the status quo was redefined.

Cho and Trent close their work by describing the holistic requirement of validating a study. They believe that it is not adequate to provide member checks and triangulation, relatively simply, ". . . we are proposing an inclusive discourse of validity in qualitative research by reflecting on what matters specific to the problem/research within our research purview." They believe that a narrative that explains the approaches to validity taken by the researchers is best so readers can judge for themselves if the research is valid. Their conclusion includes this statement:

> Many are interested in creating new senses of validity in qualitative research. Some hope that validity methods and strategies in qualitative research can be defined to the extent that they are equal to those of conventional or quantitative research methods. We believe this to be a misguided aim. We add that a variety of purposeful approaches (e.g., thick description and praxis/social change) may be combined to obtain holistically 'valid' results.

Figure 24 is based on a similar diagram presented by Cho and Trent in their report.

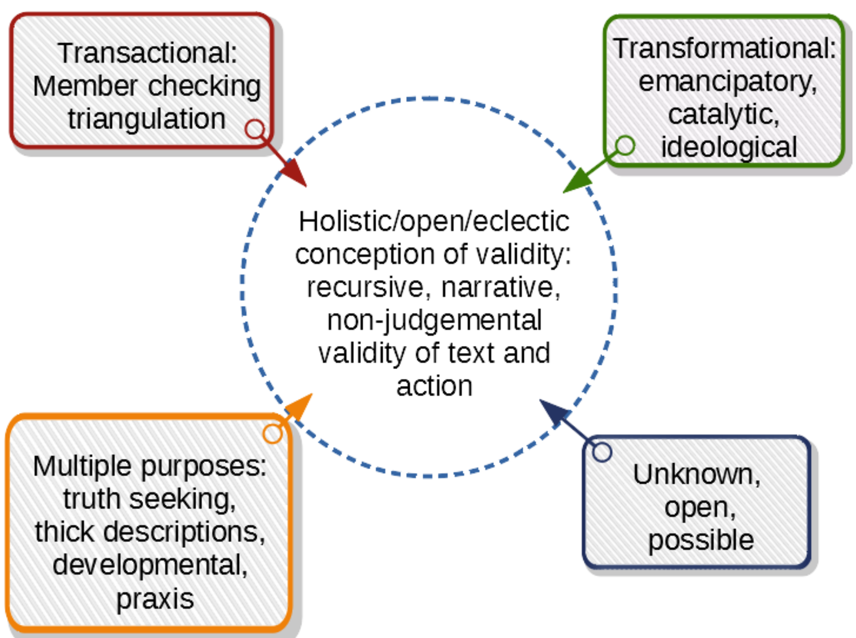
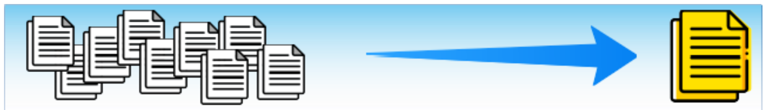


FIGURE 24: VALIDITY FOR QUALITATIVE RESEARCH



A Sampling of Research

Consumer Self-Confidence

Bearden, Hardesty, and Rose investigated consumer self-confidence (Bearden, Hardesty, & Rose, 2001). Their research project aimed to refine the conceptualization and measurement of consumer self-confidence to understand that concept better. Consumer self-confidence is "...the extent to which an individual feels capable and assured concerning his or her marketplace decisions and behaviors." The researchers conceptualized consumer self-confidence as two higher-order factors that are each made up of several dimensions.

- Decision-Making Self-Confidence
    - Information acquisition and processing. Assessed with statements like, "I know where to find the information I need before making a purchase" and "I know the right questions to ask when shopping."
    - Consideration-Set formation. Assessed with statements like, "I can tell which brands meet my expectations" and "I know which stores to shop."
    - Personal outcomes. Assessed with statements like, "I frequently agonize over what to buy" and "I never seem to buy the right thing for me."
    - Social outcomes. Assessed with statements like, "I can give good presents" and "I impress other people with the purchases I make."
- Protection
    - Persuasion Knowledge. Assessed with statements like, "I know when an offer is 'too good to be true"' and "I can separate fact from fantasy in advertising."
    - Marketplace Interfaces. Assessed with statements like, "I am hesitant to complain when shopping" and "I have a hard time saying no to a salesperson."

After the conceptualization and measurement phase, the researchers completed seven different studies designed to affirm the reliability and validity of their work. The first step was to create a pool of 145 items generated from exploratory interviews with 43 adult consumers. The items were screened to reduce redundancy, and the remaining items were given to an expert panel of 14 marketing faculty members in a major university. That reduced the pool of statements to 97.

The pool of 97 statements was used as a survey for two different studies, one to 221 and the other to 204 adults. After those data were gathered, factor analysis was used to eliminate items determined to not contribute significantly to the concept or were unclear. This analysis reduced the pool to 39 items.

A third study was completed where the 39 items were administered to 252 undergraduate business students. The study results indicated that a two-factor higher-order model with six-factor dimensions (as noted above) best fit the data. At this point, the researchers determined that the reliability and discriminant validity were both high.

The researchers tested their model for test-retest reliability, convergent validity, and relative predictive validity for studies four and five. They found the model performed well on each of these criteria. Study six was designed to test the model to see if it would detect known group differences, and it did.

One final study was conducted where the model was tested with 106 faculty and staff members from a large state university. They were asked to decide which of several competing products they would purchase based on price and perceived quality. The researchers found that consumers who had higher self-confidence tended to select the more expensive product but chose to defer the purchase when the self-confidence level was lower.

This study is an excellent example of how researchers generate a concept and determine how to measure that concept. That measurement is then subjected to extensive testing to ensure its reliability and validity.

Resolving Contractual Breaches

Johnson and Sohi (Johnson & Sohi, 2016) studied how contractual breaches between buyers and sellers were resolved when legal enforcement was not desirable. They used a grounded theory (qualitative) approach to develop a model that describes how these disputes are resolved. For this study, they interviewed 40 supplier managers and executives in multiple industries. They identified several types of resolution alternatives used and the factors that lead to selecting any given alternative. The following were the types of resolutions they identified.

- *Integrative alternatives*. This resolution is a negotiation between the buyer and seller where some consideration, like incremental competitive business, is offered in place of legal action. The second form of this resolution extends the contract into an unrelated non-competitive bid contract.

- *Change the terms*. The initial contract is rewritten to make the terms more acceptable to both parties.
- *Resource adjustment*. Occasionally, the buyer may demand a resource adjustment to help resolve the breach. For example, buyers may demand that the seller adjust the contract to include free on-site maintenance of the purchased machinery.
- *Sell unused capacity*. The seller may have overstock or other unused capacity and offer that to help settle a contract.
- *Split payments*. The seller may offer to extend the payment period and split payments to make it easier for the buyer to complete the purchase.
- *Raise prices on other transactions*. The seller may raise prices on other seller business transactions not related to the breach of contract.
- *Quid pro quo*. In some cases, large complex contracts could involve breaches by both the seller and buyer, and they could mutually agree to ignore those breaches to complete the contract.

The research also identified external and internal factors that may figure into the resolution. As an example, the overall economy, an external factor, may temper whatever action is taken to resolve the breach. Finally, they used five different methods to enhance the validity of their study.

- *Comprehensive data treatment*. All interviews were recorded, transcribed, and entered for computer analysis. By analyzing all collected data, the goal is to decrease the potential for bias by omission.
- *Refutability*. They actively sought cases that were inconsistent with the findings. While there were a few cases where companies used unique resolution methods, the researchers found no systematic patterns, like business size, for these cases.
- *Constant comparison*. As the researchers uncovered insights during interviews, those were used to guide subsequent interview questions. This process was repeated until they reached theoretical saturation; that is, no new insights were uncovered.
- *Respondent validation (member checking)*. The researchers took their data and interpretation back to the participants so they could confirm the information. This process led to the decreased potential that the researchers could misinterpret an interview response.
- *Reflexivity*. The researchers invited an external expert with knowledge in the topic to review the project and the questions asked during interviews. This process is called “peer debriefing” and

improves the validity of the project. Also, two independent judges evaluated the coding of the interviews, and their agreement with the researchers' coding was an indicator of high reliability.

Key Takeaways

Chapter 5: Defining and Measuring Concepts
• Define measurement. • Distinguish between conceptualization and operationalization. • Define reliability and validity. • Compare and contrast reliability and validity. • Suggest ways to improve reliability and validity in research projects.

6: Data

Data are a collection of facts about some topic. For example, a "customer loyalty" program gathers data from customers on how often they shop, what they purchase on each trip, what time of day they typically shop, and all sorts of other data. When data are interpreted in some way, they become information. The types of data involved limit the types of analyses that can be done. This chapter aims to introduce various concepts about data and show how they can be analyzed.

OBJECTIVES

- Define the types of data.
- Determine the difference between data and rating scales.
- Define the three primary properties of data: distribution, excess kurtosis, and skew.
- Define "database" and find several public databases that may be useful for research projects.
- Select the correct statistical test for the project.
- Define the statistical tests used for central measures, spread, frequency, correlation, parametric hypothesis testing, and nonparametric hypothesis testing.
- Define data mining and discuss the three most common forms of data mining: clustering, decision trees, and market basket.

Types of Data

Psychologist Stanley Smith Stevens defined four generic types of data divided into two groups (Stevens & others, 1946).

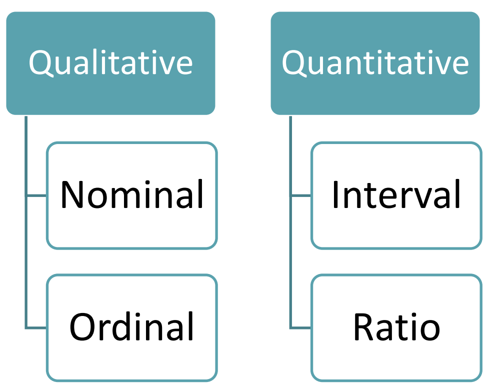


Figure 25: Data Groups and Types

- Qualitative data group observations into a limited number of categories, for example, type of pet (cat, dog, bird) or place of residence (Arizona, California). Qualitative data are analyzed using non-parametric tests, like Kruskal-Wallis H and Mann-Whitney U, since it does not have quantitative characteristics like means or standard deviations. Qualitative data can be further divided into two sub-types, nominal and ordinal.

- Nominal data contain categories that do not overlap and have no meaningful order; they are merely labels for attributes. Examples of nominal data include occupations (custodial, accounting, sales) and blood type (A, B, AB, O). A particular type of nominal data is binary, or dichotomous, where there are only two possible responses, like "yes" and "no." Nominal data are sometimes stored in a database using numbers, but they cannot be treated like numeric data. For example, binary data, like "Do you rent or own your home?" can be stored as "1 = rent, 2 = own," but the numbers, in this case, have no numeric significance and could be replaced by words like "Rent" and "Own."
- Ordinal data are categorical data, but, unlike nominal, the categories imply some sort of order (which is why it is called "ordinal"). One example of ordinal data is the "star" rating system for movies. A five-star movie is somehow better than a four-star movie, but there is no way to quantify the difference between those two categories. As another example, it is common for hospital staff members to ask patients to rate their pain levels on a scale of one to ten. If a patient reports a pain level of "seven" but after treatment later reports a pain level of "five," then the pain has decreased. However, it would be impossible to quantify the exact difference between those two levels. Ordinal scales are most used with Likert-type survey questions where the responses are selections like "Strongly Agree," "Agree," "Neutral," "Disagree," "Strongly Disagree." Ordinal data are also used when numeric data are grouped. For example, if a dataset included respondents' ages, those numbers could be grouped into categories like "20−29" and "30−39." Those groups would typically be stored in the dataset as a single number, so maybe "2" would represent the ages "20−29," making it ordinal data.

- Quantitative data are numeric, typically counts or measures, like a person's age, a tree's height, or a truck's weight. Quantitative data are measured with scales with equal divisions, so any two values can be compared. Quantitative data are discrete when represented by integers, like the count of words in a document, or continuous when represented by fractional numbers, like a person's height. Because quantitative data include characteristics like means and standard deviations, it is analyzed using parametric tests, like T-tests and Analysis of Variance (ANOVA). Quantitative data can be further divided into two sub-types, interval and ratio.

- Interval data use numbers to represent quantities where the distance between any two quantities can be calculated. One example is a temperature scale where the difference between 80°and 90°is the same as the difference between 60°and 70°. It is important to note that interval data do not include an actual zero point; thus, zero degrees Celsius does not mean "no temperature." Without a zero point, it is not reasonable to make a statement like 20° is twice as hot as 10°.[19]
- Ratio data use numbers to describe a specific, measurable distance between two quantities; however, unlike interval data, ratio data have an actual zero point. An excellent example of ratio data is the sales report for an automobile dealership. Because the data are a simple count of the number of automobiles sold, it is possible to compare one month with another. Also, since the scale has a zero point, no sales for the month, it is possible to state that one month had twice the sales of another.

Rating Scale

Researchers need to determine a rating scale, also called levels of measure, to record data gathered about an attribute when working with qualitative data. For example, male-female-other, M-F-O, and 1-2-3 are three potential rating scales for a gender attribute. A researcher could use any of these scales or devise a completely different one if the scale is used consistently throughout the entire research project. It is easy to imagine that many rating scales exist, but the most common ones are binary, Likert, semantic differential, and Guttman.

[19] To be precise, temperature does have a zero point, called "absolute zero." However, the value of zero degrees Celsius or Fahrenheit is not "no temperature."

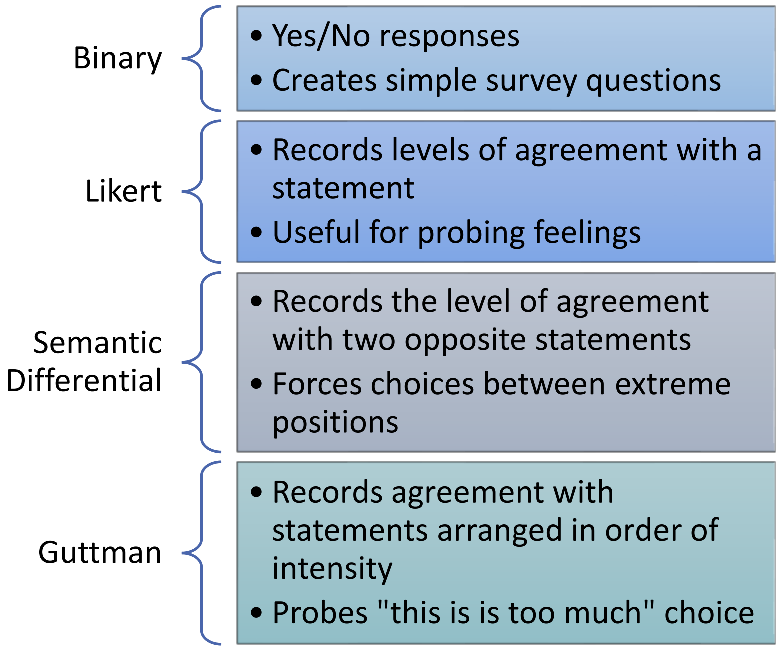


FIGURE 26: RATING SCALES

Binary. Binary scales are nominal scales consisting of binary items that assume one of only two possible values, such as yes or no, true or false, and so on. For example, a typical binary scale for a "political activism" construct may consist of the six binary items shown in Table 2: Political Activism Binary Scale. Each item in this scale is a binary item. The total number of "yes" selections, a value from 0 to 6, can be used to measure political activism. Binary scales can also employ other values, such as male or female for gender and full-time or part-time employment status. If an employment status item is modified to allow for more than two possible values (e.g., unemployed, full-time, part-time, and retired), it is no longer binary but remains a nominal item.

Question	Yes	No
Have you ever written a letter to a public official?	○	○
Have you ever signed a political petition?	○	○
Have you ever donated money to a political cause?	○	○
Have you ever donated money to a candidate running for public office?	○	○
Have you ever written a political letter to the editor of a newspaper?	○	○
Have you ever persuaded someone to change his/her voting plans?	○	○

TABLE 2: POLITICAL ACTIVISM BINARY SCALE

Likert. Designed by Rensis Likert, this is a prevalent rating scale for measuring ordinal data in business research. This scale includes Likert items that are worded statements to which respondents can indicate their extent of agreement or disagreement on a five or seven-point scale ranging from "strongly disagree" to "strongly agree." The values used on the scale are sometimes called "Anchors." A typical example of a six-item Likert scale for the "employment self-esteem" construct is shown in Table 3: Likert Scale for Employee Self-Esteem. Likert scales are summed scales; that is, the overall scale score may be a summation of the attribute values of each item as selected by a respondent.

Statement	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I feel good about my job.	○	○	○	○	○
I get along well with others at work.	○	○	○	○	○
I am proud of my relationship with my supervisor.	○	○	○	○	○
I feel like I am contributing to work.	○	○	○	○	○
I can tell that my coworkers respect me.	○	○	○	○	○

TABLE 3: LIKERT SCALE FOR EMPLOYEE SELF-ESTEEM

Likert items allow for more granularity (more finely tuned response) than binary items, including whether respondents are neutral to the statement. Three or nine values may also be used, but it is essential to use an odd number of values to allow for a "neutral" (or "neither agree nor disagree") anchor. Some studies have used a "forced choice approach" to force respondents to agree or disagree with the Likert statement by dropping the neutral mid-point and using an even number of values. Using forced choice may not be a good strategy because some people may be neutral to a given statement, and this approach does not allow them to record their neutral stance. A vital characteristic of a Likert scale is that even though the statements vary in different items or indicators, the anchors ("strongly disagree" to "strongly agree") remain the same. Likert scales are ordinal scales because the anchors are not necessarily equidistant, even though sometimes they are treated like interval scales.

Semantic Differential. This approach is a multi-item scale where respondents are asked to indicate their opinions or feelings toward a single statement using different adjectives framed as opposites. For instance, the construct "attitude toward health insurance" can be measured using three items shown in Table 4: Semantic Differential Scale. As in the Likert scale, the overall scale score may be a summation of individual item scores. Notice that in Likert scales, the statement changes, but the anchors remain the same across items. However, the statement remains constant in semantic differential scales, while the anchors (adjective pairs) change across items.

A semantic differential scale is an excellent technique for measuring people's attitudes or feelings toward objects, events, or behaviors.

How would you rate your opinion on health insurance?						
	Very Much	Much	Neutral	Much	Very Much	
Good	○	○	○	○	○	Bad
Useful	○	○	○	○	○	Useless
Caring	○	○	○	○	○	Uncaring

TABLE 4: SEMANTIC DIFFERENTIAL SCALE

Guttman. Designed by Louis Guttman, this composite scale uses a series of items arranged to increase the intensity of the construct of interest, from least intense to most. For example, the construct "attitude toward immigrants" can be measured using five items shown in Table 5: Guttman Scale. Each statement is weighted (not indicated on the scale) to indicate the intensity of that item. The weighted combination of each response is used as an aggregate measure of an observation.

How do you rate your opinion on the following statements about immigrants?		
Do you mind immigrants being citizens of your country?	Yes	No
Do you mind immigrants living in your neighborhood?	Yes	No
Would you mind living next door to an immigrant?	Yes	No
Would you mind having an immigrant as your close friend?	Yes	No
Would you mind if someone in your family married an immigrant?	Yes	No

TABLE 5: GUTTMAN SCALE

Properties of Data

About the Normal Distribution (Bell Curve)

When the quantitative data gathered from some statistical project is plotted on a graph, they often form a normal distribution (sometimes called a "bell curve" due to its shape). For example, consider the Scholastic Aptitude Test (SAT), administered to more than 1.5 million high school students every year. Figure 27 was created with fake data but illustrated the results expected of a typical SAT administration.

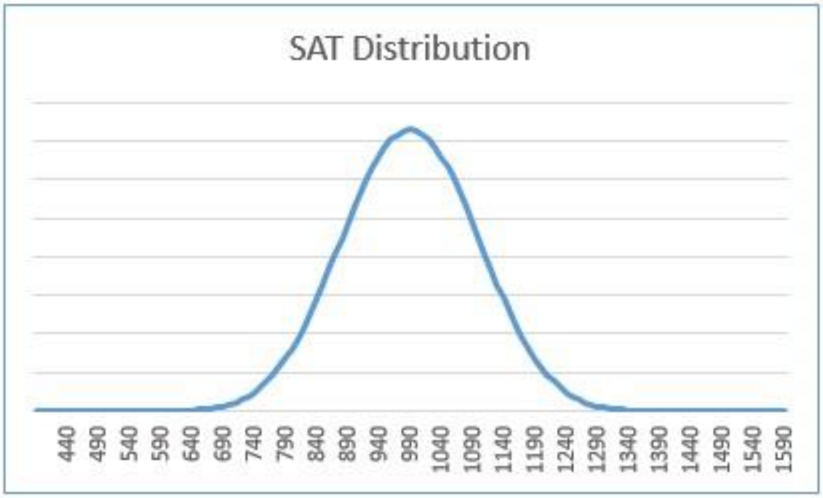


FIGURE 27: NORMAL DISTRIBUTION

SAT scores are between 400 and 1600 as listed across the X-Axis. The number of students who achieve each score is plotted on the Y-Axis. Since the most common score is 1000, that is at the peak of the curve. Remarkably few students scored below 650 or above 1300, and the curve is near the lower bound beyond those points. Figure 27 illustrates a normal distribution where most scores are bunched near the center of the graph, and only a few are at the extremes.

The normal distribution is important because it permits researchers to use specific statistical techniques to test a hypothesis. For example, perhaps a researcher hypothesized that the graduation rate at university "A" would be higher than at university "B" because students' SAT scores were higher at university "A." Since SAT scores have a normal distribution, the researcher could use statistical tests, like a t-test, to support or refute the hypothesis. However, if the data were not normally distributed, the researcher would need to use a different method to compare the students.

Excess Kurtosis

One way to mathematically describe a normal distribution is to calculate the length of the bell curve's tails, called its excess kurtosis. For a normal distribution, the excess kurtosis is 0.00; a positive value indicates long tails, while a negative value indicates short tails. Intuitively, many people believe the excess kurtosis represents the "peakedness" of the curve since longer tails would tend to lead to a more peaked graph. However, excess kurtosis measures the data outliers, which would only be present in the graph's tails. Therefore, it is difficult to categorically state that some level of excess kurtosis is good or bad. A graph with longer tails is desired in some cases, but that would indicate a problem in other cases.

Following are four examples of excess kurtosis. Notice that as the excess kurtosis increases, the tails become longer.

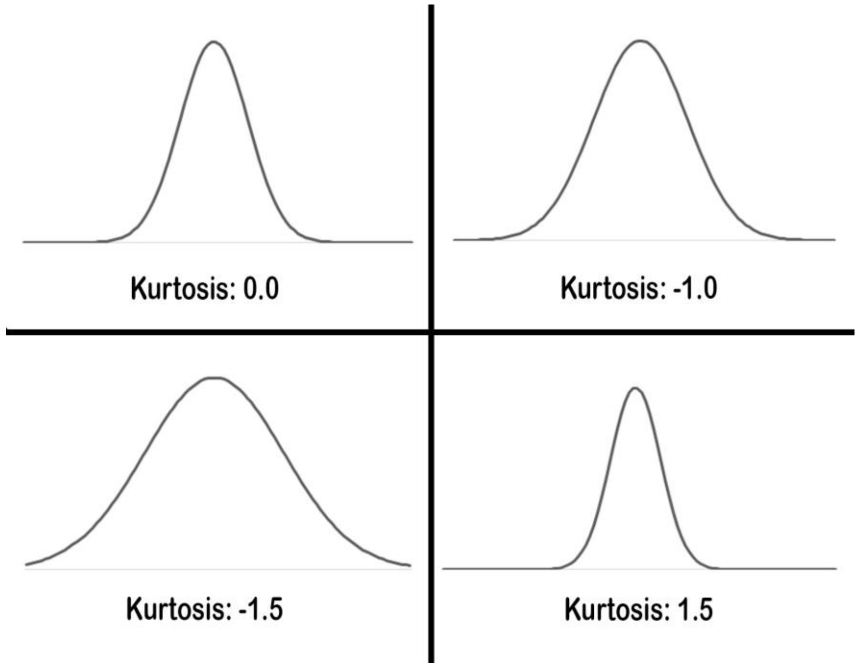


FIGURE 28: KURTOSIS IN A NORMAL DISTRIBUTION

Skew

The second numerical measure of a normal distribution is its skew, which measures the symmetry of the curve about its mean. For example, the normal distribution in Figure 27 has a skew of 0.00. A positive skew indicates that the tail on the right side is longer, which means that several data points are on the far-right side of the graph, “pulling” the tail out that direction. A negative skew indicates that the tail on the left side of the graph is longer. Following are four examples of skew.

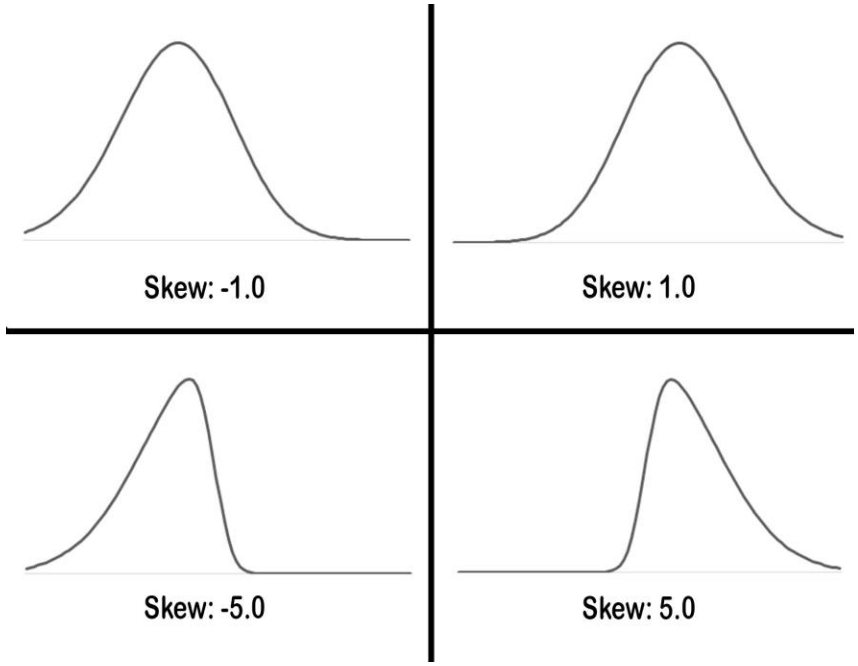


FIGURE 29: SKEW IN A NORMAL DISTRIBUTION

Databases

When many data points are gathered into a single location, it is referred to as a database. A database is not a new concept. Fifty years ago, a library would maintain hundreds of 3X5 cards that contained information about the books in the library (like the title, author, and subject). Those cards were stored in a wooden cabinet called the "card catalog," and customers could find information about whatever book they wanted. Today, databases are often contained in electronic form on the Internet to be accessed from customers' home computers, tablets, or even phones.

Data in a database are typically stored in tables that resemble spreadsheet rows and columns where each row is one record (or observation) about some phenomenon. Each column is one descriptor of that record, maybe a color attribute. For example, imagine a database containing information about the people who work at a particular company. Each row would contain data about one person, and each column would contain a single aspect of that person's employment, like *name*.[20] A database is designed to deliver answers to questions through a lookup process. For example, if the CEO of a company wanted to know the birth date of someone in the accounting department, those data could be easily found.

[20] Of course, databases are much more complex than described in this paragraph suggests, but this is not a text about databases, so the simple explanation offered is adequate for this context.

One common problem with a database is "dirty data," which contains missing values or errors. For example, it is easy for a data entry clerk to enter something like "1000000" instead of "100000" (notice the number zeros in each number) for a person's salary and create an "outlier" in the data. Another common problem are missing data. For example, if employees are asked to update their personal information but a few people could not remember the date they joined the company, they would leave that field blank and create missing data.

Dirty data makes it challenging to analyze the database. For example, if a researcher wanted to report the median salary for the workers in a factory, but ten percent of the salaries were missing from the database, the median would not be accurate. There are several methods statisticians use to mitigate the problems caused by dirty data, but those are beyond the scope of this text.

Public Databases

There are hundreds of publicly available databases that can be used for research. As one example, the United States Census Bureau maintains a vast database that contains information about the people of the United States.[21] The data at that site is freely available to anyone who wants to use it, and the site is organized in a way that makes the information is quick and easy to find. As an example, it is not difficult to discover that among adults in the United States, 28% have a high school diploma, 4% have an associate degree, 22% have a bachelor's degree, 9% have a master's degree, 2% have a Doctoral degree, and the rest fall elsewhere on the education spectrum[22]. The US Census Bureau has advanced tools available that permit researchers to focus their search significantly.

When using a public database, researchers must be concerned with bias. For example, if the database includes people's attitudes towards work, how will a researcher know if the gathered data were from a well-designed, neutral survey or just a convenience sample? In general, databases found at governmental websites (with URLs that end with .gov) or education websites (with URLs that end with .edu) would more likely be bias-free. In contrast, databases from .com sites would need to be carefully scrutinized.

Sometimes, students will find a website with links to journal articles or chapters from books. While these are valuable resources for a researcher, they are not the same as a database that contains raw

[21] The US Census Bureau's website is at https://www.census.gov/.
[22] These values are based on 2020 data.

data from a survey, experiment, or other activity. Journal articles provide good information for a literature review but would not be appropriate for an online database source.

Using Public Databases

As an example of using a public database, imagine that the CEO of "BASVFOODS" is interested in opening a neighborhood market in a small town they have never serviced before. The CEO can gather data from the US Census Bureau to determine things like the new site's median household size, income, and education level. The CEO could then compare those data with similar data from a town with a successful store to help inform the decision.

Statistical Test Selection

Once the data are gathered, it is imperative to execute appropriate statistical processes to see if they contain anything of interest. There are hundreds of tests that can be used, and researchers must consider both the goal of the analysis and the type of data being analyzed. The first step is to classify the variables being evaluated.

Quantitative data are numeric and generated through measurement or counting. Continuous data can be any value, including fractions or decimals, like a person's height or the length of time it takes to complete a task. Discrete data are commonly found by counting, like the number of people at an event.

Qualitative data are non-numeric and are often generated by people checking boxes on a survey. Ordinal data have some implied order, like a student's class (like senior or junior), and nominal data have no order, like a respondent's gender.

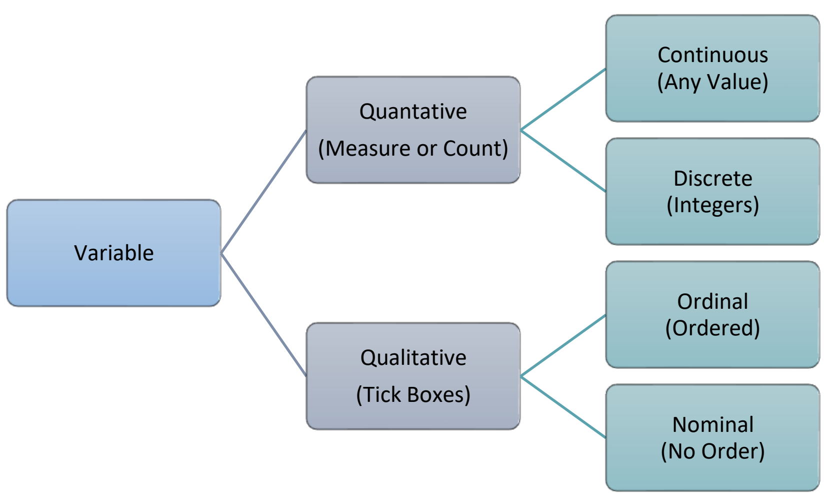


FIGURE 30: TYPES OF DATA

Once researchers know the types of variables they are working with, the statistical tests they can use become evident.

Summary Statistics

One of the most accessible types of analysis to complete is determining the center and spread of a dataset. This analysis is also one of the first that research readers expect to find. Here is a guide for what sort of central measure and spread to report.

- For normally distributed continuous data, use the mean and standard deviation.
- For skewed continuous or discrete data, use the median and Interquartile Range (IQR).
- For ordinal data, use the median and Interquartile Range (IQR).
- For nominal data, use the mode as a central measure, but there is no spread.

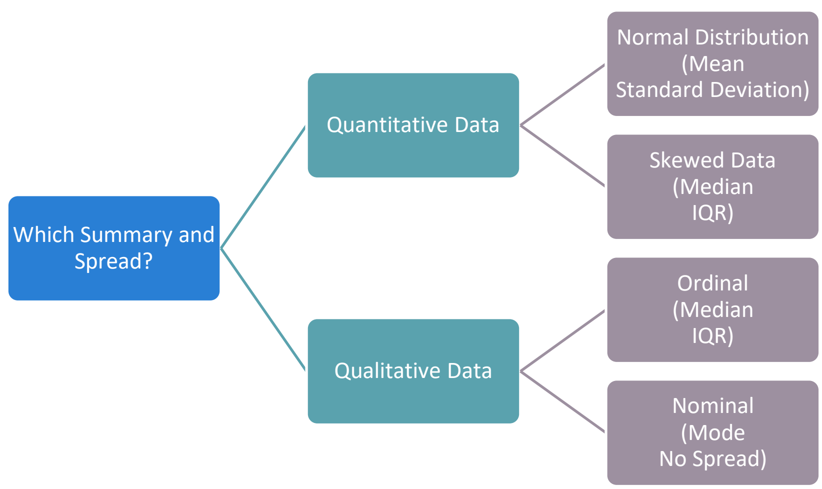


FIGURE 31: CENTRAL MEASURE AND SPREAD

Parametric vs. Nonparametric Tests

Typically, a researcher posits a hypothesis and then conducts research to either support or refute that hypothesis. The type of hypothesis test needed depends on the type of variables being tested. There are two broad categories of variables in a hypothesis test: independent and dependent.

Independent variables (explanatory variables) cause something to happen; they explain why some outcome was observed. For example, if the hypothesis is that women spend more on groceries than men, then the independent variable is the sex of the shopper. If the hypothesis is that elderly drivers are more dangerous drivers than younger drivers, then the independent variable is the driver's age.

Dependent variables (outcome variables) are those that are the outcome for whatever is being observed. For example, if the hypothesis is that women spend more on groceries than men, then the dependent variable is money spent. If the hypothesis is that elderly drivers are more dangerous than younger drivers, then the dependent variable is the number of reported accidents.

The chart in Figure 32 is used to determine if the hypothesis test should be parametric or nonparametric.

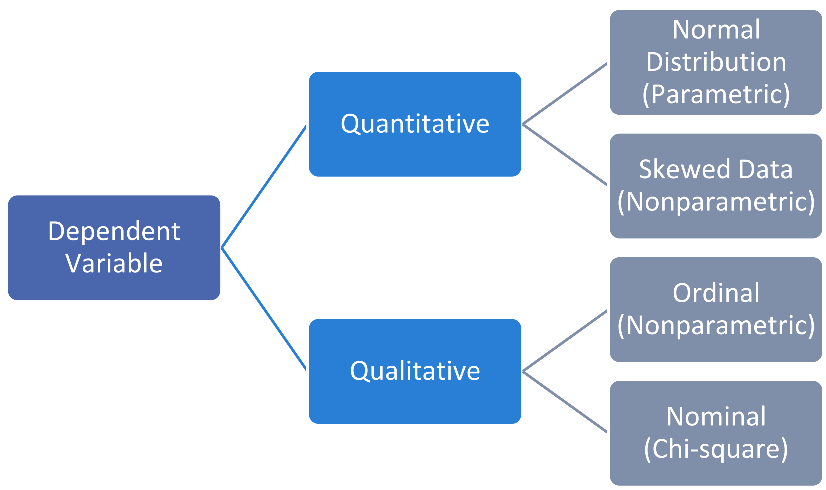


FIGURE 32: PARAMETRIC VS. NONPARAMETRIC SELECTION

Parametric tests assume that data have a normal distribution. Parametric tests are more powerful than nonparametric tests and are more likely to detect relationships or differences that exist.

Nonparametric tests are used when the data are not normally distributed; that is, the data are either skewed or qualitative. Nonparametric techniques are customarily based on ranks and are usually less powerful than parametric tests.

Chi-square is a test that is only used when the dependent variable is nominal. A chi-square test determines if there is a significant difference in the observed results compared to those expected. For example, if a coin were tossed 100 times, it would be expected to land "heads" half of the time. However, if the actual observation were that "heads" came up 75% of the time, then the chi-square statistic would indicate a significant difference between what was expected and observed.

Hypothesis Test Selection

Researchers often hypothesize that some treatment will lead to an outcome. To test that, they will apply the treatment to one group but not another and then compare the two groups to determine any difference. Figure 33 lists the statistical tests commonly used to compare the means of two or more groups. In the figure, "independent" groups have no overlapping members while "matched" groups test the same members two times. For example, suppose researchers wanted to know how if a movie affected people. In that case, they could survey people leaving two different theaters (independent groups) or survey the same people before and after seeing the movie (matched groups).

Comparing	Dep Var	Ind Var	Para	NonPara
2 independent groups	Quant	Binary	Indep t-test	Mann-Whitney
3+ independent groups	Quant	Nom	ANOVA	Kruskal-Wallis
2 matched groups	Quant	Time	Paired t-test	Wilcoxon
3+ measures same subject	Quant	Time	Rep ANOVA	Friedman

FIGURE 33: TESTS TO COMPARE TWO OR MORE SAMPLES

Another common research goal is to determine any association (or "correlation") between two or more variables. Further, a correlation may be able to predict an outcome for a new observation. For example, a researcher may run an experiment where students work with a tutor once a week and then detect improved test scores. In this case, "tutoring time" would be correlated with "test scores" and may be used to predict a new student's test score based on the amount of tutoring. Figure 34 lists the hypothesis tests that are commonly used to find correlations and predictions between two groups.

Comparing	Dep Var	Ind Var	Para	NonPara
2 Continuous	Quant	Quant	Pearson's r	Spearman's rho
Prediction	Quant	Any	Regression	None
Prediction	Nominal	Any	Log Regression	None
2 Qual Vars	Qual	Qual	None	Chi-square

FIGURE 34: TESTS OF ASSOCIATION BETWEEN TWO SAMPLES

Statistical Test Sampler

While there are hundreds of statistical procedures available, this section covers those that are commonly used.

Central Measures

Three different central measures are commonly used, depending on the type of data being summarized. Calculating these values is usually done with software since some of the calculations are complex.

- **Mean**. An arithmetic mean is calculated by adding all the terms and then dividing by the number of terms. This process is taught in elementary school as calculating the "average." However, if the terms have wildly different values, then a geometric mean is a better choice. A geometric mean multiplies the values together, and then the n-root is taken, where "n" is the number of terms in the data. Finally, if the mean of a series of rates is needed, then a harmonic mean is used. For a harmonic mean, all the terms are reciprocated, a mean is found of those reciprocated terms, and then that calculated mean is reciprocated.

- **Median**. A median is found by putting all the terms in numeric order and then selecting the middle term. A median is helpful if the dataset includes outliers, a few values far outside the other terms. Medians are frequently used to report house values since a few newer houses may be worth far more than the older houses in an area. If the dataset has an even number of terms, so there is no middle term, then the median is found by taking the mean of the middle two terms.
- **Mode**. The mode is used for nominal data and is the count of the most common term. For example, if a count of the number of undergraduate students in each class is made and it turns out that there are more seniors than any other group, then the mode would be "senior."

Spread

Two measures of spread are commonly used.

- **Range**. The range, or "dispersion," is the difference between the highest value and lowest value. For example, if a survey recorded respondents' ages and the maximum was 70 while the minimum was 30, the range would be 40.
- **Standard deviation**. This often-misunderstood value is nothing more than an indicator of how much variation exists in the data or how "scattered out" the values are. In general, the greater the standard deviation, the more variation there is in the data. For example, imagine a professor administered an examination and found the mean to be 70% with a standard deviation of 15. Then the professor changed something about the class the following semester, then administered the same examination and found the mean was 85% with a standard deviation of 5. This result would indicate that the second group of scores were generally higher and had much less variation, or they were grouped much "tighter," and that would likely be good news for the professor.

Frequency Tables

Discrete or qualitative data items are usually reported in frequency tables where the counts for a particular item are displayed. When a frequency table has two dimensions, it is usually called a "crosstab" or "pivot table." For example, Figure 35 contains an exit poll from the 2016 presidential election (Politics, 2016).

Party	Clinton	Trump	Other
Democrats	89%	8%	3%
Republicans	8%	88%	4%
Independents	42%	46%	12%

FIGURE 35: 2016 EXIT POLL

Correlation

Correlation describes a relationship between the independent (or X-Axis) and dependent (or Y-Axis) variables in a research project. A correlation is a number between −1.0 and +1.0 where the extremes indicate a more meaningful value. Thus, two variables with a correlation of +0.65 have a closer association than two variables with a correlation of +0.23. Correlations are often displayed in a matrix. Figure 36 shows the Miles per Gallon, Displacement, and Horsepower for a group of automobiles. This table shows a negative correlation between horsepower and miles per gallon, which means that cars with greater horsepower get fewer miles per gallon. On the other hand, there is a positive correlation between displacement and horsepower, so more engine displacement (larger engines) has higher horsepower.

	MPG	Disp	HP
MPG	+1.00	-0.85	-0.78
Disp	-0.85	+1.00	+0.79
HP	-0.78	+0.79	+1.00

FIGURE 36: SELECTED AUTOMOBILE CORRELATIONS

Parametric Hypothesis Tests

Some analysis techniques are only helpful with parametric data, and while there are dozens of statistical processes that will work with parametric data, two are seen most frequently.

- **T-test**. A t-test is used to analyze the difference in two groups of samples that are normally distributed. For example, a researcher may hypothesize a significant difference in the ages of people in two towns. Once people's ages have been recorded, a t-test could be used to detect a significant difference in the mean age for those two towns.

  There are two varieties of t-test commonly used, depending on the type of data being analyzed. If the two compared groups are independent, an "independent t-test" would be used. However, researchers often use the same group tested at two different times. For example, a medical trial may measure a factor (like blood pressure), apply some treatment, and then measure the factor a second time. In that case, a "paired t-test" would be used where the first trial's values would be compared to the second.

- **Analysis of Variance (ANOVA)**. An ANOVA is like a t-test but is used to analyze three or more groups to detect a significant difference.

The result of either a t-test or ANOVA is a p-value ("probability value"). A p-value describes the probability that some result was caused by pure chance and not from applied treatment. P-values are expressed as percentages, and, typically, researchers expect a p-value under 0.05 (5%) to declare the result is significant. If the calculated p-value is above 0.05, then the researcher declares that no significant result was found.

Nonparametric Hypothesis Tests

Some analysis techniques are only helpful with nonparametric data. While dozens of statistical processes will work with nonparametric data, two are seen most frequently.

- **Mann-Whitney U**. This test is used to determine any significant differences in two data groups that are not normally distributed, often categorical. As an example of using a Mann-Whitney U test, Gluck et al. used it in *How Short Is Too Short? Implications of Length and Framing on the Effectiveness of Privacy Notices* (Gluck, et al., 2016). This study postulated that people did not often read or understand long privacy notices. They conducted an experiment where they presented groups of people with short and long privacy notices and then assessed their understanding. A Mann-Whitney U test was used to compare the groups, and, not surprisingly, people who used the shorter form seemed to understand their privacy rights better than those who used a longer form. This experiment also varied the framing of the privacy policy (positive vs. negative). That part of the experiment used a Kruskal-Wallis H test to compare the results, but there was no effect due to framing differences. This study is an excellent example of using both Mann-Whitney U and Kruskal-Wallis H in a single analysis.
- **Kruskal-Wallis H**. This test is used to determine any significant differences in three or more data groups that are not normally distributed, often categorical. As an example of using a Kruskal-Wallis H test, Titlebaum and Lawrence published *Perceived Motivations for Corporate Suite Ownership in the "Big Four" Leagues* (Titlebaum & Lawrence, 2016). This study looked at 29 different reasons (like entertain clients, support the community, and business-to-business networking) corporations purchase luxury suites in any of the four major sports leagues (NFL,

> MLB, NBA, and NHL). They surveyed corporate leadership and then applied a Kruskal-Wallis H test to the survey results to see which reason was most important. They found four significant reasons: 1) Entertaining Employees, 2) Supporting the Community, 3) Perception of the Company in the Community, and 4) Customized Gifts for Suite Owners. They also compared the leagues, two at a time, to see if the reasons were different between the leagues and used a Mann-Whitney U test for this part of the analysis. They found that MLB had a significantly higher rank for “Entertaining Employees” than other sports. This study is an excellent example of using both Kruskal-Wallis H and Mann-Whitney U in a single analysis.

The result of either a Mann-Whitney U or Kruskal-Wallis H is a p-value (“probability value”). A p-value describes the probability that some result was caused by pure chance and not from the applied treatment. P-values are expressed as percentages, and, typically, researchers expect a p-value under 0.05 (5%) to declare the result is significant. If the calculated p-value is above 0.05, then the researcher declares that no significant result was found.

Data Mining

Data mining is a relatively new statistical technique that attempts to extract (or “mine”) information from massive databases. Raval (Raval, 2012) published an excellent overview of data mining techniques. As but one example, a retail store mines customer purchases, shopping habits, and other data to plan sales, send coupons, and make other suggestions that will, hopefully, lead to increased spending. While there are more than a dozen data mining techniques, these three are most common.

Clustering

It is beneficial to “cluster” customers in some way so advertising (and sales) can be more effective. For example, if a large grocery store chain can determine that most customers in one region share some common trait, it becomes easier to market to that region.

Figure 37 shows a scatter plot with three clusters. This plot was generated from dummy data, but it shows the type of clustering that a researcher may find in a dataset. The location of a specific cluster compared to the others would perhaps drive a marketing campaign focused on the customers' specific properties in that one cluster.

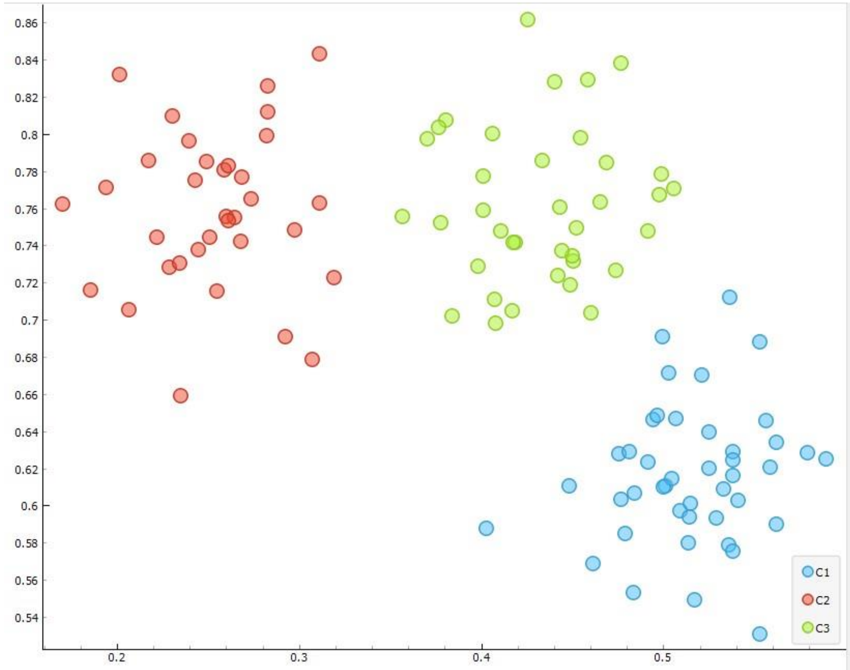


FIGURE 37: EXAMPLE OF CLUSTERING

As an example of using clustering, Vogel, Greiser, and Mattfeld published *Understanding Bike-Sharing Systems using Data Mining: Exploring Activity Patterns* (Vogel, Greiser, & Mattfeld, 2011). They analyzed more than 760,000 pickups and returns at bike stations in Vienna from 2008–2009. Their goal was to determine if the stations could be clustered in some way by usage patterns. They used cluster analysis to divide the city's stations into five groups; for example, the "returns morning-pickups evening" group showed an exceptionally high number of morning returns and evening pickups. They provided the city with information that was designed to help in future station placement.

Decision Tree

A decision tree organizes known data so predictions can be made on new data. Figure 38 is taken from a dataset that shows the fate of all the passengers on the Titanic. Starting at the top, few males survived. There were no surviving females with more than 4.5 siblings or spouses on board, so large families perished. Notice, though, nearly all first-class female passengers with few family members on board survived. A decision tree like this could be used to predict whether a given passenger survived.

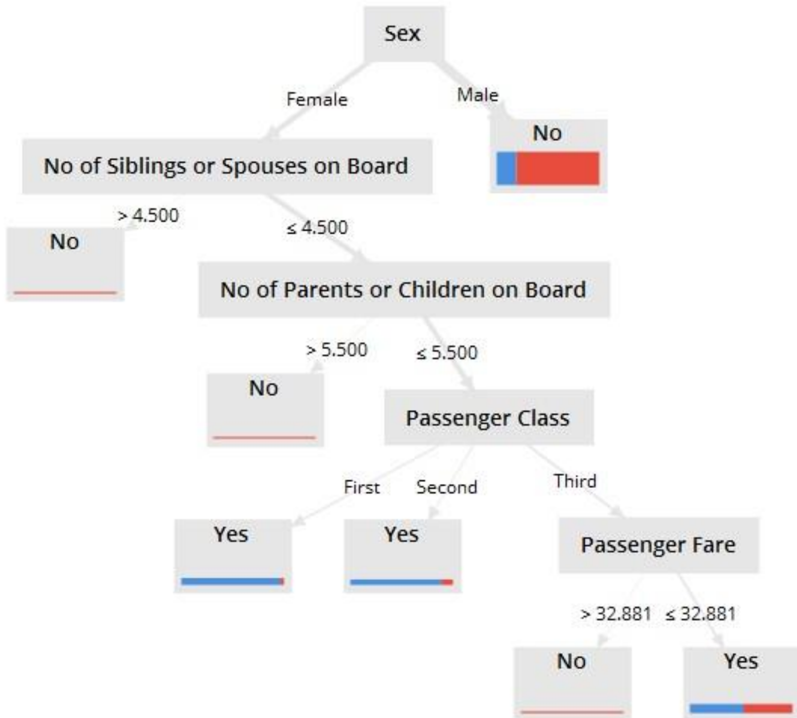


FIGURE 38: DECISION TREE

As an example of using a decision tree, the Grenada, Spain, public transportation system was studied (de Oña, de Oña, & López, 2016). Three thousand six hundred sixty-four passenger surveys conducted from 2008 until 2011 were mined. The surveys asked passengers some demographic questions (like age) and various bus service aspects. The researchers first created four clusters of passenger types using cluster analysis, “Young Students,” “Working Women,” “Sporadic Users,” and “Elderly Passengers.” Then they generated five different decision trees, one for each group and one overall, to define the difference between “poor” and “good” service ratings from the passengers. For example, in the Overall tree, service frequency was the first, most important, branch on the tree; but for Young Students, the first branch was punctuality. This study shows how two different data mining techniques, cluster analysis and decision tree building, can be used in a single study.

Market Basket

A market basket analysis evaluates the products that customers purchase simultaneously (they are in the same “market basket”) and then uses those data to drive store organization and sales decisions. For example, it is well known that customers who purchase beer also tend to purchase potato chips. What may be surprising is that customers who purchase beer also tend to purchase baby diapers[23]. A market basket analysis does not attempt to determine why these relationships exist, just that they do exist. A grocery store owner may choose to put beer and chips in the same sale or create a display near the front of the store with both beer and chips.

Figure 39 is a graphic representation of the rule set found in Figure 40, created from a market basket analysis of a bakery. The goal was to determine what sorts of products customers purchased with cherry tarts. *Rule 53* (the first line in Figure 40) indicates that apricot danish and cherry tarts tend to be purchased together, and that would be valuable information for a bakery owner. The explanation of the numbers displayed (like "Support" and "Confidence") is beyond the scope of this class. However, business owners may consider working with a research team who could complete this sort of analysis.

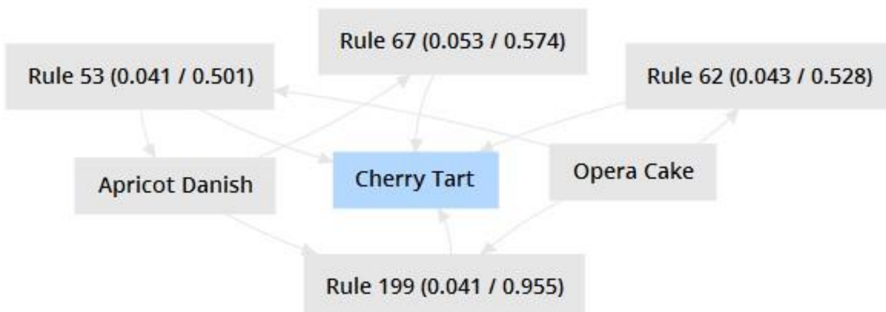


FIGURE 39: MARKET BASKET GRAPH

No.	Premises	Conclusion	Support	Confidence	LaPlace	Gain	p-s	Lift	Convicti...
53	Opera Cake	Cherry Tart, Apricot Danish	0.041	0.501	0.962	-0.123	0.037	9.431	1.897
62	Opera Cake	Cherry Tart	0.043	0.528	0.964	-0.121	0.036	5.671	1.923
67	Apricot Danish	Cherry Tart	0.053	0.574	0.964	-0.132	0.044	6.156	2.126
199	Apricot Danish, Opera Cake	Cherry Tart	0.041	0.955	0.998	-0.045	0.037	10.255	20.322

FIGURE 40: MARKET BASKET RULES

As an example of a market basket analysis, Musalem, Aburto, and Bosch completed a study of a mid-sized supermarket in Latin America (Musalem, Aburto, & Bosch, 2018). That market sells approximately 7,000 products, and the average basket size contains products from 3.6 categories. They reviewed the register receipts from July 2000 and recorded the products sold, units sold, date, and time. They

[23] It is fun to speculate why these two items are related. Maybe parents run out of diapers in the middle of the night and while they are at the store they also by some beer. Maybe they make a beer run to the store and then remember that they may also need diapers. At any rate, if a store puts diapers on sale, they may also want to consider a beer sale!

classified the products into four categories: non-perishable (like cereals, flour, and noodles), immediate consumption (like meat, milk, and cheese), hygiene (like shampoo, conditioner, and diapers), and hedonic (like ice cream, beer, and candy). They found that items in the hygiene basket were associated with sizable transactions, but not many of them. The non-perishable basket had many different categories and high transaction size, so these buyers are essential to the market. This study is an excellent example of how market basket analysis can be used to improve sales.

Key Takeaways

Figure 41 illustrates the relationship between the various types of data and the rating scales commonly used to work with those data types. Positivist researchers tend to gather quantitative data and use parametric statistical tests on interval and ratio data. Conversely, interpretivist researchers tend to gather qualitative data and use nonparametric statistical tests on nominal and ordinal data. Nominal data are typically gathered with binary and semantic difference rating scales, while ordinal data are typically gathered with Likert and Guttman rating scales.

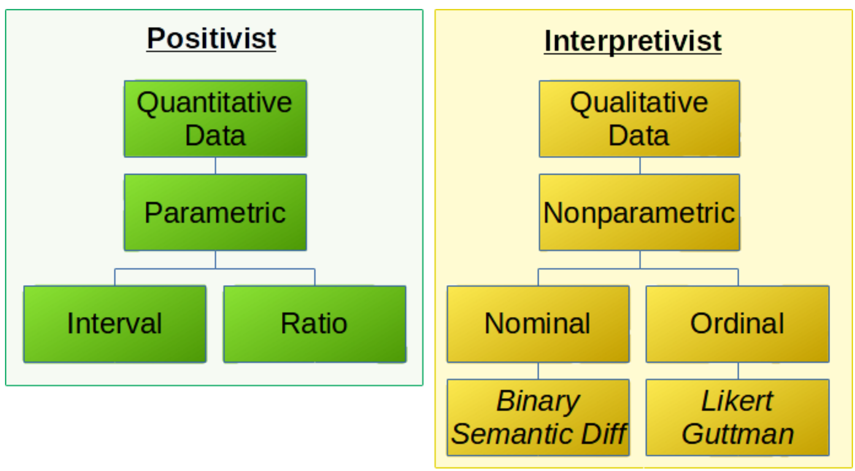


FIGURE 41: RESEARCH PHILOSOPHY AND DATA TYPES

CHAPTER 6: DATA

- Data have three commonly reported properties: distribution, excess kurtosis, and skew.
- Databases are collections of data, and many public databases can be helpful for research projects.
- Many statistical tests are used for central measures, spread, frequency, correlation, parametric hypothesis testing, and nonparametric hypothesis testing.
- Data mining is the process of extracting intelligence from massive data sets, and the three most common forms of data mining are clustering, decision trees, and market basket.

7: Sampling

Imagine that a research team undertakes a project to determine the number of t-shirts to order for an upcoming concert in a large stadium. The concert promoter needs to know how many small/medium/large t-shirts should be ordered. To figure that out, the researchers decide to conduct a survey of the fans at a similar concert a few weeks earlier and ask them what size t-shirt they prefer. So, how many fans should they survey? They could ask every fan in the stadium if they hired enough poll takers, but that would be expensive and wasteful. Perhaps it is enough to ask 10% of the crowd – maybe just 1%. Once the researchers figure out how many people to survey, they need to figure out how to select them. Do they select everyone in a specific section? Do they ask the first 100 people who come through a specific gate? These are questions that concern sampling, and whatever decisions the researchers make concerning the stadium crowd may profoundly change the collected data, not necessarily for the best. This chapter concerns sampling and offers information about determining sample size and what techniques can reduce bias in the sample.

OBJECTIVES
• Discuss sampling techniques used in quantitative research projects.
• Discuss sampling techniques used in qualitative research projects.
• Define the various statistical terms used with sampling.
• Determine the sample size required to improve the validity of a research project.
• Discuss the "WEIRD" problem with sampling.

Introduction

Business and marketing research generally concerns inferring behavior patterns within specific populations where a population is a cluster of people, events, things, or other phenomena of interest. Unfortunately, populations may be relatively large and vaguely defined, such as "the American people" or "grocery stores in the United States." Even more specific populations can be too large to study directly, like all shoppers over 18 within a specific region. Usually, it is not feasible to study an entire population because of cost constraints, so a representative sample is selected from the population for observation and analysis.

A sample is a small subset of the population from which data are gathered to make statistical inferences about the entire population. Sampling strategies are designed to allow researchers to make claims about populations much larger than the sample with a fair amount of confidence. Researchers need to choose a genuinely representative sample of the population, so the inferences derived can be generalized to the entire population. Improper and biased sampling is the primary reason for the often divergent and erroneous inferences reported in opinion polls conducted by different groups before every United States Presidential election.

The sampling process involves several stages.

1. **Define the target population.** A population can be defined as all people or items (unit of analysis) with interest characteristics. The unit of analysis may be a person, group, organization, country, object, or any other entity about which scientific inferences can be drawn. Sometimes the population is evident. For example, suppose a manufacturer wants to determine whether finished goods manufactured on a particular production line meet certain quality specifications. In that case, the population consists of the entire set of finished goods manufactured on that line. At other times, the target population may be a little harder to understand. Suppose the research project is to identify the primary motivators of shopping behavior among high school students. In that case, the target population may be high school students, store managers selling products to those students, or even the product manufacturers who design the packaging.
2. **Determine the sampling frame.** The sampling frame is an accessible section of the target population (often a list with contact information) from which a sample can be drawn. If the target population is “professional employees,” then the sampling frame would likely be the professional employees at a local company. It would not be possible to access all professional employees worldwide. If the target population is “business organizations,” then the Fortune 500 list may be an acceptable sampling frame.

   Note that sampling frames may not represent the population at large, so inferences derived by that sample may not be generalizable to the population. For instance, if the target population is the employees of small businesses and the sampling frame is employees at automotive tire companies in the American midwest, then findings from that group may not be generalizable to the American workforce at large, let alone the global workplace. Similarly, the Fortune 500 list includes the 500 largest American enterprises, which is not representative of American firms in general, most of which are medium and small-sized firms. Also, note that the targeted research population may not be the same as the population from which a sample is drawn. For example, suppose a researcher wants to determine the success rate of a new “quit smoking” program. In that case, the target population is the universe of smokers who had access to this program, an unknown population. Hence, the researcher may sample patients arriving at a local medical facility for smoking cessation treatment, some of whom may not have had exposure to this particular “quit smoking” program, in which case, the sampling frame does not correspond to the population of interest.

3. **Select a sample.** Once the sampling frame is defined, the actual sample must be drawn using a well-defined technique. Sampling techniques can be grouped into probability sampling (often called "random sampling") and non-probability sampling. Probability sampling is essential if the results are generalized, but there may be unique circumstances where non-probability sampling can also be justified.

As an example of this process, consider Figure 42. Imagine that a researcher intends to conduct a project where the intent is to draw some conclusions about professional workers in a large city. That population would be much too large to survey directly so that a sampling frame would be created of, perhaps, the workers in one part of the city. From that frame, specific workers would be selected for observation.

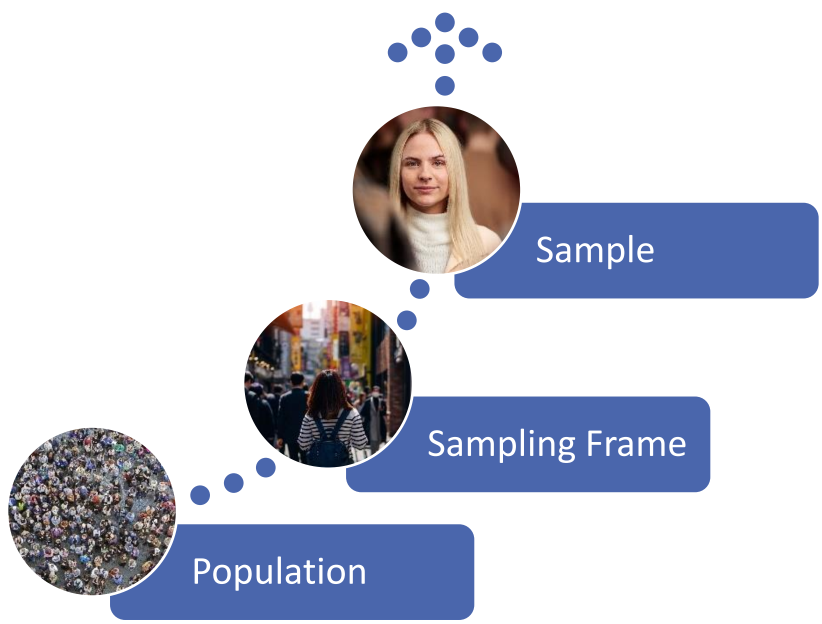


FIGURE 42: SAMPLE SELECTION PROCESS

Both qualitative and quantitative researchers use sampling techniques, but because qualitative and quantitative research goals differ, the sampling procedures are also different.

Sampling in Quantitative Research

Quantitative researchers are often interested in generalizing populations bigger than their study samples and rely on probability sampling. Probability sampling is a technique in which every unit in the population has a chance (nonzero probability) of being selected to be sampled, and this chance can be accurately determined. Every element in the sample frame must have a known chance to be sampled because that makes the sample representative of the entire population. If, for example, a researcher is investigating something about the differences between men and women shoppers, then it would be important that the sample contained both men and women in about the same proportions as in the entire population. Obtaining a representative sample is important because a critical goal of studies that rely on probability samples is generalizability. Generalizability is perhaps the key feature that distinguishes probability samples from non-probability samples. Generalizability refers to the idea that the results in the sample will be indicative of some trait in the entire population.

A parametric sample aims to form a normal distribution with negligible skew and excess kurtosis[24]. Statistics generated from the sample, such as the mean or standard deviation, are unbiased estimates of population parameters as long as the samples are appropriately selected. Therefore, statistical analysis of the sample can be reasonably applied to the entire population.

All probability sampling has two attributes in common: 1) every unit in the population has a known non-zero probability of being sampled, and 2) the sampling procedure involves random selection. The different types of probability sampling techniques include the following.

Simple Random Sampling

In this technique, all possible subsets of a population (more accurately, of the sampling frame) have an equal probability of being selected, making the sample an unbiased estimate of the population.

Simple random sampling involves randomly selecting respondents from a sampling frame, usually employing a table of random numbers[25] or a computerized random number generator. For instance, to randomly select 200 firms to survey from a list of 1000, the list could be entered into a spreadsheet program like Excel then the "random" function could be used to assign a random number to each of the

---

[24] See Chapter 6 for a definition of skew and excess kurtosis.
[25] A table of random numbers can be generated at https://stattrek.com/statistics/random-number-generator.aspx.

1000 firms[26]. Next, the list is sorted by those random numbers, and the first 200 firms on that sorted list would be selected. Simple random is the simplest of all probability sampling techniques, and that simplicity is also its greatest strength. Because the sampling frame is not subdivided or partitioned, the sample is unbiased, and the inferences are the most generalizable compared to other probability sampling techniques.

Figure 43 illustrates simple random sampling. A few individuals in the entire population are selected (circled) at random for the study.

FIGURE 43: SIMPLE RANDOM SAMPLING

Systematic Sampling

In this technique, the sampling frame is ordered according to some criteria, and then elements are selected at regular intervals through that ordered list. Systematic sampling involves a random start and then proceeds with selecting every kth element from that point onwards. The value of k is easily calculated by k = N/n, where k (the sampling ratio) is calculated by dividing the size of the sampling frame, N, by the desired number of samples, n. It is essential that the starting point is not automatically the first item in the list but is randomly chosen from within the first k elements. For example, imagine selecting 200 firms from a list of 1000; k is equal to 5 (k = 1000/200). The list of 1000 firms would be sorted in increasing (or decreasing) order by some size-related criterion, like employee count or annual revenues. Then, one of the first five (the value of k) firms on the list would be randomly selected, and every fifth firm after that. This process ensures that there is no over-representation of large or small

[26] That the "random" function in Excel (or any other computer program) is not truly random. However, it is usually random enough for most research projects.

firms in the sample, but, instead, firms of all sizes are generally uniformly represented. In other words, the sample is representative of the population, at least based on the sorting criterion.

There is one specific instance in which systematic sampling should be avoided. If the sampling frame has a pattern, then a particular bias — referred to as the *periodicity problem* — can be introduced with a systemic sampling strategy. Imagine, for example, that a researcher wanted to observe how people use the outdoor spaces on a college campus. Perhaps the researcher selects "random" dates in September to make observations. However, all observations may be on the same day of the week due to the weekday cycle. When working with patterned data, it is best not to use systematic sampling.

Figure 44 illustrates systematic sampling. In this case, k = 5, so every fifth person is chosen. At random, person three was chosen to start the pattern, and then every fifth person after that is chosen.

FIGURE 44: SYSTEMATIC SAMPLING

Stratified Sampling

In stratified sampling, the sampling frame is divided into homogeneous, non-overlapping subgroups, called "strata," and a simple random sample is drawn from each subgroup. If selecting 200 firms from a list of 1000, the firms could be stratified by size, perhaps "large" (more than 500 employees), "medium" (50−500 employees), and "small" (fewer than 50 employees). Then, 67 firms would be randomly selected from each subgroup to create a sample of 200 firms. However, since there are many more small firms than large firms, sampling an equal number from each group will favor large firms with fewer representatives in the sampling frame. This technique is non-proportional stratified sampling. The sample size within each subgroup does not reflect the proportions in the population, and the smaller subgroup (large-sized firms) is oversampled.

An alternative technique is to select subgroup samples in proportion to their numbers in the population. For instance, if there are 100 “large” firms, 300 “medium” firms, and 600 “small” firms, then 20 should be selected from the “large” group, 60 from the “medium” group, and 120 from the “small” group in order to maintain the correct proportion of firms. This technique is called proportional stratified sampling.

Figure 45 illustrates stratified sampling. The population is divided into five subgroups, and three elements were chosen from each subgroup. This illustration is an example of non-proportional stratified sampling since the same number are selected from each subgroup, though, in this case, each subgroup is the same size.

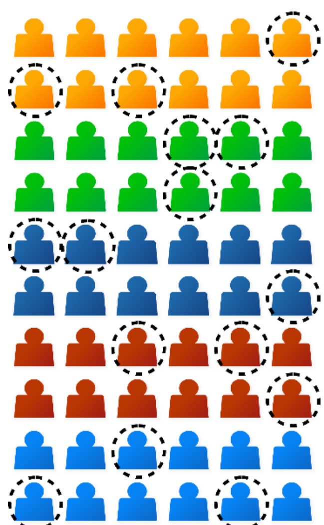


FIGURE 45: STRATIFIED SAMPLING

Cluster Sampling

If the population is dispersed over a broad geographic region, it may not be feasible to sample the entire population randomly. In such a case, it may be reasonable to divide the population into “clusters” (typically along geographic boundaries), randomly select only a few clusters, and then measure *all* units within that cluster. For instance, to sample small city governments in California, rather than travel all over the state to interview key city officials (as would be necessary with a simple random sample), the cities could be clustered by county. Then all small-city officials in a randomly selected group of counties would be interviewed. However, depending on between-cluster differences, the variability of sample results from a cluster sample will generally be higher than for a simple random sample so that the results would be less generalizable to the population.

Figure 46 illustrates cluster sampling. In this case, the population was divided into five clusters, and then clusters three and four were randomly selected for the research project.

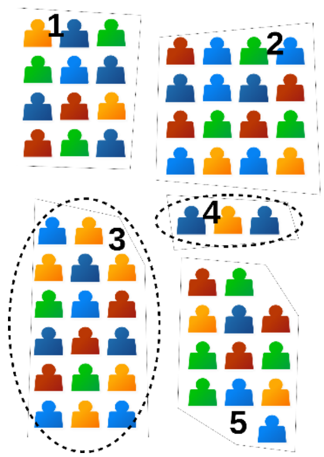


FIGURE 46: CLUSTER SAMPLING

Matched Pair Sampling

Sometimes, researchers may want to compare two subgroups within one population based on a specific criterion; for example, why are some firms consistently more profitable than others? The sampling frame of firms could be categorized into "highly profitable" and "moderately profitable" firms based on gross margins, earnings per share, or other profitability measures. Then, a simple random sample of firms in one subgroup is selected. Each of those firms would be matched with one in the other subgroup based on its size, industry segment, or other criteria. Now, each firm can be studied in greater detail to look for reasons to explain their differences. Such matched pair sampling technique is often an ideal way of understanding bipolar differences between subgroups within a given population.

Figure 47 illustrates matched pair sampling. In this case, one blue, green, and red element was selected at random and then matched with a similar element for comparison.

FIGURE 47: MATCHED PAIR SAMPLING

Multistage Sampling

The sampling techniques described previously are all examples of single-stage sampling; however, they can be combined to conduct multistage sampling. For instance, a list of firms can be stratified based on size, and then a systematic sampling can be conducted within each stratum. This process would be a two-stage combination of stratified and systematic sampling. As a second example, the school districts in the state of Florida could be grouped into clusters, then a simple random sample of schools taken within each cluster, then a simple random sample of grade levels within each school, and, finally, a simple random sample of students taken from each grade level. This complex sample would constitute a four-stage sampling process consisting of cluster and simple random sampling.

Sampling in Qualitative Research

Qualitative researchers typically make sampling choices that enable them to deepen their understanding of the phenomenon being studied. Non-probability sampling is a sampling technique in which some population units have zero chance of selection or where the probability of selection cannot be accurately determined. Typically, units are selected based on specific non-random criteria, such as quota or convenience. Because selection is not random, the sampling error cannot be estimated in non-probability sampling, making it subject to sampling bias. Also, the samples will not exhibit a normal distribution, so statistics like mean and standard deviation are of no value. As a result, information from a non-probability sample cannot be generalized back to the entire population.

Non-probability samples are ideal during the design phase of a research project. For example, a survey can be administered to only a few people who seem to resemble the population being investigated to help solve potential problems with the survey. A non-probability sample is also suitable for an exploratory research project where the goal is to determine if there is a need to conduct a more extensive project. Researchers also use non-probability sampling as the primary means of data collection in extensive qualitative research projects. The researcher's goal in such projects is to understand a specific unit rather than a more general understanding of the population. Thus, researchers interested in contributing to the theoretical understanding of some phenomena might use non-probability samples.

Types of non-probability sampling techniques include the following.

Convenience Sampling

This type of sampling is sometimes called “accidental” or “opportunity.” This sample is drawn from that part of the population close to hand, readily available, or convenient. Another form of convenience sampling is “volunteer” sampling, where the only samples are from subjects who volunteer for the project. As an example of convenience sampling, if a researcher stood outside a shopping center and surveyed shoppers as they walk in, it would be a convenience sample. This sample is non-probability because shoppers at other shopping centers are systematically excluded from the survey. The survey results may reflect the unique characteristics of this particular shopping center, such as the nature of its stores, the demographic profile of its patrons, or its location. However, they would not be representative of the opinions of the shopper population at large. Hence, the scientific generalizability of such observations will be very limited. Other examples of convenience sampling are sampling students registered in a particular class or sampling patients arriving at a specific medical clinic. This type of sampling is most useful for pilot testing, where the goal is instrument testing or measurement validation rather than obtaining generalizable inferences.

Figure 48 illustrates convenience sampling where the researcher in the top left corner only samples the nearby elements and ignores the rest.

FIGURE 48: CONVENIENCE SAMPLING

Quota Sampling

The population is segmented into mutually exclusive subgroups (just as in stratified sampling). Then a non-random set of observations is chosen from each subgroup to meet a predefined quota. In proportional quota sampling, the proportion of respondents in each subgroup should match that of the population. For instance, if the American population consists of 70% Caucasians, 15% Hispanic-Americans, and 13% African-Americans, and researchers wished to understand their voting preferences in a sample of 100 people, they could ask people their voting preferences outside a shopping center. However, they would have to stop asking Hispanic people when they have 15 responses from that subgroup even as they continue sampling other ethnic groups. The goal is that the ethnic composition of the sample matches that of the general American population. Nonproportional quota sampling is less restrictive in that it does not have to achieve a proportional representation but does, perhaps, meet a minimum size in each subgroup. In this case, researchers may decide for whatever reason to have 50 respondents from each of the three ethnic subgroups (Caucasians, Hispanic-Americans, and African-Americans) and then stop when the quota for each subgroup is reached. Neither type of quota sampling will represent the American population since a sample from a shopping center in New York would be different from Kansas. The nonproportional technique is even less representative of the population, but it allows capturing the opinions of marginalized groups.

Figure 49 illustrates quota sampling. A random sample is drawn from each subgroup, but the bigger subgroups have more samples than the smaller subgroups.

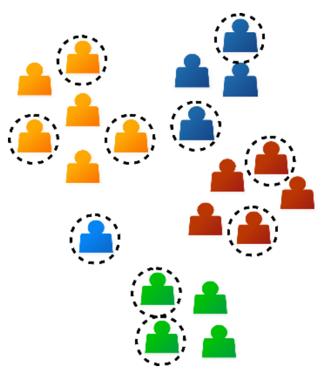


FIGURE 49: QUOTA SAMPLING

Snowball Sampling

In snowball sampling, a researcher identifies a few respondents who match the criteria for inclusion in the study and then asks those respondents to recommend others they know who would also meet the selection criteria. Snowball sampling is sometimes referred to as “chain referral sampling” since each respondent is also asked to provide the names of other potential respondents. After a few rounds, the researcher would have many respondents, like a snowball getting more massive as it rolls down a mountain. For example, suppose a researcher wishes to survey computer network administrators but only knows one or two. In that case, those administrators could recommend their peers, and that group could recommend others. Although this method hardly leads to representative samples, it may sometimes be the only way to discover hard-to-reach subjects or when no sampling frame is available.

Figure 50 illustrates snowball sampling. In this case, the researcher asked only one person, who recommended one other, who recommended two more, and so forth. In the end, the researcher sampled eight people but had only one to start the process.

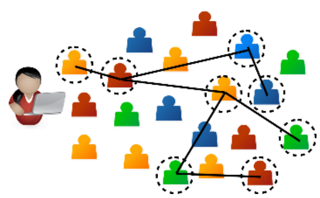


FIGURE 50: SNOWBALL SAMPLING

Purposive Sampling

To draw a purposive sample, a researcher begins with specific attributes to examine and then seeks research participants who exhibit those attributes. For example, a researcher studying students' satisfaction with their living quarters on campus would want to be sure to include students who stay in each of the different types or locations of on-campus housing. If only students from one of the ten dorms on campus are included, then essential details about the experiences of students who live in the other nine dorms would be missed.

Expert Sampling

This sampling technique chooses respondents in a non-random manner based on their expertise on the studied phenomenon. For instance, to understand the impacts of a new governmental policy such as the Sarbanes-Oxley Act, a group of corporate accountants familiar with this act would be sampled. The advantage of this approach is that since experts tend to be more familiar with the subject matter than nonexperts, opinions from a sample of experts are more credible than a sample that includes both experts and non-experts. However, the findings are still not generalizable to the overall population at large.

Summary of Sampling Techniques

Type	Strengths	Weaknesses
Random	• Widely accepted • Reasonable chance of getting a representative sample • High reliability and validity	• May not capture minority groups in the population • It may be impractical in some studies
Systematic	• If the list randomized, bias unlikely	• If the list is not randomized, bias likely
Stratified	• Avoids bias sometimes caused by random	• Takes time and resources • All population characteristics must be accounted for
Cluster	• Decreased chance for bias • Can sample geographically diverse regions	• Time-consuming • Difficult to set up
Matched Pair	• Improves validity, especially for comparison studies	• Difficult to find matching pairs
Multistage	• A thorough sampling of a population • Can drill down to the target in a population	• Difficult to set up • A researcher must be skilled in all sampling methods used
Convenience	• Easy and inexpensive	• Sample subject to bias
Quota	• Assures samples from all target population groups	• Prone to bias
Snowball	• A thorough sampling of a population	• Time-consuming and challenging to complete
Purposive	• Samples the best possible subjects in the population	• Prone to bias
Expert	• Provides potentially better responses	• It can be difficult and expensive to find experts • It takes time to elicit responses

TABLE 6: SUMMARY OF SAMPLING TECHNIQUES

Statistics of Sampling

The preceding discussion introduced terms such as population, parameter, sample statistic, and sampling bias. This section defines these terms and illustrates how they are used and reported in a research project.

For this discussion, consider the data seen in Figure 51. This data set represents the responses to a restaurant web survey typical in many fast food places. In this data, the attributes (columns) are the customer identification, the store number, the date and time a purchase was recorded, the purchase made (in coded form), and the answers to four questions, Q1-Q4. One column, illustrated in green, represents all responses to a single attribute, Q1 in this case. One observation (row) is highlighted in blue and represents all of the responses by a single customer.

FIGURE 51: STASTICAL TERMS

A specific observation of a phenomenon, such as a person's answer to item Q2, is called a response. In other words, a response is one point of measurement provided by a sampled unit (a customer). Researchers would expect each respondent to provide different responses to the items in a survey. All responses for a single item, or a single column of data, can be analyzed in many different ways. If the item has nominal data, data in categories, then the responses can be graphed into a frequency distribution based on their frequency of occurrences. For this survey, the responses to Q2 can only be A, B, C, or D since those were the only selections possible, making Q2 nominal data. A frequency distribution for Q2 could look something like Figure 52.

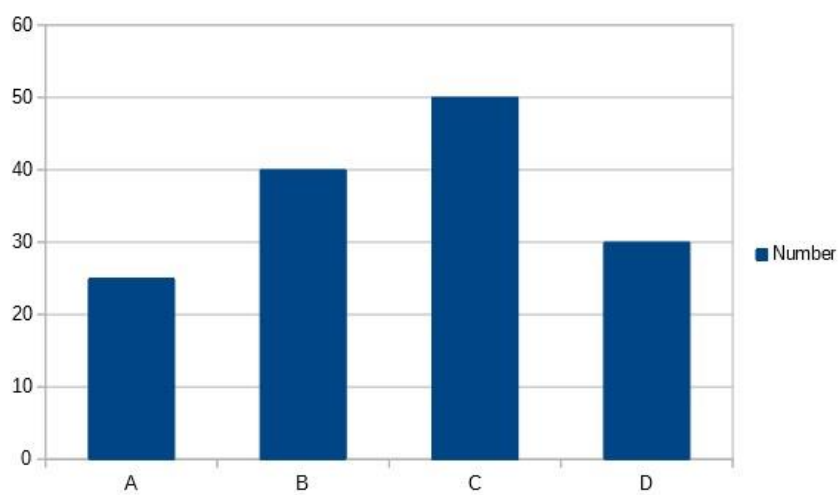


FIGURE 52: FREQUENCY DISTRIBUTION

Thus, the researcher could report that 40 people chose response B. In contrast, 50 selected response C. If a sample has many responses, the frequency distribution resembles a normal distribution. That can then be used to estimate the overall characteristics of the entire sample. Continuous data, such as measurements, can have various characteristics calculated, like the sample mean or standard deviation. These sample estimates are called sample statistics (a "statistic" is a value that is estimated from observed data).

Populations also have means and standard deviations that could be obtained if the entire population could be sampled. However, since the entire population can usually not be sampled, population characteristics are unknown and are called population parameters (not "statistic" because they are not statistically estimated from the data). Sample statistics may differ from population parameters if the sample is not perfectly representative of the population. For example, the sample mean might be 45 while the population parameter is 48, and the difference between the two is called "sampling error." Suppose sample size could be gradually increased until it approaches that of the population. In that case, the sampling error will decrease, and a sample statistic will approximate the corresponding population parameter. In reality, though, it is usually impossible to have a sample size equal to the population.

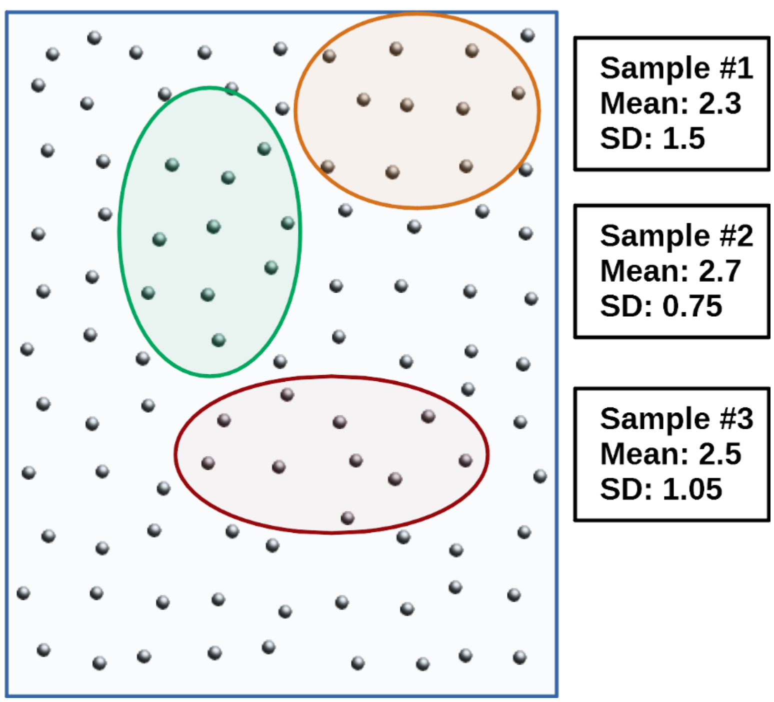


FIGURE 53: SAMPLE STATISTICS

Suppose a sample is truly representative of the population. In that case, the estimated sample statistics would be identical to corresponding population parameters. In most cases, though, a sampling distribution is used to estimate how closely the statistics align with the parameters.

Figure 53 illustrates the data from a research project. Imagine that three different random samples are drawn from the population, and the mean and standard deviation are calculated for each sample. If each random sample was perfectly representative of the population, then the means and standard deviations from the three samples will be identical and equal to the population parameter. However, this is extremely unlikely since each random sample will constitute a different population subset, so their calculated statistics will be slightly different from each other.

If the mean of the three sample means is calculated, then it would be possible to calculate the variability (or spread) of those means, and that calculation is called the standard error of the mean. Imagine the standard error of the mean for Figure 53 is 0.115[27].

As the number of samples drawn from the population increases, the variance between the means of those samples tends to decrease, and they approach the mean of the entire population. Also, as the number of samples increases, the standard error of the mean decreases and approaches zero. The mean value of a sample is presumed to be an estimate of the population parameter. The standard error of the mean makes it possible to estimate confidence intervals for how well the statistical mean predicts the population parameter. Since the standard error is similar to the standard deviation for a group of samples, it can be said that:

- (Sample statistic + one standard error) represents a 68% confidence interval for the population parameter.
- (Sample statistic + two standard errors) represents a 95% confidence interval for the population parameter.
- (Sample statistic + three standard errors) represents a 99% confidence interval for the population parameter.

In Figure 53, the mean of the three samples is 2.5, and the standard error is 0.115. Therefore, any sample mean in the range of 2.385− 2.615 has a 68% chance of predicting the population mean, any sample mean in the range of 2.27−2.73 has a 95% chance of predicting the population mean, and any sample mean in the range of 2.155−2.845 has a 99% chance of predicting the population mean. Since sample one falls in the 95% confidence band, a researcher could predict that 2.3 (the mean of that sample) predicts the mean of the entire population with a 95% confidence level.

[27] Students interested in how this number was derived can find help online. One source is an online calculator: https://www.calculator.net/standard-deviation-calculator.h tml

Sample Size

When working with sampling, one of the first questions researchers face is the sample size. In other words, how many observations are needed to have a helpful sample? The larger the sample, the more accurately it will represent the population, and the more robust the statistical analysis is possible. Unfortunately, the answer to the question about sample size is a bit nebulous. Many researchers mention that one rule of thumb is a minimum of 30, but that is a gross oversimplification and does not work well except in classroom-size projects for university students. When determining the correct sample size, a number that is too small may exclude parts of the population, while a number that is too large may become unwieldy for the researcher.

Determining the sample size is a mix of both mathematics and experience. Some formulas will help researchers determine sample size, but those must be tempered with experience and understanding the population being studied. For example, suppose a formula indicated that a sample size of 250 is appropriate for a population of 500 people. However, suppose the researcher knows that the population is very diverse with many racial or ethnic groups. In that case, the sample may need to be increased to ensure that all groups are fairly represented.

Numerous sample size tables have been published to make researchers' jobs more straightforward, and it is easy to refer to these tables to determine the correct sample size. Table 7 shows an elementary sample size table, adapted from Determining Sample Size by Glenn Israel (Israel, 1992).

Population Size	±5%	±7%	±10%
100	81	67	51
125	96	78	56
150	110	86	61
175	122	94	64
200	134	101	67
225	144	107	70
250	154	112	72
275	163	117	74
300	172	121	76
325	180	125	77
350	187	129	78
375	194	132	80
400	201	135	81
425	207	138	82
450	212	140	82

TABLE 7: SAMPLE SIZE FOR 95% CONFIDENCE LEVEL

To read Table 7, imagine that a researcher is doing a project that involves surveying people.

- The population size is in the first column. Thus, the first data row is for a population of 100 people.
- The level of precision (p-value) is on the top row. That is the range for the population parameter compared to the statistic. Thus, if the sample mean of the research project is 60, and the p-value is ±5%, then the population mean will be between 55 and 65.
- The sample size is in columns 2-4. This many people would need to be surveyed to create a confidence level of 95%. That means that if 100 samples were drawn from the population, then 81 surveys would result in the same result 95 times.

A Word of Caution

When reading a research report, it is easy to overlook the need to ask where research participants came from and how many were observed, in other words, how the sample was determined. It is easy to focus only on the interesting "stuff" in the findings, but understanding the procedures used for selecting study participants is critical.

Students who have ever taken an introductory psychology or sociology class at a large university are frequently drafted into research projects since that is an easy group for graduate students to access. However, that access comes at a cost: sample representativeness. One study of top academic journals in psychology found that over two-thirds (68%) of participants in studies published by those journals were based on samples drawn in the United States (Arnett, 2008). Further, the study found that two-thirds of the work derived from United States samples published in the Journal of Personality and Social Psychology was based on samples made up of American undergraduates taking psychology courses.

These findings raise an interesting question: What do we learn from social scientific studies, and about whom do we learn it? That concern was raised by Joseph Henrich and colleagues (Henrich, Heine, & Norenzayan, 2010). They point out that behavioral scientists make sweeping claims about human nature based on samples drawn only from WEIRD (Western, Educated, Industrialized, Rich, and Democratic) societies. Too often, the claims are based on even narrower samples, as is the case with studies relying on samples drawn from college classrooms.

As it turns out, many robust findings of the nature of human behavior when it comes to fairness, cooperation, visual perception, trust, and other behaviors are based on studies that excluded participants from outside the United States and sometimes excluded anyone outside the college classroom. This exclusion raises questions about what we know about human behavior instead of United States undergraduate behavior. Of course, not all research findings are based on samples of WEIRD folks like college students. Thus, it is essential to pay attention to the samples on which studies are based and the claims made about whom those studies apply.

A related but slightly different concern is sampling bias. Bias in sampling occurs when the elements selected for inclusion in a study do not represent the larger population from which they were drawn. For example, an online poll conducted by a newspaper about a local issue will not represent the public. Those without access to the Internet, who do not read that paper's website, and do not have the time or interest will not answer the question.

Another thing to keep in mind is that just because a sample may be representative in all respects that a researcher thinks are relevant, relevant aspects may not be considered. For example, peoples' phones would seem to have little to do with their voting preferences. However, if pollsters making predictions about the results of the 2008 presidential election not included both cell phones and landlines, their predictions would have underestimated Barack Obama's lead since he was more favored by cell-only users than McCain (Keeter, Dimock, & Christian, 2008).

When evaluating a research report, remember that sample quality is determined only by the sample obtained, not by the sampling method itself. A researcher may administer a survey to a representative sample by correctly employing a random selection technique. However, if only a handful of the people sampled respond to the survey, the researcher will have to be very careful about the claims made. Another issue is that researchers sometimes discuss their findings as though they apply groups other than the population sampled. This tendency is usually innocent since it is human nature to inflate the findings. However, consumers of those findings must be attentive to this sort of (hopefully unintentional) bait and switch.

At their core, questions about sample quality should address who has been sampled, how they were sampled, and why they were sampled. Being able to answer those questions will improve the understanding of the research results.

Key Takeaways

This chapter introduced several issues related to sampling in a research project.

Chapter 7: Sampling
• Discuss sampling techniques used in quantitative research projects. • Discuss sampling techniques used in qualitative research projects. • Define the various statistical terms used with sampling. • Determine the sample size required to improve the validity of a research project. • Discuss the “WEIRD” problem with sampling.

8: Survey Research

How do retailers know what sorts of products their customers are likely to purchase? They use surveys to ask questions. Unfortunately, creating an unbiased survey that asks the right questions is a much more complex task than it may seem. This chapter introduces the art and science of survey design, including various surveys, writing questions, and analyzing the results.

OBJECTIVES

- Differentiate between cross-sectional and longitudinal surveys.
- Compare and contrast the four types of longitudinal surveys: trend, panel, cohort, and retrospective.
- Describe the characteristics of compelling questions.
- Define the various question response options.
- Describe tips for creating good questionnaires.
- Discuss how survey data are analyzed.
- Discuss bias in survey research.

Introduction

Survey research involves standardized questionnaires or interviews to collect data about people and their preferences, thoughts, and behaviors systematically. Although census surveys were conducted as early as Ancient Egypt, using a survey as a formal research method was pioneered in the 1930 and 40s by sociologist Paul Lazarsfeld to examine the effects of the radio on political opinion formation of the United States. This method has since become a prevalent method for quantitative research in business and social sciences. Because most students have completed many surveys, they often underestimate the skill and effort needed to create a valid survey. The process is time-consuming, tedious, and requires many revisions.

The survey method is best suited for studies that have individual people as the unit of analysis. Although other units of analysis, such as groups, organizations, or dyads (pairs of organizations, such as buyers and sellers), are also studied using surveys, such studies often use a specific person from each unit as a "key informant" or a "proxy" for that unit. Such surveys may be subject to respondent bias if the informant chosen does not have adequate knowledge or has a biased opinion about the phenomenon of interest. For instance, Chief Executive Officers may not adequately know employee's perceptions or teamwork in their own companies. They may therefore be the wrong informant for studies of team dynamics or employee self-esteem.

Strengths

- Measures unobserved data
- Remote data collection
- Results generalizable
- High reliability
- Can detect minor effects
- Economical in time and cost

Weaknesses

- Subject to many biases
- Cannot change instrument during use
- Questions can be tricky to write

FIGURE 54: SURVEY STRENGTHS AND WEAKNESSES

Survey research has several inherent strengths compared to other research methods.

Surveys are an excellent vehicle for measuring a wide variety of unobservable data, such as people's preferences (e. g., political orientation), traits (e. g., self-esteem), attitudes (e. g., toward immigrants), beliefs (e. g., about a new law), behaviors (e. g., smoking or drinking behavior), or factual information (e. g., income).

Survey research is also ideal for remotely collecting data about a population that is too large to observe directly. A large area, such as an entire country, can be covered using mail-in, electronic mail, internet, or telephone surveys using meticulous sampling to ensure that the population is adequately represented in a small sample. Due to their unobtrusive nature and the ability to respond at one's convenience, questionnaire surveys are preferred by some respondents.

Surveys are more easily generalized than other research techniques since data can be collected from extensive samples at a relatively low cost. Because surveys are standardized in that the same questions, phrased in precisely the same way, are posed to all participants, they tend to have higher reliability than other methods of gathering data. Interviews may be the only way of reaching certain population groups such as the homeless or illegal immigrants, for which there is no sampling frame available.

Extensive sample surveys may allow the detection of minor effects while analyzing multiple variables. Depending on the survey design, comprehensive sample surveys may also allow comparative analysis of population subgroups (i.e., within-group and between-group analysis). Survey research is more economical in terms of researcher time, effort, and cost than most other experimental research and case research methods.

At the same time, survey research also has some disadvantages. It is subject to many biases such as non-response bias, sampling bias, social desirability bias, and recall bias (all of these are discussed later in this chapter). While surveys are flexible in the sense that any number of questions on any number of topics may be asked, the researcher is also stuck with that instrument even if it is later shown to contain confusing items. Survey questions must be written such that a broad range of people will understand each of them. Because of this, survey results may suffer from validity concerns not found in more flexible methods.

Types of Surveys

Surveys are categorized by either time, how often a survey is conducted, or administration, how it is delivered to respondents. This section develops both types of concepts.

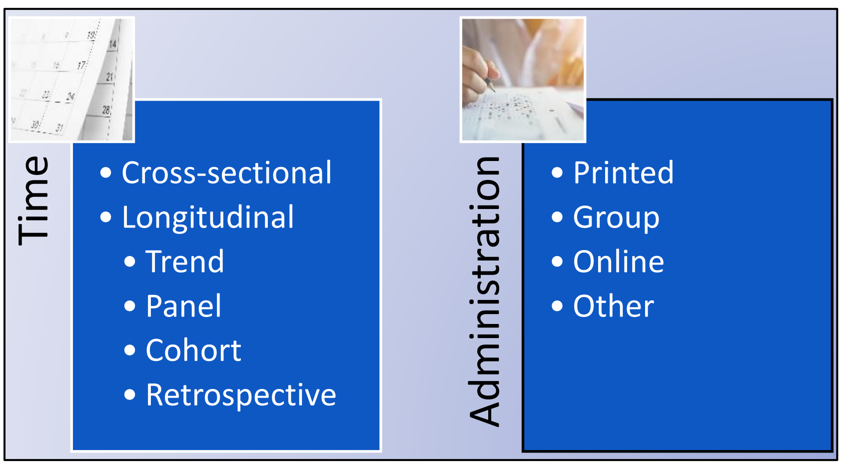


FIGURE 55: TYPES OF SURVEYS

Time

In terms of time, there are two main types of surveys: cross-sectional and longitudinal.

CROSS-SECTIONAL

Cross-sectional surveys are those that are administered at just one point in time. These surveys offer researchers a snapshot in time and explain how things are at a particular point in time. These surveys are called "cross-sectional" since they will take a snap-shot across multiple analytical units. For example, a survey may be administered to staff members in the human resources department of five different companies or customers of several different movie theaters on the same evening.

An example of a cross-sectional survey is a study of e-cigarette use among adolescents conducted by Dutra and Glantz (Dutra & Glantz, 2014). They used a cross-sectional survey of more than 40,000 students from more than two hundred middle and high schools across the United States. They determined that “...ever e-cigarette use was associated with higher odds of ever smoking cigarettes...”

Another example of a cross-sectional survey, Jørgensen et al. (Jørgensen, Villadsen, Burr, Punnett, & Holtermann, 2016), investigated if workplace health promotions depend on the work environment. They surveyed 10,605 Danish workers and determined that lower participation in health promotions is dependent on when they are offered (during or after work), the social support at work for the programs, and whether their work has high physical demands.

One problem with cross-sectional surveys is that the events, opinions, behaviors, and other phenomena that such surveys are designed to assess are generally not stagnant. Thus, generalizing from a cross-sectional survey about the way things are can be tricky. Perhaps something can be concluded about how things were when the survey was administered, but it is not easy to know whether things remained that way afterward. For example, imagine how Americans might have responded to a survey about terrorism on September 10, 2001, compared to September 12, 2001. The point is not that cross-sectional surveys are useless, but researchers must remember that these surveys are a snapshot in time.

LONGITUDINAL

Longitudinal surveys are those that include observations made over an extended period. There are four types of longitudinal surveys: trend, panel, cohort, and retrospective.

Trend survey. The first type of longitudinal survey is a trend survey. This type of study takes place over a long period, often years, and involves multiple surveys. For example, Dobrow, Ganzach, and Liu (Dobrow Riza, Ganzach, & Liu, 2015) studied job satisfaction regarding employee’s age and tenure. They surveyed 21,670 people in 34 “waves” of data collection spanning 40 years. They found that as tenure in an organization increased, people tend to be less satisfied with their jobs; however, as people age and move from job to job, their satisfaction increased.

Panel survey. Unlike a trend survey, a panel survey uses the same people each time it is administered. It is easy to imagine that administering a survey to the same 100 people every year for five years in a row can be expensive. Keeping track of where people live, when they move, and when they die takes resources that researchers often do not have. Panel surveys, however, can be powerful.

The Youth Development Study (YDS), administered by the University of Minnesota (Minnesota, 2018), is an excellent example of a panel study.

Since 1988, YDS researchers have administered an annual survey to the same 1,000 people. Study participants were in ninth grade when it began, and they are now in their thirties. Several hundred papers, articles, and books have been written using data from the YDS. One of the significant lessons learned from this panel study is that work has a largely positive impact on young people. Contrary to popular belief about the impact of work on adolescents' performance in school and transition to adulthood, work increases confidence, enhances academic success, and prepares students for success in their future careers. This panel study provided important information about the effect of work on young people.

As an example of a panel survey for business, Huhtala, Kaptein, and Feldt conducted a two-year study concerning how the ethical culture of organizations influences the well-being of managers (Huhtala, Kaptein, & Feldt, 2016). They found that managers in low or decreasing ethical cultures experienced changes in their well-being over the two years of the study.

Cohort survey. A researcher identifies a category of people in a cohort survey and then regularly surveys the people who fall into that category. The same people do not necessarily participate from year to year, but all participants must meet whatever categorical criteria fulfill the researcher's primary interest. Familiar cohorts that may be of interest to researchers include people of particular generations or those born around the same time, graduating classes, people who began work in a given industry simultaneously, or perhaps people who have some specific life experience in common. An example of this sort of research can be seen in Christine Percheski's work on cohort differences in women's employment (Percheski, 2008). Percheski compared women's employment rates across seven different generational cohorts, from Progressives born between 1906 and 1915 to Generation Xers born between 1966 and 1975. She found, among other patterns, that professional women's labor force participation had increased across all cohorts. She also found that professional women with young children from Generation X had higher labor force participation rates than similar women from previous generations. She concluded that mothers do not appear to be opting out of the workforce, as some journalists had speculated.

In another cohort study, Wright and Hinson surveyed public relations practitioners over ten years from 2005 until 2015 (Wright & Hinson, 2015). While some of the same practitioners would have participated over that entire span, it is reasonable to assume that at least a few respondents changed during that period. However, they were all public relations experts in a company. They found that Facebook and Twitter were the dominant means of public relations communication while LinkedIn and YouTube were not popular. The survey respondents, though, agreed that social media were changing the way public relations is practiced. It would be interesting to see how the perceptions of public relations experts have changed since 2016, given the heavy reliance on social media by politicians in the United States.

All three types of longitudinal surveys share the strength of permitting a researcher to make observations over time. Thus, if whatever behavior or other phenomena in which researchers are interested changes, they will capture those changes.

Retrospective survey. Retrospective surveys are similar to other longitudinal studies in that they concern changes over time, but like a cross-sectional study, they are administered only once. In a retrospective survey, participants are asked to report events from the past. Researchers can gather longitudinal-like data without actually incurring the time or expense of a longitudinal survey by having respondents report past behaviors, beliefs, or experiences. Of course, this benefit must be weighed against the possibility that people's recollections of their pasts may be faulty. Imagine, for example, that people are asked in a survey to respond to questions about where, how, and with whom they spent last Valentine's Day. Since Valentine's Day cannot be more than 12 months ago, the chances are favorable for an accurate response. However, if the question is to compare last Valentine's Day with the six previous Valentine's Days, the result would be much different.

Table 8 summarizes each of the four types of a longitudinal survey.

Type	Description
Trend	The researcher examines trends over time; the same people do not necessarily participate in the survey more than once.
Panel	The researcher surveys the same sample several times over some time.
Cohort	The researcher identifies some category of interest and then regularly surveys people who fall into that category.
Retrospective	Surveys are conducted only once, but the respondents are asked to report things from the past.

TABLE 8: COMPARE THE FOUR TYPES OF LONGITUDINAL SURVEY

Administration

One common way to administer surveys is in the form of a self-administered questionnaire. This type of survey means that research participants are given a set of questions in writing, to which they are asked to respond. Self-administered questionnaires can be delivered in printed format, in a group setting, or online.

Printed survey. Printed self-administered questionnaires may be delivered to participants in person; for example, people gathered in one place, like students in class, church members, or inmates. Researchers may also hand-deliver questionnaires by going door-to-door and either asking people to fill them out right away or making arrangements for the researcher to return to pick up completed surveys. Though the advent of online survey tools has made the door-to-door delivery of surveys nearly extinct, an occasional survey researcher may still use this method, especially around election time.

If a researcher cannot visit each sample member to deliver a survey, sending it through physical mail may be another consideration. While this mode of delivery may not be ideal (imagine how much less likely someone would be to return a survey where the researcher was not standing on the doorstep waiting), sometimes it is the only available or the most practical option.

Often survey researchers who deliver their surveys via physical mail provide some advance notice to respondents about the survey to get people thinking about and preparing to complete it. They may also follow up with a mailed reminder a few weeks after their survey has been distributed. These notices remind people to complete the survey and to thank those who returned the survey. Most survey researchers agree that this sort of follow-up is essential for improving mailed surveys' return rates (Babbie & Wagenaar, Unobtrusive research, 2010).

Group survey. The second type of survey administration is a group-administered questionnaire where respondents are brought together, and each respondent is asked to complete the questionnaire while in the room. Respondents typically sit at desks or in a study carrel and enter their responses independently without interacting with each other. This format is convenient for the researcher, and a high response rate is assured. Also, if respondents do not understand any specific question, they can ask for clarification. These surveys are most useful in an organization where it is relatively easy to assemble employees in a conference room or lunchroom, especially if corporate executives approve the survey.

Online survey. One final approach to delivering surveys is online. This delivery mechanism is becoming increasingly common, no doubt, because it is easy to use, relatively cheap, and quicker than knocking on doors or waiting for mailed surveys to be returned. Researchers may subscribe to a service offering free online survey delivery, like SurveyMonkey. One advantage of using a service like SurveyMonkey, aside from online delivery, is that results can be provided in readable formats by data analysis programs such as R and Excel. Online surveys save researchers the step of manually entering data into an analysis program, as is necessary for hard copy surveys.

There are several weaknesses in online surveys. If the survey website is not password-protected or designed to prevent multiple submissions, the responses can be easily compromised. Furthermore, sampling bias may be a significant issue since the survey cannot reach people who do not have computer or Internet access, such as the poor, seniors, and minority groups. Moreover, the respondent sample will be skewed toward a younger demographic who are online much of the time and have the time and ability to complete such surveys. Finally, computing the response rate may be problematic if the survey link is posted on Facebook, Twitter, or other social media sites instead of directly e-mailed to targeted respondents.

Online surveys often include an incentive to improve the response rate. Sometimes, the only incentive is knowing that the respondent is helping other people. It is possible, though, to provide some coupons to an online store like Amazon. Commonly, online surveys provide food, like a “free large drink,” from the restaurant administering the survey. Finally, it is possible to have respondents provide some sort of contact information, like an email address, and then have a drawing for a free tablet or some other prize. Using these sorts of rewards raises questions about the validity of the results. If people are only participating in a survey to have a chance at a prize, are they simply pattern-responding (choose all “A” answers, for example), or will they take the time to respond thoughtfully?

Other types. Researchers sometimes administer surveys by posing questions directly to respondents rather than having them read the questions independently. These types of surveys are interviews, which are discussed in Chapter 10: Interviews. It is enough at this point to mention that interview methodology differs significantly from survey research in that data are collected via personal interaction.

Designing Effective Questionnaires

Invented by Sir Francis Galton, a questionnaire is a research instrument consisting of items intended to capture responses from respondents in a standardized format. Items may be either structured or unstructured. Structured items ask respondents to select an answer from a set of choices. The responses are then aggregated into a composite scale or index for statistical analysis. On the other hand, unstructured questions ask respondents to respond in their own words using a free-flow type of entry. Questions should be designed such that respondents can read, understand, and respond to them in a meaningful way, so surveys would not be appropriate for specific demographic groups such as children or the illiterate.

Effective Questions

Writing survey questions is more of an art than a science, but three steps will help this process.

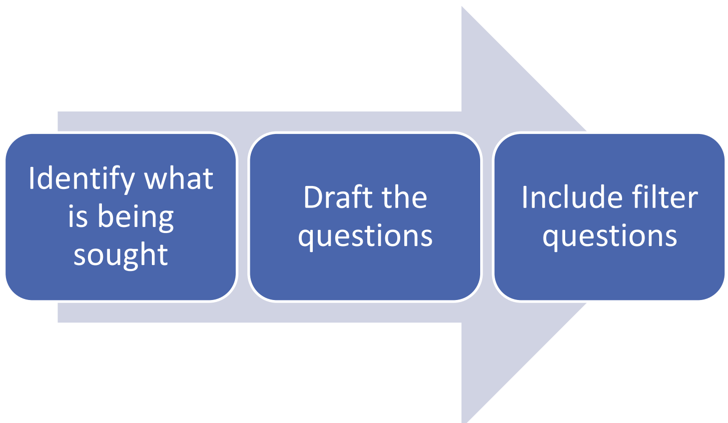
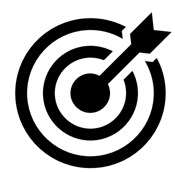


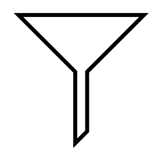
Figure 56: Writing Survey Questions



The first thing needed to write survey questions is identifying what, exactly, is being sought. Though it seems obvious, missing essential questions when designing a survey is far too frequent. Suppose researchers want to understand how students successfully transition from high school to college and what factors contribute to that success. To understand those factors, the researchers will need to include questions about all possible factors contributing to success. They would consult the literature on the topic. However, They should also take the time to brainstorm and talk with other researchers (and even high school students) about what may be necessary to transition to college. It may not be possible to include every single factor on a survey since that may make the survey pages long, but some thought would generate a list of the most critical factors.

While it is important to include questions on all essential topics in the research question, an "everything-but-the-kitchen-sink" approach is counterproductive. That approach puts an unnecessary burden on the survey respondents. Remember that respondents have agreed to volunteer their time and attention, so they deserve respect in only asking the necessary on-topic questions.

Once the topics are identified, the questions need to be drafted. Questions should be as clear and to the point as possible. Writing questions is not the time for researchers to show off their creative writing skills; a survey is a technical instrument and should be written in a way that is as direct and succinct as possible. Every question should be relevant to every person who is asked to respond. This rule means two things: first, respondents have knowledge about the topic, and second, they have experience with the events, behaviors, or feelings being probed. For example, a sample of 18-year-old respondents should not be asked how they would have advised President Bill Clinton concerning his impeachment. For one thing, few 18-year-olds are likely to have any clue about how to advise a president; moreover, today's 18-year-olds were not even alive during Clinton's impeachment, so that they would have had no experience with the event. In the example of a successful college transition, respondents must understand the phrase "transition to college" and have experienced that transition.



If a survey includes items for which only a portion of respondents will have had experience, it should include a "filter question." A filter question is designed to identify a subset of survey respondents who are asked additional questions not relevant to the entire sample. Using the successful college transition survey mentioned above, if alcohol abuse is determined to be relevant to college success, it would not be appropriate to ask, "How often did you drink alcohol during your first semester in college?" That presupposes alcohol use, and many students may abstain altogether. It would be better to ask a "filter question" like "If you drank alcohol during your first semester of college, please answer questions 13 and 14; otherwise, skip to question 15." Surveys administered online often automate filter questions, so the response will make additional questions appear or disappear, often without the respondent knowing that they answered a filter question.

Responses obtained in survey research are very sensitive to the types of questions asked. Poorly framed or ambiguous questions will likely result in meaningless responses with very little value. Dillman (Dillman, 2011) recommends several rules for creating good survey questions.

Best Practices

- Is the question clear and understandable? Survey questions should be stated in straightforward language, preferably in an active voice, and without complicated words or jargon that a typical respondent may not understand. All questions in the questionnaire should be worded similarly to make it easy for respondents to read and understand. The only exception is if the survey is targeted at a specialized group of respondents, such as doctors or lawyers who use internal jargon in their everyday environment.

- Is the question worded negatively? Negatively worded questions, such as "should your local government not raise taxes," tend to confuse respondents and lead to unreliable responses. Such questions should be avoided. More importantly, in all cases, double-negatives must be avoided.
- Is the question ambiguous? Survey questions should not include words or expressions that different respondents may interpret differently, like "justice." For instance, if the survey includes a question like, "what is your annual income," it is unclear whether it refers to only wages or also dividends, rental, and other income. Different interpretations by different respondents will lead to incomparable responses.
- Does the question have biased or value-laden words? Bias refers to any property of a question that encourages subjects to answer in a certain way. That skews observed responses. It is often difficult to anticipate in advance the biasing wording, but survey questions should be carefully scrutinized to avoid such language to the greatest extent possible.
- Is the question double-barreled? Double-barreled questions are those that can have multiple answers. For example, a question like "Are you satisfied with the hardware and software provided for your job?" may confuse respondents who may be satisfied with the hardware but not the software. It is always best to separate double-barreled questions into two or more questions.
- Is the question too general? Sometimes, questions that are too general may not accurately convey respondents' perceptions. If a survey question asked "How big is your firm," the question is so general that it could be interpreted differently by respondents, ask more specific questions like "how many people does the firm employ," or "what is the firm's annual revenue."
- Is the question too detailed? Avoid unnecessarily detailed questions that serve no specific research purpose. For instance, does the research project require the ages for each child in a household, or is just the number of children enough? Conversely, it may sometimes be better to gather too many details than not enough since analysis may benefit from the extra detail.
- Is the question presumptuous? If a survey asks, "what are the benefits of a tax cut," there is a presumption that the respondent sees the tax cut as beneficial. Many people may not view tax

cuts as beneficial since that generally leads to decreased funding for public services. Questions with built-in presumptions should be avoided on a survey.

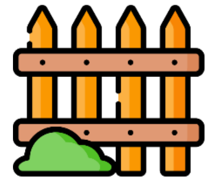
- Is the question imaginary? The hosts of television games like to ask, “if you win a million dollars on this show, how will you spend it?” Most respondents have never been faced with such a large amount of money and have never thought about it, so their answers tend to be random and trite, such as take a tour around the world. Imaginary questions have imaginary answers, and those cannot be used for research inferences.
- Do respondents have the information needed to answer the question correctly? Frequently, the assumption is that subjects have the necessary information to answer a question when, in reality, they do not. For instance, the CEO of a company should not be asked about the day-to-day operational details of their company since they do not work at that level.
- Is there a socially desirable response? Respondents usually try to answer questions in a way that will match social norms. For example, if students were asked if they cheat on exams, they would likely not admit to that behavior since cheating is not socially acceptable.

Response Options

While posing clear and understandable questions is crucial, so is providing respondents with unambiguous response options. This requirement assumes that the questions are closed-ended; respondents are only permitted to select from specific options. Writing closed-ended questions puts a burden on the researcher to provide respondents an adequate set of response options. Researchers should keep the following in mind when writing responses.

- Response options should be mutually exclusive. If a survey asks the respondents to choose an age group and the selections are “less than 20,” “20−30,” “30−40,” “40−50,” “above 50,” then which option should a 30-year-old person select since that age is in two groups? Exclusivity is one of those points about question construction that seems apparent but is easily overlooked.
- Response options should be exhaustive. Every possible response should be covered in the set of response options that are provided. For example, if a survey asks the respondents to indicate sex, then “male” and “female” may not be enough since there are other potential responses. Researchers may want to include an “other” response if some responses may not be listed in the options.

- Responses should be sensitive. For questions that ask for personal information, researchers should include an option like “do not care to respond” or “none are correct” to avoid nonresponse due to potential embarrassment.



Fence-sitting and floating are undesirable respondent behaviors for which researchers should plan. Fence-sitting is when a respondent selects “no opinion” rather than take a stance, while floating is when a respondent selects an opinion when, in fact, they may have none. These behaviors are especially evident when Likert questions are asked. As an example, consider two potential sets of responses for a question about a fictitious “Proposition 100.”

Do you agree with this statement: If Proposition 100 is passed, my taxes will increase?

1. Strongly disagree—disagree—neither agree nor disagree—agree—strongly agree
2. Strongly disagree—disagree—agree—strongly agree

The first set of responses permit a respondent to fence-sit and select a neutral opinion, while the second set forces the respondent to indicate some level of agreement. Either of these response sets could be viable depending on the goal of the research project. For questions that probe socially undesirable behavior (like cheating on exams), it may be appropriate to give respondents the option to remain neutral. At the same time, the researcher may want to force respondents to take a stance in other cases.

A matrix, which lists a set of questions that use the same response categories, creates a compact presentation that is easy to understand and encourages participation. Following is an example matrix for an imaginary set of election propositions.

Do you support these propositions?

Prop	Strongly Support	Support	Do Not Support	Strongly Do Not Support
100	○	○	○	○
115	○	○	○	○
220	○	○	○	○

FIGURE 57: EXAMPLE RESPONSE MATRIX

Responses to closed-ended questions are captured using one of the following response formats.

- **Dichotomous response.** Respondents are asked to select one of two possible choices, such as yes-no or agree-disagree. For example: *Do you think that the death penalty is justified under some circumstances: yes/no.*

- **Nominal response.** Respondents are presented with more than two unordered options. For example: *What is your employment industry: manufacturing / consumer services / retail / education / health care / tourism & hospitality / other.*
- **Ordinal response.** Respondents have more than two ordered options. For example, what is your highest education level*: high school/college degree/graduate studies.*
- **Interval-level response.** Respondents are presented with a 5-point or 7-point Likert, semantic differential, or Guttman scale.
- **Continuous response.** Respondents enter a numeric value, such as their age or tenure in a firm. These responses generally tend to be fill-in-the-blanks.

Of course, surveys need not be limited to closed-ended questions. Researchers can include open-ended questions, which do not include response options, to gather additional details. An open-ended question asks respondents to reply to the question using their own words. These questions are generally used to discover more about a survey participant's experiences or feelings about the survey topic. If, for example, a survey includes closed-ended questions asking respondents about their involvement in extracurricular activities during college, an open-ended question could ask respondents why they participated in those activities or what they gained from their participation. Allowing respondents to reply in their own words can make the experience of completing the survey more satisfying and often reveals new information that had not occurred to the researcher.

Designing Questionnaires

After constructing quality questions and clear response options, researchers also need to craft a high-quality questionnaire, the document that contains those questions and response options.

One of the first things to do is to group survey questions thematically. On a questionnaire about a successful transition from high school to college, perhaps a few questions would ask about study habits, others about friendships, and still others on exercise and eating habits. Those topics would then become the themes around which the survey is organized. There are other potential arrangements. For example, perhaps it would make more sense to present questions using a temporal arrangement, starting with life and habits before college and then after beginning college.

Once similar questions are grouped, it is time to think about how to present the groups. Most survey researchers agree that it is best to begin a survey with questions to make respondents continue (Dillman, 2011). However, there is disagreement about where to place demographic questions concerning a person's age or gender. On the one hand, placing those questions at the beginning of the questionnaire may lead respondents to think the survey is uninteresting and not worth completing. On the other hand, if the survey includes a sensitive or complex topic, such as child sexual abuse, respondents may feel encouraged to continue if the survey starts with "easy" demographic questions.

In truth, the order in which survey questions are presented is best determined by the unique characteristics of the research. Only the researcher can determine how best to order questions after consulting with people who provide feedback. To do so, researchers must consider the unique characteristics of the topic, questions, and, most importantly, the surveyed sample. Keeping in mind the characteristics and needs of the people completing the survey should help researchers determine the most appropriate order to present questions.

It is also relevant to consider the time it will take respondents to complete the questionnaire. Surveys vary in length, from just a page or two to a dozen or more pages, which means they also vary in the time it takes for completion. The length of a survey depends on several factors. First, what is it that the researcher wishes to know? Wanting to understand how grades vary by gender and year in school certainly requires fewer questions than understanding how demographic characteristics shape people's experiences in college. Keep in mind that even if a research question requires a large number of questions, researchers must keep the questionnaire as brief as possible. Any hint that useless questions have been included just for the sake of gathering "just-in-case" data will discourage respondents from completing the survey.

It is also necessary to consider how long respondents are may be willing to spend completing the questionnaire. Asking busy people in a mall to complete a survey may make some resent the intrusion, and they may not be willing to take more than a few minutes. However, people concerned about the survey topic, like a "hot button" political topic, may be willing to more time to complete it.

Some experts advise that surveys take no longer than 15 minutes to complete, while others suggest that up to 20 minutes is acceptable. As with question order, there is no clear-cut, always correct answer about questionnaire length. The study and sample's unique characteristics should be considered to determine how long to make the questionnaire.

An excellent way to estimate the time it will take respondents to complete a questionnaire is through pretesting. Pretesting creates feedback on a questionnaire so researchers can improve it before it is administered. Pretesting can be expensive and time-consuming if the survey is tested on a large sample of people who resemble the sample where it will ultimately be administered. However, there is much to gain by pretesting with a small number of people to whom researchers have easy access. By pretesting a questionnaire, researchers can find out how understandable the questions are, get feedback on question wording and order, find out whether any of the questions are offensive, and learn whether there are places where they should include filter questions. Pretests can also be timed to get an idea of the amount of time it will take respondents and whether there is "wiggle room" to add additional items.

Perhaps this goes without saying, but a questionnaire should also be attractive. A messy presentation style can confuse respondents or, at the very least, annoy them. Be brief, to the point, and as straightforward as possible. Avoid cramming too much into a single page, make the font size readable (12 points is usually good), leave a reasonable amount of space between items, and make sure all instructions are unambiguous. Think about enjoyable books, documents, articles, or web pages and try to mimic the features of those documents in the survey presentation.

Question Sequencing

In general, questions should flow logically from one to the next. Principally, questions should flow from the least sensitive to the most sensitive, from the factual and behavioral to the attitudinal, and the more general to the more specific. Above all, the survey's purpose should guide question sequencing, but here are a few best practices.

- Start with non-threatening questions that can be easily recalled. Good options are demographics (age, gender, education level) for individual-level surveys and "firmographics" (employee count, annual revenues, industry) for firm-level surveys.
- If open-ended questions are used, place them near the end of the survey.
- If the survey follows a historical sequence of events, follow a chronological order from earliest to latest.
- Ask about one topic at a time. When switching topics, use a transition, such as "The next section examines your opinions about…."
- Use filter or contingency questions as needed, such as: "If you answered 'yes' to question 5, please proceed to Section 2. If you answered 'no,' go to Section 6."

Other Golden Rules

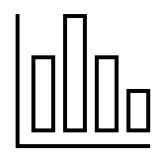
To appropriate an old saying, "do unto your respondents as you would have them do unto you." Be attentive and appreciative of respondents' time, attention, trust, and confidentiality of personal information. Always practice the following strategies for all survey research:

- Time is valuable. Be respectful of the time it takes respondents to complete the survey. Keep the survey as short as possible and limit it to only what is necessary. Respondents do not like spending more than 15 minutes on any survey, though they may take longer for something they consider essential. Longer surveys tend to dramatically lower response rates.
- Always assure respondents about the confidentiality of their responses and how the data will be used (e. g., for academic research), protected, and reported (usually in the aggregate).
- For organizational surveys, assure respondents that they will receive a copy of the final results and ensure that happens.
- Thank respondents for their participation in the study.
- Finally, always pretest the questionnaire, at least using a small convenience sample, before administering it to respondents in a field setting. Such pretesting may uncover ambiguity, lack of clarity, or biases in the questions, which should be eliminated before administering it to the intended sample.

Analysis of Survey Data

From Completed Questionnaires to Data

The goal of survey data analysis is to condense large amounts of raw data into usable and understandable information.



The number of completed questionnaires received divided by the number of questionnaires distributed is called the "response rate." For example, if 100 people were surveyed and 75 returned completed questionnaires, the response rate was 75%. Though response rates vary, and researchers do not always agree about what makes a reasonable response rate, rates in the 20−40% range are not uncommon. Much research has been done to determine how to improve the response rate, and the suggestions offered in the "Non-response bias" section, page 165, will help.

The primary concern with response rates is the problem of nonresponse bias. If only respondents with strong opinions about the study topic return questionnaires, that would bias the analysis. On the other hand, recent research indicates that surveys with low response rates did not significantly differ in findings or sample representativeness (Rindfuss, Choe, Tsuya, Bumpass, & Tamaki, 2015) (Wright G. , 2015). While the ideal response rate can be debated, researchers should aim for the highest response rate possible.

Whatever the response rate, the data must be processed into manageable and analyzable bits. One significant advantage of survey research is that the responses, especially from closed items, can be easily analyzed. For open items, researchers must first create a codebook, which is a document that outlines how the data are translated from words to numbers. For example, Andrew Richards posited an approach to qualitative analysis in a team environment (Richards & Hemphill, 2018). For his study, Richards eventually created a codebook to "…understand how physical education teachers navigate the sociopolitical realities of the contexts in which they work…". Figure 58 shows a small portion of the codebook Richards developed for his project.

Themes	Subthemes	Definitions	Examples from Transcripts
Subject Marginalization	Lack of communication	Teacher believes physical education does not matter due to lack of communication about issues that affect the physical education environment.	"My stressful day, um probably when things pop up that are not…A lot of my stresses get raised from being an activities director. If the school calls me and says now they have to-- they have kids who are not coming, they change times, or I have a different schedule. My stuff is very organized and if it's not where I think it's supposed to be and I need it, that's very stressful for me" (1019, 210-217, individual interview)
	Lack of time and resources	Teacher believes physical education does not matter due to lack of teaching contact time and resources such as materials, equipment for PE, or teaching facilities.	"It's kind of rough because I don't have my own classroom. I don't have my own computer up there. I don't have a room that I can make into a welcoming environment so that's kind of rough" (1018, 110-112, individual interview) "Right now that class is more just like babysitting. It's just a study hall, kind of boring. I don't have a classroom I'm in the gym balcony where the bleachers are at. I don't have space the kids complain" (1018, 120-122, focus group)
	Lack of support	Teacher believes physical education does not matter due to situations in which the physical educator does not feel support for ideas or initiatives.	"I think the colleagues, it wouldn't matter either way outside of the P.E. teachers, and I think the administration wouldn't care either way." (1018, 348-350, individual interview) "At the elementary level that would be a big issue. As they get a little older, you know middle school, high school it's not as much probably fun. They don't see it in their eyes as much fun. The students themselves probably wouldn't care, there'd be a handful." (1019, 307-309, focus group)

FIGURE 58: EXAMPLE CODEBOOK

Data entry is the part of the survey where the data collected are entered into a computer for analysis. Most surveys are conducted online or using electronic devices like tablets, so the data are already in digital format. It is also possible to use a data capture form that can be scanned into a computer. However, data collected using simple "paper and pencil" forms must be manually entered into a computer for analysis. This data entry is a time-consuming and, frankly, tedious task that researchers must slog through before beginning analysis.

After the data are in a digital format, researchers can use several good tools for analysis. Quantitative data can be analyzed with *R*, *SPSS*, or *Excel*. *Excel* is most familiar to students, but its application for statistical analysis is somewhat limited. *SPSS* is very powerful but expensive. *R* is available as a free download and is very powerful but has a steep learning curve.

Once the data are coded into a computer, they must be analyzed. One of the simplest types of analysis is to look for patterns in the data. Univariate analysis is the most basic form of quantitative analysis and involves finding patterns across just one variable. For continuous data, analysis tools include the central measure, histograms, and box plots. For discrete data, analysis tools include frequency distributions and bar plots. The bivariate analysis looks for relationships between two variables, and analysis tools include correlations and scatter plots. The multivariate analysis looks for relationships between three or more variables.

Another standard method of data analysis is hypothesis testing. In this case, researchers have postulated a hypothesis and then conduct research to either support or refute the hypothesis. Any of the pattern-finding tools mentioned in the previous paragraph can be used for hypothesis testing. However, other more powerful tools are also available, like an ANOVA or Kruskal-Wallis H.

Biases in Survey Research

Despite all of its strengths and advantages, survey research is often tainted with one or more of the five following systematic biases that may invalidate some of the inferences derived from such surveys.

- Non-response
- Sampling
- Social desirability
- Recall
- Common method

FIGURE 59: COMMON SURVEY BIASES

Non-response bias. Survey research is generally notorious for its low response rates. A 15–20% response rate is typical in a mail survey, even after two or three reminders. Suppose the majority of the targeted respondents fail to respond to a survey. In that case, a legitimate concern is whether people are not responding due to a systematic reason, which may raise questions about the validity of the study's results. For instance, dissatisfied customers tend to be more vocal about their experience than satisfied customers and are therefore more likely to respond to questionnaire surveys or interview requests. Hence, survey data is likely to have a higher proportion of responses from dissatisfied people than the population from which it is drawn. In this instance, not only will the results lack generalizability, but the observed outcomes may also be an artifact of the biased sample. The following strategies may be employed to improve response rates.

- **Advance notification.** A short letter sent in advance to the targeted respondents soliciting their participation in an upcoming survey can prepare them in advance and improve their propensity to respond. The letter should state the purpose and importance of the study, mode of data collection (like a phone call or a mailed survey), and appreciation for their cooperation. A variation of this technique may request the respondent to return a postage-paid postcard indicating whether they are willing to participate in the study.
- **Relevance of content.** Respondents are more likely to respond to surveys that they consider relevant or important than those that do not matter.
- **Respondent-friendly questionnaire.** Shorter survey questionnaires tend to elicit higher response rates than longer questionnaires. Furthermore, higher response rates are expected if the survey questions are straightforward, nonoffensive, and easy to understand.
- **Endorsement.** Organizational surveys should include an endorsement from a senior executive attesting to the importance of the study. Such endorsement can be in the form of a cover letter or a letter of introduction, improving the researcher's credibility.
- **Follow-up requests.** Multiple follow-up requests may coax some non-respondents to respond, even if their responses are late.
- **Interviewer training.** Response rates for interviews can be improved with trained interviewers using computerized dialers to contact potential respondents.
- **Incentives.** Response rates, at least with specific populations, may increase with the use of incentives in the form of cash or gift cards, giveaways such as pens or stress balls, entry into a lottery, draw or contest, discount coupons, the promise of contributing to charity, or other prizes.
- **Non-monetary incentives.** Businesses, in particular, are more prone to respond to non-monetary incentives than financial incentives. An example of such a non-monetary incentive is a benchmarking report comparing the business's response against the aggregate of all responses to a survey.
- **Confidentiality and privacy.** Assurances that the respondent's private data or responses will not fall into the hands of any third party may help improve response rates.

Sampling bias. Bias can be introduced in a poorly designed survey by using the wrong sample selection. Telephone surveys conducted by calling a random sample of publicly available mobile phone numbers will systematically exclude people who do not own a mobile phone. Another issue concerns people who are unable to answer the phone (for instance, they are at work) when the survey is being conducted and will include a disproportionate number of respondents who have who stay home during much of the day, such as the unemployed, the disabled, and the elderly. Likewise, online surveys tend to include many students and younger people who are constantly on the Internet and systematically exclude people with limited or no access to computers or the Internet, such as the poor and the elderly. Questionnaire surveys tend to exclude children and the illiterate, who cannot read, understand, or meaningfully respond to the questionnaire. A different kind of sampling bias relates to sampling the wrong population, such as asking teachers (or parents) about the academic learning of their students (or children) or asking CEOs about operational details in their company. Biases like these make the sample a poor representative of the intended population and reduce the generalizability of the findings.

Social desirability bias. Many respondents tend to avoid negative opinions or embarrassing comments about themselves, their employers, family, or friends. Negative questions like "do you think your project team is dysfunctional," will likely lead to untruthful responses. This tendency among respondents to "spin the truth" to portray themselves socially desirable is called the "social desirability bias," which hurts the validity of responses obtained from survey research. There is practically no way of overcoming the social desirability bias in a questionnaire survey. However, an astute interviewer may spot inconsistent answers in an interview setting and ask probing questions or use personal observations to supplement respondents' comments.

Recall bias. Responses to survey questions often depend on subjects' motivation, memory, and ability to respond. When dealing with events in the distant past, respondents may not adequately remember their motivations or behaviors. Perhaps their memory of such events may have evolved with time and is no longer retrievable. For instance, if respondents are asked to describe their utilization of computer technology one year ago or even memorable childhood events like birthdays, their response may not be accurate due to difficulties with recall. One possible way of overcoming the recall bias is by anchoring respondent's memory in specific events as they happened, rather than asking them to recall their perceptions and motivations from memory.

Common method bias. Common method bias refers to the amount of spurious covariance shared between independent and dependent variables measured simultaneously, such as in a cross-sectional survey, using the same instrument, such as a questionnaire. In such cases, the phenomenon under investigation may not be adequately separated from measurement artifacts. This bias can be reduced if the independent and dependent variables are measured at different times using a longitudinal survey design. Also, the bias is reduced if the variables are measured using different methods, such as computerized recording of dependent variables versus questionnaire-based self-rating of independent variables.

Key Takeaways

Chapter 8: Survey Research
• Differentiate between cross-sectional and longitudinal surveys. • Compare and contrast the four types of longitudinal surveys: trend, panel, cohort, and retrospective. • Describe the characteristics of compelling questions. • Define the various question response options. • Describe tips for creating good questionnaires. • Discuss how survey data are analyzed. • Discuss bias in survey research.

9: Experimental Research

Introduction

With hundreds, or maybe thousands, of products to sell, how does a merchant determine what to stock? Often that decision is the result of an experiment. The merchant may stock two or three similar items and then check to see what sells best. Part of the experiment may include manipulating the price of the items, their locations in the store, and how they are advertised. In the end, the merchant would have generated reliable data to help determine what products to stock in the future.

Experimental research, often considered the "gold standard" in research designs by researchers who are positivist in outlook, is one of the most rigorous research designs. In this design, the researcher manipulates one or more independent variables ("treatments"), subjects are randomly assigned to different treatment levels, and the results of the treatments on outcomes are observed and analyzed. The unique strength of experimental research is its internal validity due to its ability to link cause to an effect through treatment manipulation while controlling for the spurious effect of extraneous variables.

OBJECTIVES

- Describe experimental research
- Define the terms control group, treatment group, treatment manipulation, random selection, validity
- Describe how internal validity can be improved
- Describe two-group designs
- Describe factorial designs
- Describe hybrid designs
- Describe quasi-experimental designs
- List the strengths and weaknesses of experimental formats

One example of experimental research can be found in Alcott's study of behavioral interventions (Allcott & Rogers, 2014). In this study, the results of the *Opower* program were evaluated. *Opower* was a company that sent a "home energy report" to more than six million homes representing 85 utilities across the United States. The reports were designed to use social pressure to get homeowners to moderate their electricity use. For Alcott's experiment, the three different sites were compared using pre- and post- intervention electricity usage statistics. He found that the initial home energy report caused high-frequency "action and backsliding," but the cycles seem to attenuate over time. He also found that if the reports were discontinued, the effect was still relatively persistent. Finally, he found that consumers are slow to habituate. While this study was focused on a single intervention involving energy conservation, it may be interesting to speculate if similar intervention efforts in areas like dieting, exercise, and smoking cessation would experience similar results.

Experimental research is best suited for explanatory research, where the goal of the study is to examine cause-effect relationships rather than for descriptive or exploratory research. It also works well for research that involves a relatively limited and well-defined set of independent variables that can either be manipulated or controlled. The conduct of experimental research is generally found in one of two settings.

- Laboratory experiments conducted in laboratory (artificial) settings tend to be high in internal validity, but this comes at the cost of low external validity because the artificial setting in which the study is conducted may not reflect the real world. As one example of a laboratory study, Berger and Iyengar (Berger & Iyengar, 2013) studied how the medium, oral vs. written, shapes the message in word-of-mouth communications. As part of their research, they conducted three laboratory studies and then collected field data to validate the laboratory studies. They found a difference in how people communicated about brand-name products when the channel was oral rather than written. They found, specifically, that written communication (like online chat) led to people mentioning more exciting products and brands than when they talked about them orally.
- Field experiments conducted in field settings such as in an organization can be high in external validity. However, such experiments are relatively rare because of the difficulties of manipulating treatments and controlling for extraneous effects in a field setting. As an example of a field experiment, Cai, Chen, and Fang (Cai, Chen, & Fang, 2009) observed diners in restaurants to determine if knowing what other people had ordered influenced what they in turn ordered. They found that the demand for the top five dishes was increased by about 13−18% when the popularity ratings were revealed to customers.

It is common for researchers to combine experimentation settings in one project so one setting can be validated in another setting. Gainsbury and Blaszczynski (Gainsbury & Blaszczynski, 2011) specifically researched the question of whether laboratory and field experiments came to the same result. They set up a gambling laboratory and had 127 students play an electronic gambling machine that included a warning about gambling presented in two different ways: pop-up and static sign. They then interviewed the gamblers to see what they remembered about the warnings. Then they went to a casino and repeated the experiment with gamblers who happened to be present. They found that the “venue participants” were less likely to complete surveys, and when they did, they provided less information than the laboratory participants. They concluded that while both settings provided valuable information about gambling behavior, care must be taken in generalizing the laboratory results to the “real world.”

Experimental research designs can be grouped into two broad categories: actual and quasi-experimental designs. Both designs require manipulating a treatment, but actual experiments also require random participant assignment while quasi-experiments do not. As an example of a quasi-experiment, McElroy and Morrow (McElroy & Morrow, 2010) evaluated the result of an office redesign project in a large financial services company. They compared two groups of employees, those who were reassigned to a new office design and those who were not. They surveyed 127 employees who had moved and 144 who had not and then compared the perceptions of the work space and organization’s culture for both groups. They found that those who moved reported less workspace and more distractions than those who did not, but the results varied by the respondent's age. They also reported more favorable perceptions of the organization’s culture and work-related attitudes with no age difference. Because the participates had not been randomly assigned to groups, this was a quasi-experimental design.

Basic Concepts

All experimentation designs include the following four characteristics.

FIGURE 60: EXPERIMENTAL DESIGN CHARACTERISTICS

Treatment and control groups. In experimental research, some subjects are administered one or more experimental stimuli called treatments (the treatment group) while other subjects are not given such a stimulus (the control group). The treatment may be considered successful if subjects in the treatment group rate more favorably on outcomes than control group subjects. Multiple experimental stimuli may be administered, in which case there may be more than one treatment group. For example, to test an advertisement's effectiveness, researchers may expose three different groups of people to different levels of advertising. One treatment group would view a printed article with an embedded advertisement. A second treatment group would view a television program with an embedded advertisement. The third group, the control group, would view media with no advertising. After the exposure period, the groups could be surveyed to determine which remembered the advertising best.

Treatment manipulation. Treatments are the unique feature of experimental research that sets this design apart from all other research methods. Treatment manipulation helps control for the “cause” in cause-effect relationships. Treatments must be checked using various pilot tests before the start of the experiment. Any measurements conducted before the treatment is administered to the subjects are called pretest measures, and those conducted after the treatment are post-test measures.

Random selection and assignment. For the experiment to be valid, the treatment and control groups must be as nearly equivalent as possible. When working with humans, it is impossible to have identical groups—even identical twins have some differences. However, as much as possible, researchers must make an effort to make the groups equivalent. One primary way to accomplish this is through both random selection and random assignment of test subjects. Random selection is the process of randomly drawing a sample from a population or a sampling frame. This approach is typically employed in survey research and assures that each unit in the population has an equal chance of being selected into the sample. Random assignment is the process of randomly assigning subjects to experimental or control groups. Random assignment is a standard practice in actual experimental research to ensure that treatment groups are similar to each other and the control group before treatment administration. Random selection is related to sampling and is more closely related to the external validity of findings. In contrast, the random assignment of participants is related to design and improves internal validity. By including random selection and random assignment in the experiment’s design, researchers attempt to mitigate potential differences in two groups of subjects.

Threats to internal validity. Although experimental designs are considered more rigorous than other research methods in terms of the internal validity of their inferences (by their ability to control causes through treatment manipulation), they are not immune to internal validity threats. As an example, imagine that researchers wanted to study the effect of a unique mathematics tutoring program. They would want to plan for the following internal validity threats.

- **History threat** is the possibility that the observed outcomes (dependent variables) are caused by extraneous or historical events rather than experimental treatment. For instance, students' post-tutoring mathematics score improvement may have been caused by their preparation for a mathematics examination rather than the tutoring program.
- **Maturation threat** is the possibility that observed effects are caused by natural maturation of subjects (their general improvement in their intellectual ability to understand complex concepts) rather than the experimental treatment. Perhaps the mathematics students' post-tutoring scores improved because they matured during the study and better understand mathematics concepts.
- **Test threat** is a threat in pre/post designs where their pretest responses condition subjects' posttest responses. For instance, if students remember their answers from the pretest evaluation, they may repeat them in the post-test exam.
- **Instrumentation threat** is the possibility that the difference between pretest and post-test scores is not due to the mathematics tutoring program but to changes in the administered test, such as the post-test having a higher or lower degree of difficulty.
- **Mortality threat** is the possibility that subjects may be dropping out of the study due to a systematic reason not necessarily related to the treatment. If the low-performing students drop out, the post-test results will be artificially inflated by the preponderance of high-performing students.
- **Regression threat** ("regression to the mean") is the tendency of a post-test to regress toward the measure's mean rather than the anticipated direction. For instance, if students in the mathematics study scored high on the pretest, they will score lower on the post-test (closer to the mean) because their initial high scores could have been a statistical aberration. It would also be true that students who under-performed on the pretest would tend to score closer to the mean on the post-test, which would falsely appear to be due to the tutoring treatment.

Two-Group Experimental Designs

The most straightforward experimental designs are two group designs involving one treatment group and one control group. They are ideal for testing the effects of a single independent variable manipulated as a treatment. The two basic two-group designs are the pretest/post-test control group design and the post-test-only control group design, while variations may include covariance designs. These designs are often depicted using a standardized notation, as found throughout the remainder of this chapter. In that notation, *R* represents the random assignment of subjects to groups. *X* represents the group's treatment. *O* represents pretest or post-test observations of the dependent variable (with different subscripts to distinguish between pretest and post-test observations).

Pretest/Post-test Control Group

Subjects are randomly assigned to treatment and control groups, and initial measurement of the dependent variables of interest (pretest) is conducted. The treatment group is administered a treatment (the independent variable), and the dependent variables are measured again (post-test). The notation for this design is shown in Table 9, where each group's process should be read from left to right. Thus, the Treatment Group (top line) is first assigned using a random process; then a pretest is administered (Observation 1), then treatment of some sort is applied; finally, a post-test is administered. The Control Group (bottom line) follows the same process, except there is no applied treatment.

R	O_1	X	O_2	Treatment Group
R	O_3		O_4	Control Group

TABLE 9: PRETEST/POST-TEST DESIGN

The effect, *E*, of the treatment is the difference between the scores of the two groups, as shown in Equation 1.

$$E = (O_2 - O_1) - (O_4 - O_3) \qquad \text{Equation 1}$$

Statistical analysis of this design applies an ANOVA between the treatment and control groups. The pretest/post-test design mitigates several threats to internal validity, such as maturation, testing, and regression, since these threats can be expected to influence both treatment and control groups similarly. The selection threat is controlled via the random assignment of members. However, additional threats to internal validity may exist. For instance, mortality can be a problem if there are differential dropout rates between the two groups. The pretest measurement may bias the post-test measurement (mainly if the pretest introduces unusual topics or content).

Post-Test-only Control Group

This design is a simpler version of the pretest/post-test design where pretest measurements are omitted. The design notation is shown in Table 10.

R	X	O_1	Treatment Group
R		O_2	Control Group

TABLE 10: POST-TEST-ONLY DESIGN

The treatment effect is measured as the difference in the post-test scores between the two groups, as shown in Equation 2.

$$E = (O_1 - O_2) \qquad \text{Equation 2}$$

The appropriate statistical analysis of this design is a two-group ANOVA. The simplicity of this design makes it more attractive than the pretest/post-test design in terms of internal validity. This design controls maturation, testing, regression, selection, and pretest/post-test interaction, though the mortality threat may continue to exist.

Covariance

Sometimes, the measure of a dependent variable may be influenced by an extraneous variable called a covariate. Covariates are those variables that are not of central interest to an experimental study but should be controlled to eliminate their potential effect on the dependent variable. This design allows for more accurate detection of the effects of the independent variables. A covariance design is a particular type of pretest/post-test control group design. The pretest measure is essentially a measurement of the covariates rather than the dependent variables. The design notation is shown in Table 11, where *C* represents the covariates:

R	C	X	O1	Treatment Group
R	C		O2	Control Group

TABLE 11: COVARIATE DESIGN

Because the pretest measure is not a measurement of the dependent variable but rather a covariate, the treatment effect is measured as the difference in the post-test scores between the treatment and control groups as in Equation 3.

$$E = (O_1 - O_2) \qquad \text{Equation 3}$$

Due to the presence of covariates, the correct statistical analysis of this design is a two-group ANCOVA. This design has all the advantages of a post-test-only design but improved internal validity due to controlling covariates. Covariance designs can also be extended to pretest/post-test control group design.

Factorial Designs

Two-group designs are inadequate if the research project requires manipulation of two or more independent variables (treatments). In such cases, four or higher-group designs are required. Such designs, quite popular in experimental research, are commonly called factorial designs. Each independent variable is called a factor, and each sub-division of a factor is called a level. Factorial designs enable researchers to examine the individual effect of each treatment on the dependent variables (called main effects) and their combined effect (called interaction effects).

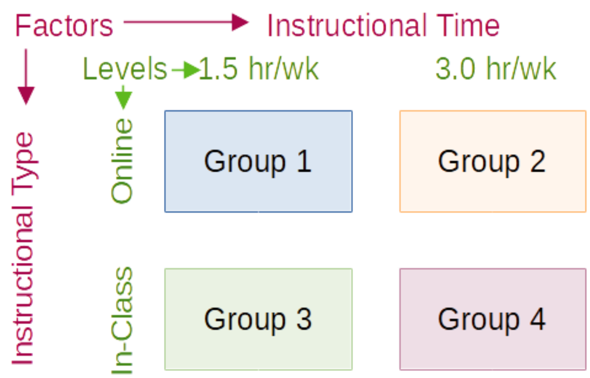


FIGURE 61: 2X2 FACTORIAL DESIGN

The most basic factorial design is a 2x2 factorial design, consisting of two treatments, each with two levels (high/low or present/absent). For instance, suppose a research project compares the learning outcomes of two different instructional techniques (online and in-class) and instruction time (1.5 or 3 hours per week). In this case, there are two factors: instructional type and instructional time, each with two levels (in-class and online for instructional type, and 1.5 and 3 hours/week for instructional time), as shown in Figure 61. If another level of instructional time (maybe 6 hours/week) is desired, then the second factor will consist of three levels, and a 2x3 factorial design is required. On the other hand, if a third factor, such as group work (present versus absent), is desired, then a 2x2x2 factorial design is needed. Each number in this notation is a factor, and the value of that number is the number of levels.

Factorial designs can also be depicted using a design notation, such as that shown in Table 16.

R	X_{11}	O_1	Group 1
R	X_{12}	O_2	Group 2
R	X_{21}	O_3	Group 3
R	X_{22}	O_4	Group 4

TABLE 12: 2X2 FACTORIAL DESIGN

R represents a random assignment of subjects to treatment groups. X represents the treatment groups where the subscript is the level of each factor, so 11 is 1.5 hr/wk and online. *O* represent observations of the dependent variable. Notice that the 2x2 factorial design has four treatment groups, corresponding to the four combinations of the two levels of each factor. A 2x3 design will have six treatment groups, and a 2x2x2 design will have eight treatment groups.

As a rule of thumb, each cell in a factorial design should have a minimum sample size of 20 (this estimate is based on medium effect sizes). So a 2x2x2 factorial design requires a minimum total sample size of 160 subjects, with at least 20 subjects in each cell. The cost of data collection can increase substantially with more levels or factors in a factorial design. Sometimes, cells in these factorial designs may not receive any treatment due to resource constraints, called incomplete factorial designs. However, incomplete designs decrease the ability to draw inferences about those factors.

In a factorial design, the main effect exists if the dependent variable shows a significant difference between multiple levels of one factor and at all levels of other factors. No change in the dependent variable across factor levels is the null case (baseline), from which main effects are evaluated. In the above example, perhaps the main effect of instructional type, instructional time, or both can be seen on learning outcomes. An interaction effect exists when the effect of differences in one factor depends on a second factor. In the above example, if instructional type affects learning outcomes more significantly for 3 hours/week of instructional time than for 1.5 hours/week, then there is an interaction effect between instructional type and instructional time on learning outcomes. Note that interaction effects dominate and make main effects irrelevant, so it is not meaningful to interpret main effects if interaction effects are significant.

Hybrid Experimental Designs

Hybrid designs are those that are formed by combining features of more conventional designs. Three such hybrid designs are randomized blocks design, Solomon four-group design, and switched replications design.

Randomized Block

The randomized block design is a variation of the pretest/post-test control group or the post-test-only control group design. The subject population is grouped into relatively homogeneous subgroups (called blocks) within which the experiment is replicated. For instance, to replicate the same post-test-only design among university students and working professionals (two homogeneous blocks), subjects in both blocks are randomly split between a treatment group (receiving the same treatment) and a control group (see Table 13). The purpose of this design is to reduce the “noise” or variance in data that may be attributable to differences between the blocks so that the actual effect of interest can be detected more accurately.

University Students	R	X	O_1
University Students	R		O_2
Professionals	R	X	O_3
Professionals	R		O_4

TABLE 13: RANDOMIZED BLOCK DESIGN

Solomon Four-group

In this design, the sample is divided into two treatment groups and two control groups. One treatment group and one control group receive the pretest, and the other two groups do not. This design represents a combination of the pretest/post-test and post-test-only designs. It is intended to test for the potential biasing effect of pretest measurement on posttest measures that tends to occur in pretest/post-test designs but not in post-test only designs. The design notation is shown in Table 14.

R	O_1	X	O_2
R	O_3		O_4
R		X	O_5
R			O_6

TABLE 14: SOLOMON FOUR-GROUP DESIGN

Switched Replication

This design is a two-group design implemented in two phases with three waves of measurement. The treatment group in the first phase serves as the control group in the second phase, and the control group in the first phase becomes the treatment group in the second phase, as illustrated in Table 15. In other words, the original design is repeated or replicated temporally with treatment/control roles switched between the two groups. By the end of the study, all participants will receive the treatment during the first or the second phase. This design is most feasible in organizational contexts where organizational programs (e.g., employee training) are implemented in a phased manner or are repeated at regular intervals.

R	O_1	X	O_2		O_3
R	O_4		O_5	X	O_6

TABLE 15: SWITCHED REPLICATION DESIGN

Quasi-Experimental Designs

Quasi-experimental designs are almost identical to actual experimental designs but lacking one key ingredient: random assignment. For instance, one entire class section is used as the treatment group, and another section of the same class is used as the control group. This lack of random assignment potentially results in non-equivalent groups, such as one group possessing greater mastery of a specific content than the other group, say by having a better teacher in a previous semester, introducing the possibility of selection bias. Quasi-experimental designs are therefore inferior to actual experimental designs in interval validity due to the presence of a variety of selection-related threats:

- **Selection-maturation threat**. The treatment and control groups maturing at different rates
- **Selection-history threat**. The treatment and control groups being impacted differently by extraneous or historical events
- **Selection-regression threat**. The treatment and control groups regressing toward the mean between pretest and post-test at different rates
- **Selection-instrumentation threat**. The treatment and control groups responding differently to the measurement
- **Selection-testing**. The treatment and control groups responding differently to the pretest
- **Selection mortality**. The treatment and control groups are demonstrating differential dropout rates.

Given these selection threats, it is generally preferable to avoid quasi-experimental designs.

Many actual experimental designs can be converted to quasi-experimental designs by omitting random assignments. For instance, the quasi-equivalent version of the pretest/post-test control group design is called the nonequivalent groups design, as shown in Table 16, with random assignment *R* replaced by non-equivalent (non-random) assignment *N*.

N	O_1	X	O_2	Treatment Group
N	O_3		O_4	Control Group

TABLE 16: NONEQUIVALENT GROUPS DESIGN

Likewise, the quasi-experimental version of switched replication design is called Nonequivalent Switched Replication Design, as illustrated in Table 17.

N	O_1	X	O_2		O_3
N	O_4		O_5	X	O_6

TABLE 17: NONEQUIVALENT SWITCHED REPLICATION DESIGN

In addition, there are quite a few unique non-equivalent designs without corresponding true experimental design cousins, including those listed below.

Regression-discontinuity

This design is a non-equivalent pretest/post-test design where subjects are assigned to a treatment or control group based on a cutoff score on a pre-program measure. For instance, severely ill patients may be assigned to a treatment group to test the efficacy of a new drug or treatment protocol, and those who are mildly ill are assigned to the control group. In another example, students who lag on standardized test scores may be selected for a remedial curriculum program to improve their performance. However, those who score high are not selected for the program. The design notation can be represented as in Table 18, where *C* represents the cutoff score:

C	O_1	X	O_2	Treatment Group
C	O_3		O_4	Control Group

TABLE 18: REGRESSION-DISCONTINUITY DESIGN

Because of the use of a cutoff score, the observed results may be a function of the cutoff score rather than the treatment, which introduces a new threat to internal validity. However, using the cutoff score also ensures that limited or costly resources are distributed to people who need them the most rather than randomly across a population while simultaneously allowing a quasi-experimental treatment. The control group scores in the regression-discontinuity design do not serve as a benchmark for comparing treatment group scores, given the systematic non-equivalence between the two groups. Instead, if there is no discontinuity between pretest and post-test scores in the control group, but there is in the treatment group, it is evidence of the treatment effect.

Since the regression-discontinuity design does not include random assignment, researchers should plan for 2.75 times as many participants as in a random design. Thus, if a researcher determines that 100 participants are needed for some experiment using random assignment procedures, then 275 should be found for a regression-discontinuity design experiment.

Finally, since researchers can determine which participants get the treatment, this design mitigates some of the ethical issues surrounding a random-assignment experimental design. Participants who need the treatment are placed in that group while those in the control group do not.

Proxy Pretest

This design, shown in Table 19, looks very similar to the nonequivalent groups (pretest/post-test) design, with one critical difference: the pretest score is collected after the treatment is administered. A typical application of this design is when a researcher is brought in to test the program's efficacy (e. g., an educational program) after the program has already started. Pretest data is not available. Under such circumstances, the best option for the researcher is often to use a different prerecorded measure, such as students' grade point averages before the start of the program, as a proxy for pretest data. A variation of the proxy pretest design uses subjects' post-test recollection of pretest data, which may be subject to recall bias. However, it may provide a measure of perceived gain or change in the dependent variable.

N	O_1	X	O_2	Treatment Group
N	O_3		O_4	Control Group

TABLE 19: PROXY PRETEST DESIGN

Separate Pretest/Post-test Samples

This design is helpful if it is not possible to collect pretest and post-test data from the same subjects for some reason. As shown in Table 20, there are four groups in this design, but two groups come from a single nonequivalent group, while the other two groups come from a different non-equivalent group. For instance, to test customer satisfaction with a new online service implemented in one city but not in another, customers in the first city serve as the treatment group. Those in the second city constitute the control group. Suppose it is not possible to obtain pretest and post-test measures from the same customers. In that case, customer satisfaction is measured at one point in time. The new service program is implemented, and then customer satisfaction (with a different set of customers) is measured after the program has started. Customer satisfaction is also measured in the control group simultaneously but without the new program implementation. The design is not particularly strong because changes in any specific customer's satisfaction score before and after the implementation cannot be examined, only the average customer satisfaction scores. Despite the lower internal validity, this design may still help collect quasi-experimental data when pretest and post-test data are not available from the same subjects.

N_1	O_1	
N_1	X	O_2
N_2	O_3	
N_2		O_4

TABLE 20: SEPARATE PRETEST/POST-TEST SAMPLES DESIGN

Nonequivalent Dependent Variable

This design is a single-group pre/post-quasi-experimental design with two outcome measures. One measure is theoretically expected to be influenced by the treatment, and the other measure is not. For instance, a new calculus curriculum for high school students would be anticipated to influence post-test calculus scores but not algebra scores. However, the post-test algebra scores may still vary due to extraneous factors such as history or maturation. Hence, the pre/post algebra scores can be used as a control measure, while the pre/post-calculus scores are considered the treatment measure. The design notation, shown in Table 21, indicates the single group, *N*, followed by pretest O_1 and post-test O_2 for both calculus and algebra for the same group of students. This design weakens internal validity, but its advantage lies in not using a separate control group.

A variation of the nonequivalent dependent variable design is a pattern matching design. This variation employs multiple outcome variables and a theory that explains how much each variable will be affected by the treatment. The researcher can then examine if the theoretical prediction is matched in actual observations. The pattern-matching technique effectively alleviates internal validity concerns in the original nonequivalent dependent variable design based on the degree of correspondence between theoretical and observed patterns.

N	O_1	X	O_2
N	O_3		O_4

TABLE 21: NONEQUIVALENT DEPENDENT VARIABLE DESIGN

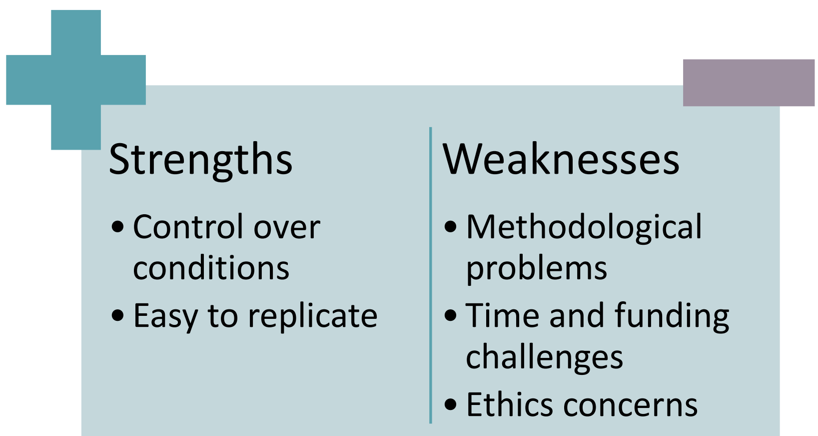


FIGURE 62: STRENGTHS AND WEAKNESSES OF EXPERIMENTAL DESIGN

Strengths and Weaknesses of Experimental Research

Strengths

A strength of the experimental method, particularly in cases where experiments are conducted in lab settings, is that the researcher has substantial control over the conditions to which participants are subjected. Experiments are also generally easier to replicate than other data collection methods, mainly when an experiment has been conducted in a lab setting.

Proper experimental treatment design is a critical task in experimental design because the treatment is the raison d'etre of the experimental method and must never be rushed or neglected. These steps help in the design of an adequate and appropriate task.

- use pre-validated tasks if available
- conduct treatment manipulation checks to check for the adequacy of such tasks by debriefing subjects after performing the assigned task
- conduct pilot tests – repeatedly, if necessary
- use tasks that are simpler and familiar for the respondents than tasks that are complex or unfamiliar

Weaknesses

On the other hand, experimental research is one of the most challenging research designs and should not be taken lightly. This type of research is often beset with a multitude of methodological problems, like those listed below.

- Though experimental research requires theories for framing hypotheses for testing, much of current experimental research is atheoretical. Without theories, the hypotheses being tested tend to be ad hoc, possibly illogical, and meaningless.
- Many of the measurement instruments used in experimental research are not tested for reliability and validity and are incomparable across studies. Consequently, results generated using such instruments are also incomparable.
- Much experimental research uses inappropriate research designs, such as irrelevant dependent variables, no interaction effects, no experimental controls, and nonequivalent stimulus across treatment groups. Findings from such studies tend to lack internal validity and are highly suspect.
- The treatments (tasks) used in experimental research may be diverse, incomparable, inconsistent across studies, and sometimes inappropriate for the subject population. For instance, undergraduate student subjects are often asked to pretend that they are marketing managers and asked to perform a complex budget allocation task with no experience or expertise. The use of such inappropriate tasks introduces new threats to internal validity (i. e., subject's performance may be an artifact of the content or difficulty of the task setting), generates non-interpretable and meaningless findings, and makes integrating findings across studies impossible.

Time, other resources such as funding, and even the topic may limit a researcher's ability to experiment. For researchers in the medical and health sciences, experimenting could require denying needed treatment to patients, which is an explicit ethical consideration. Even those whose research may not involve administering medications or treatments may be limited in their ability to conduct a classic experiment. For example, in business or economics experiments, it may not be ethical to provide a financial or another reward to members of the experimental group but not the control group.

One last concern for researchers opting for an experimental design is ethics. While research ethics was more thoroughly covered in Chapter 3, there is an ethical concern for experimentation that is not found in other research designs. As an extreme example of this ethical concern, imagine a drug manufacturer has found a drug that seems to cure Alzheimer's. To test the efficacy of that drug, they would run an experiment where a sample of people with Alzheimer's would be divided at random into two experimental groups. One group would be given the treatment while the other would not. While this type of drug testing is common, there would seem to be an ethical issue in providing one group of people with a potential cure but withholding that drug from another group. To be sure, business research does not rise to the level of denying a group of customers a life-saving drug. However, there could be harm to a group of customers denied some benefit given to another group.

Key Takeaways

Chapter 9: Experimental Research
• Describe experimental research • Define the terms control group, treatment group, treatment manipulation, random selection, validity • Describe how internal validity can be improved • Describe two-group designs • Describe factorial designs • Describe hybrid designs • Describe quasi-experimental designs • List the strengths and weaknesses of experimental formats

Qualitative Methods

Qualitative methods are based on the evaluation of non-numeric data, like photographs and text documents. These methods include activities like fieldwork, unobtrusive, and interpretive research methods.

10: Interviews

What Is Interview Research?

Today's young men are delaying their entry into adulthood. That is a way of saying they are "totally confused," "cannot commit to their relationships, work, or lives," and are "obsessed with never wanting to grow up." At least, according to sociologist Michael Kimmel. He interviewed 400 young men, ages 16 to 26, over four years across the United States to learn how they made the transition from adolescence into adulthood. Since the results of Kimmel's research were published in 2008 (Kimmel, 2008), his work has made quite a splash. Featured in news reports, on blogs, and in many book reviews, some claim Kimmel's research "could save the humanity of many young men."[28] Whatever is correct about Kimmel's research, one thing remains true: We surely would not know nearly as much as we now do about the lives of many young American men were it not for interview research.

Objectives
• Discuss the role of the interviewer • Conducting and analyzing a qualitative interview • Conducting and analyzing a quantitative interview • Discuss the strengths and weaknesses for both qualitative and quantitative interviews • Discuss using focus groups in a research project

Knowing how to create and conduct a good interview is essential for researchers, especially those interested in qualitative research.

Beyond research, many professionals use interviews in their daily routines.

- Market researchers learn how to increase sales.
- Journalists gather information from people as diverse as VIPs to random people on the street.
- Television hosts help viewers get to know guests on their shows.
- Employers make decisions about job offers.
- Radio hosts greet call-in participants.

[28] This quote is from a review of *Guyland* by Arlie Hochschild. It was found at Guyland: The Perilous World Where Boys Become Men by Michael Kimmel, Paperback | Barnes & Noble® (barnesandnoble.com)

From the research perspective, interviews are a data collection method that involves two or more people exchanging information through a series of questions and answers. A researcher designs the questions to elicit information from interview participant(s) on a specific topic or set of topics. Typically interviews involve an in-person meeting between two people, an interviewer and an interviewee. Nevertheless, interviews need not be limited to two people, nor must they occur in person.

Interviews are an excellent way to gather detailed information. They also have an advantage over surveys; with a survey, if a participant's response sparks some follow-up question, researchers generally do not have an opportunity to ask for more information. What they get is what they get. In an interview, however, because researchers are talking with the study participants in real-time, they can ask follow-up questions and help clarify the responses. Thus, interviews are a valuable method to determine the "story" behind the responses in a written survey.

Interviews are also helpful when:

- the research topic is rather complex
- the question being asked requires an explanation
- the answers to the questions may not be immediately apparent to participants who may need some time to work through their responses.

Also, if the research topic is one about which people will likely have a lot to say or desire to explain or describe some process, interviews may be the best method.

In sum, interview research is instrumental when the following are true.

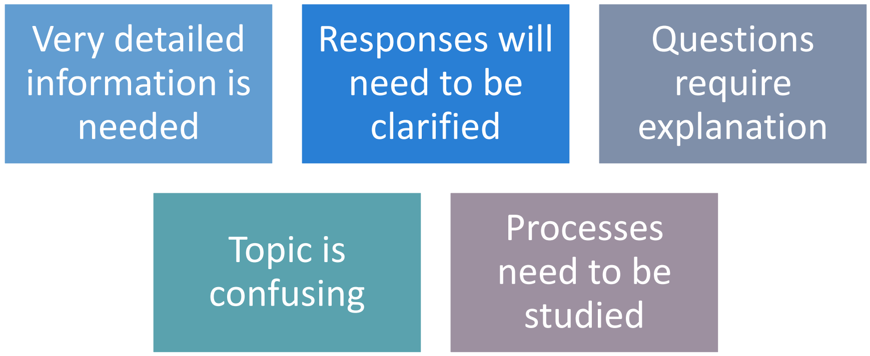


FIGURE 63: INDICATORS FOR INTERVIEW RESEARCH

Role of the Interviewer

The interviewer has a complex and multi-faceted role in the interview process and includes the following tasks:

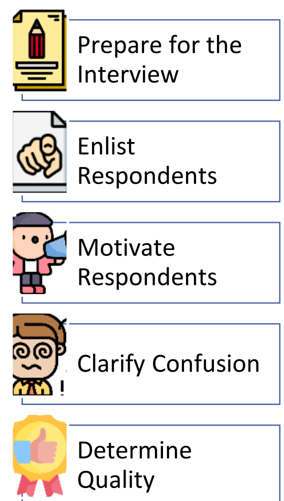


FIGURE 64: INTERVIEWER PREPARATION

- **Prepare for the interview**. Since the interviewer is at the forefront of the data collection effort, the data collected depends heavily on how well the interviewer is trained to do the job. The interviewer must be trained in the interview process and the survey method and be familiar with the purpose of the study, how responses will be stored and used, and sources of interviewer bias. Interviewers should also rehearse and time the interview before the formal study.
- **Locate and enlist the cooperation of respondents**. Particularly in cases where the interview will occur in the participant's home, the interviewer must locate the address and work around the respondents' schedule, sometimes at undesirable times such as weekends. They should also be like a salesperson and "sell" the idea of participating in the study.
- **Motivate respondents**. Respondents often feed off the motivation of the interviewer. If the interviewer is disinterested or inattentive, respondents will not be motivated to provide helpful or informative responses. The interviewer must demonstrate enthusiasm about the study,

communicate the importance of the research to respondents, and be attentive to respondents' needs throughout the interview.

- **Clarify any confusion or concerns**. Interviewers must be able to think on their feet and address unanticipated concerns or objections raised by respondents. Additionally, they should ask probing questions as necessary even if such questions are not in the script.
- **Determine the quality of the response**. The interviewer is in the best position to judge the quality of information collected and supplement verbal responses obtained by recording personal observations of gestures and body language.

Qualitative Interview Techniques

Qualitative interviews are sometimes called "intensive" or "in-depth" interviews. These interviews are semistructured, where researchers will have a particular topic for the interview. However, questions are open-ended and may not be asked in the same way or precisely the same order to each respondent. During qualitative interviews, the primary aim is to hear from respondents about what they think is important and to hear it in their own words. This section considers conducting qualitative interviews, analyzing interview data, and the strengths and weaknesses of this method.

Conducting Qualitative Interviews

Qualitative interviews might feel more like a conversation than an interview with respondents. However, the researcher usually guides the conversation to gather information on a specific topic from a respondent. Qualitative and quantitative interviews are different because qualitative interviews are semistructured and contain open-ended questions, where quantitative interviews are structured and contain close-ended questions. Open-ended questions are more demanding of participants than closed-ended questions since they require participants to develop their own words, phrases, or sentences to respond.

In a qualitative interview, researchers usually use a list of topics or questions to be covered during the interview. The list is called a "guide" because it is that—it is used to guide the interview, but it is not set in stone. Think of an interview guide like an agenda for a meeting, it contains the goals to be accomplished, but it is not critical if an item is skipped or if the order is changed somewhat.

Interview guides outline essential issues, but because participants are asked to provide answers in their own words and raise points they believe are essential, each interview is likely to flow differently. The opening question in an in-depth interview may be the same across all interviews, but from that point on, what the participant says will shape how the interview proceeds. Many researchers believe that this free flow of topics makes in-depth interviewing exciting. It is also what makes in-depth interviewing rather challenging to conduct. It takes a skilled interviewer to ask questions and listen to respondents. Interviewers must also pick up on cues about when to follow up, move on, and let the participant speak without guidance or interruption.

Interview guides tend to list topics or even specific questions. However, the format of an interview guide might depend on the researcher's style, experience, and comfort level as an interviewer or with the topic. For example, in interviews of young people about their experiences with workplace sexual harassment, the guide may be topic-based with few specific questions in the guide. Instead, it could contain only an outline of essential topics listed in the most sensible order for discussion, noted on a sheet of paper.

Of course, interview guides do not appear out of thin air. They are the result of thoughtful and careful work on the part of a researcher. The topics and questions are organized thematically and in the order in which they are likely to proceed. However, the flow of a qualitative interview is in part determined by what a respondent has to say. Sometimes qualitative interviewers may create two versions of the interview guide. One version contains a brief outline of the interview, perhaps with just topic headings, and another version contains detailed questions underneath each topic heading. In this case, the researcher might use the very detailed guide to prepare and practice the interviews and then bring the brief outline to the interview. Instead of a very long list of detailed questions, bringing an outline to the interview encourages the researcher to listen to what the participant is saying. An overly detailed interview guide will be challenging to navigate through during an interview. It could give respondents the impression that the interviewer is more interested in the questions than in the participant's answers.

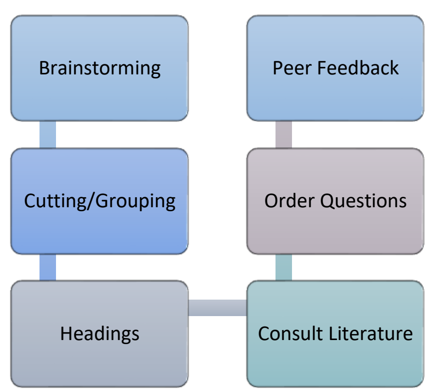


FIGURE 65: STEPS TO CREATE AN INTERVIEW GUIDE

Brainstorming, often with colleagues, is an excellent first step in constructing an interview guide. There are no rules at the brainstorming stage—list all the topics and questions that come to mind when thinking about the research question. Once a good list is created, it can be pared down by cutting questions and topics that seem redundant and then grouping like questions and topics together. It is at this point that headings for grouped categories are developed. Another important avenue of approach is to consult scholarly literature to determine what kinds of questions other interviewers have asked in similar studies. As with quantitative survey research, it is best not to place very sensitive or potentially controversial questions at the beginning of the qualitative interview guide. Participants need the opportunity to warm up to the interview and to feel comfortable talking with the interviewer. Finally, it is crucial to get feedback on the interview guide as it is being developed. Researchers should ask peers for guidance and suggestions once they develop what they think is a pretty strong guide. The chances are that peer reviewers will find ways to improve the guide.

There are a few guidelines worth noting about the specific questions in the guide.

- Avoid questions that can be answered with a simple yes or no.
- If yes/no questions must be asked, then include follow-up questions. One of the benefits of qualitative interviews is that participants can be asked for more information.
- While follow-up questions are appropriate, "why" should be avoided since this particular question can be construed as confrontational. Instead of "why," something like, "Could you tell me a little more about that?" is a good option.
- Leading questions should be avoided. For example, rather than asking, "Don't you agree that people who spend money frivolously are selfish?" ask, "What comes to mind when you hear that someone has spent money frivolously?"
- Keep most, if not all, questions open-ended. The key to a successful qualitative interview is to allow participants to share information in their own words and in their way.
- Interviewers should not monopolize the time discussing their thoughts or personal stories; instead, they should ask questions as succinctly as possible and let the participants have plenty of time to respond.

After the interview guide is constructed, the interviewer is still not ready to begin conducting interviews. The researcher next has to decide how to collect and maintain the information that participants provide.

Commonly, qualitative interviewers take audio recordings of the interviews they conduct. Recording interviews allows researchers to focus on their interaction with the interview participant rather than being distracted by taking notes. Of course, not all participants will feel comfortable being recorded, and sometimes, even the interviewer may feel that the subject is so sensitive that recording would be inappropriate. If this is the case, it is up to the researcher to balance note-taking with listening.

Practicing the interview in advance is crucial. Ideally, researchers should interview one or two peers, or even friends, who are willing to participate in trial runs. Even better are a few people who are similar in at least some ways to the sample. The trial runs can provide feedback on the questions and the demeanor of the interviewer.

Analysis of Qualitative Interview Data

Analysis of qualitative interview data typically begins with a set of transcripts of the interviews. Ideally, researchers who recorded the interview can have the recordings transcribed, so a written verbatim record is available. Interviewers who relied on notes taken during the interview should write a full version of the notes as quickly as possible after the interview while the session is still fresh in mind. It is usually also helpful to note non-verbal items such as body language, tone of voice, or unusually long pauses before an answer.

While third-party transcribers are easily found, it may be best for the interviewer to transcribe the recordings personally. Often, things can be recalled and noted about nonverbal behaviors and interactions relevant to the analysis, but that could not be picked up by the audio recording alone. For example, interviewees may roll their eyes, wipe tears from their faces, and even make obscene gestures that speak volumes about their feelings. However, those indicators would have been lost if the interviewer had not transcribed the recording personally.

The analysis aims to reach inferences, lessons, or conclusions by condensing large amounts of data into more manageable bits of understandable information. Analysis of qualitative interview data usually is inductive research and moves from the specific observations an interviewer collects to identifying patterns across those observations. Qualitative interviewers typically begin by reading through transcripts of their interviews and identifying codes, a shorthand representation of some complex set of issues or ideas. This phase of the research is often referred to as “coding.” It involves reading and rereading (and re-reading again) all of the interview transcripts until the researcher has a clear idea about what sorts of themes come up across the interviews.

Qualitative researcher and textbook author Kristin Esterberg (Esterberg, 2002) describes coding as a multistage process. She suggests that there are two types of coding: open coding and focused coding. To analyze qualitative interview data, researchers can begin by open coding transcripts. They read through each transcript, line by line, and make a note of whatever categories or themes emerge. At this stage, it is essential that they not let the original research question or expectations about what they think they may find cloud their ability to see new categories or themes. This process is called open coding for a reason; they must keep an open mind. Open coding usually requires multiple passes. As researchers read through the transcripts, they begin to see commonalities across the categories or themes. Then, they can begin focused coding.

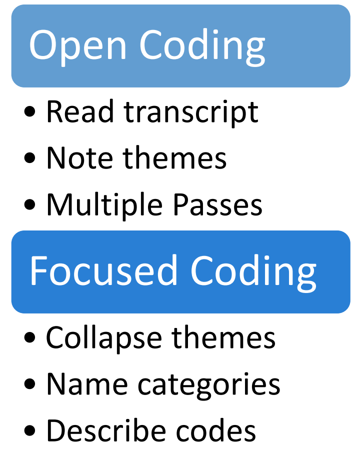


Figure 66: Open And Focused Coding

Focused coding involves collapsing or narrowing themes and categories identified in open coding by reading through the notes made while conducting open coding. Researchers identify themes or categories that seem to be related, perhaps merging some or redefining others. Then they give each theme or category a name or code. Then, they identify passages of data representing the emerging codes by reading through the transcripts yet again (and probably again). They also might write up brief definitions or descriptions of each code to interpret the data and develop a way to discuss the findings.

As tedious and laborious as it might seem to read through hundreds of pages of transcripts multiple times, sometimes getting started with the coding process is the hardest part. In their text on analyzing qualitative data, Lofland (Lofland, 1995) identifies a set of questions helpful in the coding process.

1. Of what topic, unit, or aspect is this an instance?
2. What question about a topic does this item of data suggest?
3. What answer to a question does this item of data suggest (i.e., what proposition is suggested)?

Qualitative data can be analyzed with tools like *NVivo*, *RQDA*, and *QDA Miner Lite*[29]. *NVivo* is very powerful but expensive. *RQDA* is an *R* package that is useful for qualitative data analysis. Since it is part

[29] NVivo information can be found at http://www.qsrinternational.com, RQDA at http://rqda.r-forge.r-project.org/, and QDA Miner Lite at https://provalisresearch.com/products/qualitative-data-analysis-software/freeware/

of the *R* system, it could be easily used in a mixed methods project where *R* is used for quantitative and *RQDA* is used for qualitative analysis. *QDA Miner Lite* is a free Windows software application installed and used on a local computer. These programs are specifically designed to assist qualitative researchers with organizing, managing, sorting, and analyzing large amounts of qualitative data. The programs work by allowing researchers to import interview transcripts and then label or code passages, cut and paste passages, search for different words or phrases, and organize complex interrelationships among passages and codes.

For example, the following excerpt from a paper analyzing the electronic gaming industry in two jurisdictions (Buchanan & Jones, 2010) summarizes how analyzing qualitative interview data often works.

> Data were collected through these combined methods, and while analysis was undertaken using NVivo, the analysis was guided by these methods. Thirty-eight in-depth interviews were undertaken with gaming operators and gaming machine manufacturers in both the Nevada (USA) and NSW (Australian) jurisdictions during 2005−2006. Interview data were augmented through observation, resulting in a rich collection of data. The data were coded and initially entered into 'nodes' within the NVivo program. A pre-defined set of themes was derived from topic areas of the interviews. Each theme then became a node. As each interview was read, additional themes were identified and nodes created for each theme. The nodes were fleshed out as data were extracted from each interview referring to the same theme. Thus a range of themes was created as a result of going through the data and coding according to themes within each transcript. Once all data had been placed into various nodes, themes were checked through the matrix function within NVivo to ensure that the various themes were distinct from each other and that there was no redundancy.

Strengths and Weaknesses of Qualitative Interviews

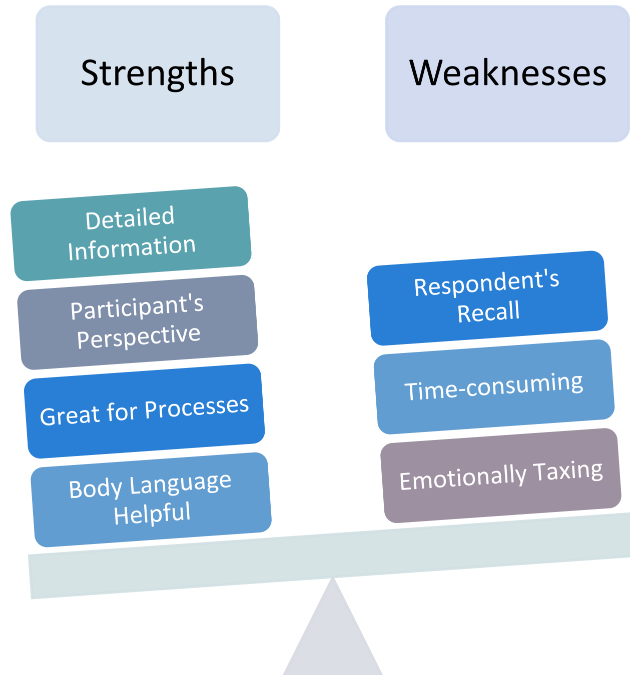


FIGURE 67: STRENGTHS AND WEAKNESSES OF QUALITATIVE INTERVIEWS

As the preceding sections have suggested, qualitative interviews are an excellent way to gather detailed information. Whatever topic is of interest to researchers employing this method can be explored in much more depth than with almost any other method. Participants are allowed to elaborate in a way that is not possible with other methods like survey research. Moreover, they also can share information with researchers in their own words and from their perspectives rather than being asked to fit those perspectives into limited response options provided by the researcher. Because qualitative interviews are designed to elicit detailed information, they are beneficial when a researcher aims to study processes; the "how" of various phenomena. However, another sometimes overlooked advantage of qualitative interviews is that they occur in person so researchers can make observations beyond those that a respondent is orally reporting. A respondent's body language, and even her or his choice of time and location for the interview, might provide a researcher valuable data.

Of course, all these benefits come with some drawbacks.

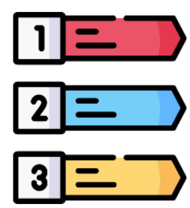
Qualitative interviews rely on respondents' ability to accurately and honestly recall whatever details about their lives, circumstances, thoughts, opinions, or behaviors are being probed. Further, qualitative interviewing is time-intensive and can be pretty expensive. Creating an interview guide, identifying a sample, and conducting interviews are just the beginning. Transcribing interviews is labor-intensive, and that is before coding even begins. It is also common to offer respondents some monetary incentive or thank-you for participating since researchers ask for more participants' time than if they had emailed them a questionnaire. Conducting qualitative interviews is not only labor-intensive but also potentially emotionally taxing. It may be that the researcher will hear stories that are shocking, infuriating, and sad, so researchers embarking on a qualitative interview project should keep in mind their abilities to hear stories that may be difficult to hear.

Quantitative Interview Techniques

Quantitative interviews are similar to qualitative interviews in that they involve some researcher/respondent interaction. However, conducting and analyzing findings from quantitative interviews differs in several ways from that of qualitative interviews. Each approach also comes with its own unique set of strengths and weaknesses.

Conducting Quantitative Interviews

Much of what was covered earlier in this chapter and Chapter 8, Surveys, applies to quantitative interviews. Quantitative interviews are sometimes referred to as "survey interviews" because they resemble survey-style question-and-answer formats. The difference between quantitative interviews and surveys is that questions and answer options are read to respondents in an interview rather than completing a questionnaire independently. As with questionnaires, the questions posed in a standardized interview tend to be closed-ended. There are instances in which a quantitative interviewer might pose a few open-ended questions as well. In these cases, the coding process works somewhat differently than coding in-depth interview data.



In quantitative interviews, an interview schedule leads researchers as they pose questions and answer options to respondents. An interview schedule is usually more rigid than a qualitative interview guide. It contains the list of questions and answer options that the researcher will read to respondents. Whereas qualitative researchers emphasize respondents' roles in helping to determine how an interview progresses, in a quantitative interview, consistency in how questions and answer options are presented is critical. The aim is to pose every question-and-answer option in the very same way to every respondent. This consistency minimizes interviewer effect or possible changes in how an interviewee responds based on how or when the interviewer presents questions and answer options.

Quantitative interviews may be recorded, but because questions tend to be closed-ended, taking notes is less disruptive than during a qualitative interview. If a quantitative interview contains open-ended questions, however, recording the interview is advised. It may also be helpful to record quantitative interviews if a researcher wishes to assess possible interviewer effects. Noticeable differences in responses might be more attributable to the interviewer effect than to any fundamental respondent differences. Having a recording of the interview can help researchers make such determinations.

Quantitative interviewers are usually concerned with gathering data from a large, representative sample but collecting data from many people via interviews can be laborious. Technological advances in telephone interviewing procedures, like computer-assisted telephone interviewing (CATI), can assist quantitative interviewers in this process. These programs allow an interviewer to enter responses directly into a computer as they are provided, thus saving hours that would otherwise have to be spent entering data into an analysis program by hand. Also available are Automated Computer Telephone Interviewing (ACTI) programs where a program generates questions with a realistic voice and then uses voice recognition to record responses. Unfortunately, due to the pervasive increase in "push polling" for election campaigns, many respondents are unwilling to speak to a researcher on the phone.

Analysis of Quantitative Interview Data

As with survey data, the analysis of quantitative interview data usually involves coding response options numerically, entering numeric responses into a data analysis computer program, and then running various statistical processes to identify patterns across responses. Chapter 8 describes the coding process for quantitative data. However, what happens when quantitative interviews ask open-ended questions? In this case, responses are typically numerically coded, just as closed-ended questions are, but the process is a little more complex than simply giving a "no" a label of 0 and a "yes" a label of 1.

In some cases, quantitatively coding open-ended interview questions may work inductively. In this case, rather than ending with descriptions and excerpts, the researcher will assign a numerical value to responses and analyze those data statistically. Keep in mind that with quantitative methods, the aim is to represent and condense data into numbers. Conversely, coding open-ended interview questions can be a deductive process. The researcher may begin with a list of likely responses that have a numerical value assigned. Then researchers will review participants' open-ended responses and designate the numerical value that most closely matches the value of the expected response.

Strengths and Weaknesses of Quantitative Interviews

The strengths and weaknesses of quantitative interviews tend to be compared to those of hard copy questionnaires. For example, response rates tend to be higher with interviews than with mailed questionnaires. That makes sense—most people find it easier to say "no" to a piece of mailed paper than to a person. Quantitative interviews can also help reduce respondent confusion. Suppose respondents are unsure about the meaning of a question or answer option on a questionnaire. In that case, they probably will not have the opportunity to get clarification from the researcher. On the other hand, an interview allows the researcher to clarify or explain any items that may be confusing.

As with every method of data collection, there are also drawbacks to conducting quantitative interviews. Perhaps the largest, and of most concern to quantitative researchers, is the interviewer effect. While questions on hard copy questionnaires may create an impression based on how they are presented, having a person administer questions introduces a slew of additional variables that might influence a respondent. Consistency is critical with quantitative data collection—and human beings are not known for their consistency. Interviewing respondents is also much more time-consuming and expensive than mailing questionnaires. Thus quantitative researchers may opt for written questionnaires over interviews because they will reach a large sample at a much lower cost than interact personally with every respondent.

Issues to Consider

In quantitative and qualitative interviews, researchers interact with their subjects, creating three complexities that deserve attention.

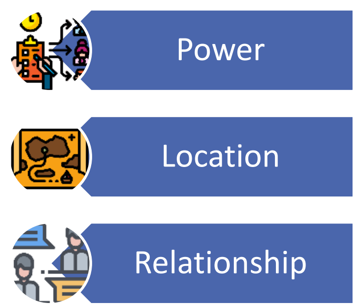


FIGURE 68: INTERVIEWER COMPLEXITIES

Power

First and foremost, interviewers must be aware of and attentive to the power differential between themselves and interview participants. The interviewer sets the agenda and leads the conversation. While qualitative interviewers allow participants some control over which or to what extent various topics are discussed, the researcher is in charge (at least, most respondents will perceive the dynamic). Researchers are asking people to reveal things about themselves that they may not typically share with others. Moreover, researchers are generally not reciprocating by revealing much about themselves. All these factors shape the power dynamics of an interview.

Several excellent articles have been written dealing with issues of power in research and data collection. Anyan (Anyan, 2013) offered several suggestions for overcoming the power imbalance between researchers and participants during the data gathering phase, including the ". . . interviewer must court the interviewee, enhance the sense of rapport between them and build a sympathetic relationship and a sense of mutual trust in the research interview." During data analysis, researchers may want to consider letting interviewees interpret what they meant during the interview. "The willingness to share the data analysis process with the interviewee or letting them join the final stages of writing is in the hands of the interviewer." However, researchers must be vigilant not to let the interviewee shape the outcome of the research project; researchers have an ethical obligation to maintain standards that the average interviewee would not understand.

Another easy way to balance the power differential between researchers and interview participants is to make the intent of the research project very clear. Sharing the rationale for conducting the research and the questions that frame the project may help keep a proper balance of power. Participants should also understand how the data will be stored and used; and how their privacy will be protected. Many of these details are stipulated by the oversight group's procedures and requirements. However, even if they are not, researchers should be attentive to how sharing information with participants can help balance the power differences between themselves and those who participate in the research project.

There are no easy answers for handling the power differential between the researcher and participants, and even professional researchers disagree on the best approach for doing so. It is nevertheless an issue to be attentive to when conducting research, particularly those involving interpersonal interactions and relationships with research participants.

Location

One way to balance the power between researcher and respondent is to interview in the participant's chosen location, where respondents will feel most comfortable answering questions. Interviews can occur in many locations—in respondents' homes or offices, researchers' homes or offices, coffee shops, restaurants, public parks, or hotel lobbies, to name just a few possibilities. Each location comes with its own set of benefits and challenges. It may be best to allow the participant to choose the most convenient and most comfortable location. It is also essential to find a location where there will be few distractions. For example, some coffee shops and restaurants are so loud that recording the interview can be challenging. Other locations may present different sorts of distractions. For example, it may be that parents will spend more time attending to their children during an interview than responding to questions. Interviewers should be prepared to suggest a few possible locations and note avoiding distractions when asking participants to choose a location.

A location's accessibility and safety must balance the respondent's control in choosing. It is conceivable, for example, that a participant's home could be decorated wall to wall with posters representing violently racist messages. Even if the interview topic has nothing to do with home decor, the interviewer's discomfort may distract from the task at hand. While it is essential to conduct interviews in a comfortable location for respondents, doing so should never come at the expense of the interviewer's welfare or safety.

Researcher-Respondent Relationship

Finally, a unique feature of interviews is that they require social interaction, which means that, to some extent, a relationship is formed between interviewer and interviewee. While there may be some differences in how the researcher-respondent relationship works depending on whether the interviews are qualitative or quantitative, one essential relationship element is the same: respect. A good rapport between the interviewer and the participant is crucial to successful interviewing. Rapport is the sense of connection established between the interviewer and participant. Some argue that this term is too clinical, and perhaps it implies that a researcher tricks a participant into thinking they are socially closer than they are. While it is unfortunately true that some researchers might believe this implication, that is not the sense for rapport that researchers should attempt to establish with their subjects. Instead, as already mentioned, the key is respect.

There are no big secrets or tricks for how to show respect for research participants. At its core, the interview interaction should not differ from any other social interaction in which interviewers show gratitude for a person's time and respect for a person's humanity. Interviewers must interview in a culturally sensitive way. In some cases, this might mean educating themselves about the study population and even receiving training to help them learn to communicate with the research participants. Interviewers should not judge participants; they are there to listen. Participants have been kind enough to give them their time and attention. Even if interviewers disagree strongly with what a participant shares in an interview, their job as researchers is to gather the information being shared, not to make personal judgments about it.

Developing good rapport requires good listening. Listening during an interview is an active, not a passive, practice. Active listening means that interviewers participate with the respondent by showing that they understand and follow whatever is being shared. The questions asked to respondents should indicate that interviewers heard what they have just said. Active listening probably means that interviewers will probe the respondent for more information from time to time throughout the interview. A probe is a request for more information. Both qualitative and quantitative interviewers probe respondents, though the way they probe usually differs. In quantitative interviews, probing should be uniform. Often quantitative interviewers will predetermine what sorts of probes they will use. Interviewers should not use probes that might appear to agree or disagree with what respondents said. So "yes" or "I agree" or even a questioning "hmmm" should be avoided. Instead, responses like a simple "thank you" to indicate that the response was heard are more neutral. A "yes" or "no" response should be used if, and only if, a respondent explicitly asks if they were heard or understood.

In some ways, qualitative interviews better lend themselves to following up with respondents and asking them to explain, describe, or otherwise provide more information. Qualitative interviewing techniques are designed to go with the flow and take whatever direction the respondent goes during the interview. Nevertheless, it is worth the interviewer's time to come up with helpful probes in advance of an interview, even in the case of a qualitative interview. They do not want to find themselves stumped or speechless after a respondent has just said something that needs to be probed further. Preparing for unexpected responses is another reason that practicing an interview in advance with people similar to those in the sample is a good idea.

Conducting the Interview

Before the interview, interviewers should prepare a "kit" to carry to the interview session, including a cover letter from the principal investigator or sponsor, several copies of the survey instrument, photo identification, and a telephone number for respondents to verify the interviewer's authenticity. The interviewer should set up an appointment and then be on time. To start the interview, interviewers should speak in an imperative and confident tone, such as "I would like to take a few minutes of your time to interview you for a critical study," instead of "May I come in to do an interview?" They should introduce themselves, present personal credentials, explain the purpose of the study in a few sentences, and assure confidentiality of respondents' comments, all in less than a minute. No big words or jargon should be used, and no details should be provided unless specifically requested. If interviewers wish to record the interview, they should ask for the respondent's explicit permission before starting. Even if the interview is recorded, the interview must take notes on crucial issues, probes, or verbatim phrases.

During quantitative interviews, interviewers should follow the script and ask questions exactly as written without changing the order of questions or skipping questions that may have been answered earlier. It is not uncommon for the participant to answer a question that will be asked later in the interview. In this case, the interviewer should ask the question every time it comes up and record the response as if it was never answered before. Later, the interviewer can use both answers to represent the interview's outcome better. Interviewers can be more expansive during qualitative interviews and guide the discussion along a different path than anticipated. Any issues with the questions should be worked out during rehearsal before the actual interview sessions. In any case, interviewers should not finish the respondent's sentences. If the respondent gives a brief cursory answer, the interviewer should probe further to elicit a more thoughtful, thorough response. Some practical probing techniques are:

- **The silent probe**. Just pausing and waiting (without going into the next question) may suggest that the interviewer is waiting for a more detailed response.
- **Overt encouragement**. Occasional "uh-huh" or "okay" may encourage the respondent to go into more detail. However, the interviewer must not express approval or disapproval of what was said by the respondent.
- **Ask for elaboration**. Phrases such as "can you elaborate on that" or "tell me more about what you think" will encourage a better response.
- **Reflection**. Interviewers can try the psychotherapist's trick of repeating what the respondent said. For instance, "What I hear is that you found that experience very traumatic," and then pause and wait for the respondent to elaborate.

After the interview is completed, interviewers should thank respondents for their time, tell them when to expect the results, and not leave hastily. Immediately after leaving, they should write down any notes or key observations that may help interpret the respondent's comments.

Focus Groups

Focus groups resemble qualitative interviews in that a researcher may prepare an interview guide in advance and interact with participants by asking them questions. However, anyone who has conducted both one-on-one interviews and focus groups knows that each is unique. In an interview, one member (the participant) is usually the most active while the other (the researcher) is the listener, conversation guide, and question asker. On the other hand, focus groups are planned discussions designed to elicit group interaction and "collects data through group interaction on a topic determined by the researcher." (Morgan, 1996) In this case, the researcher may play a less active role than in a one-on-one interview. The researcher aims to get participants talking to each other and to observe interactions among participants.

Focus groups are typically more dynamic than interviews. The researcher takes the role of moderator, posing questions or topics for discussion, but then lets the group members discuss the question or topic among themselves. Participants may ask each other follow-up questions, agree or disagree with one another, display body language that indicates something about their feelings, or even pose questions not previously conceived by the researcher. It is just these sorts of interactions and displays that are of interest to the researcher. A researcher conducting focus groups collects data on more than people's direct responses to the questions; the group interaction is a critical focal point. Due to the nature and unpredictability of group interaction and the fact that focus group researchers generally want to enhance group interaction, focus groups tend to be qualitative rather than quantitative.

There are numerous examples of marketing and business research using focus group methodology.

Focus groups were used in a 2011 study of the use of visual tobacco warnings. The researchers determined that the European Union graphic warnings on cigarette packages were more effective than text warnings (Gallopel-Morvan, Gabriel, Le Gall-Ely, Rieunier, & Urien, 2011). They used six focus group discussions conducted in Rennes, Paris, and Brest with 50 people (26 smokers, 24 non-smokers, 25 women, 25 men).

An interesting study published by Wutich et al.in 2010 compared the results of a focus group with an open-ended self-administered questionnaire among water management decision-makers in Phoenix, Arizona (Wutich, Lant, White, Larson, & Gartin, 2010). She found that the focus group was no better than the questionnaires when the questions were only moderately sensitive. However, the focus group was better ". . . for very sensitive topics when there appeared to be an opportunity to exchange important information or solve a pressing problem."

In 2013, Sylvetsky et al. published a focus group study where the effectiveness of advertising for the development of an obesity awareness campaign aimed at young people (Sylvetsky, et al., 2013). She conducted ten focus group discussions in two regions of Georgia, United States. The groups of children, aged 9−14, were led in discussions concerning healthy food choices and lifestyles. Researchers identified three themes.

- "My Mom wants me to eat healthy foods like broccoli, but it looks nasty and tastes gross."
- "Obesity is a problem, but it does not apply to me."
- "Everyone is made differently, and it does not matter if you are fat."

Government officials and political campaign workers use focus groups to learn how the public feels about a particular issue or candidate. One of the earliest documented uses of focus groups came from World War II when researchers used them to assess troop training materials' effectiveness and various propaganda efforts (Merton & Kendall, 1946). Market researchers quickly adopted this method of collecting data to learn about human beliefs and behaviors. Within social science, the use of focus groups did not take off until the 1980s, when demographers and communication researchers began to appreciate their use in understanding knowledge, attitudes, and communication. Beyond various applied research projects, social scientists also use focus groups in theory development projects like those mentioned above.

Focus groups share many of the strengths and weaknesses of one-on-one qualitative interviews. Both methods can yield very detailed, in-depth information, are excellent for studying social and business processes, and provide researchers with an opportunity to hear what participants say and observe what they do in terms of their body language. Focus groups offer the added benefit of giving researchers a chance to collect data on human interaction by observing how group participants respond and react to one another. Like one-on-one qualitative interviews, focus groups can also be quite expensive and time-consuming. However, there may be some time savings with focus groups as it takes fewer group events than one-on-one interviews to gather data from the same number of people. Another potential drawback of focus groups is that one or two participants might dominate the group, silencing other participants. Careful planning and skillful moderation on the researcher's part are crucial for avoiding or dealing with such possibilities. The various strengths and weaknesses of focus group research are summarized below.

Strengths

- Yield detailed, in-depth data
- Less time-consuming then one-on-one interviews
- Useful for studying social processes
- Allows researchers to observe body language
- Allows researchers to observe group interactions

Weaknesses

- Expensive
- More time-consuming than a survey
- Some participants may dominate the group

FIGURE 69: FOCUS GROUP STRENGTHS AND WEAKNESSES

As mentioned, careful planning and skillful moderating are two crucial considerations in using focus groups as data collection methods. In some ways, focus groups require more advanced planning than other qualitative data collection methods such as one-on-one interviews or field research. Researchers must take care to form focus groups whose members want to interact with one another and control the event's timing so that participants are not asked nor expected to stay for a longer time than they have agreed to participate. The researcher should also be prepared to inform focus group participants of their responsibility to maintain confidentiality. Nevertheless, while researchers can and should encourage all focus group members to maintain confidentiality, they should also clarify to participants that the unique nature of the group setting prevents them from being able to promise that confidentiality will be maintained.

Group size should be determined in part by the topic of the interview and the researcher's sense of the likelihood that participants will have much to say without much prompting. If the topic is something the participants feel passionately about and will have much to say, a group of three to five is ideal. Groups more extensive than that, especially for heated topics, can quickly become unmanageable. Some recommend that a group of about six to ten participants is ideal for focus group research, while others recommend that groups include anywhere from three to twelve participants. The size of the focus group is ultimately the researcher's decision. When forming groups and deciding how large or small to make them, researchers must consider what they know about the topic and participants' potential interest in, passion for, and feelings about the topic. They must also consider their comfort level and experience in conducting focus groups.

It may seem counter-intuitive, but in general, it is better to form focus groups consisting of participants who do not know one another than to create groups consisting of friends, relatives, or acquaintances. Group members who do not know each other may share some taken-for-granted knowledge or assumptions. In business research, it is precisely the taken-for-granted that is often of interest; thus, the focus group researcher should avoid setting up interactions where participants may be discouraged from questioning or raising issues they take for granted. However, groups should not be so heterogeneous that participants will be unlikely to feel comfortable talking with one another.

Focus group researchers must carefully consider the composition of the groups they put together. Morgan suggests that "homogeneity in background and not homogeneity in attitudes" should be the goal (Morgan, 1996). Participants must feel comfortable speaking up and have enough differences to facilitate a productive discussion. Whatever composition researchers design for their focus groups, the vital point to keep in mind is that multiple social contexts shape focus group dynamics. Participants' silences and their speech may be shaped by gender, race, class, sexuality, age, or other background characteristics or social dynamics, all of which might be suppressed or exacerbated depending on the group's composition.

In addition to the importance of planning, focus groups also require skillful moderation. While a researcher certainly does not want to be viewed as a stick-in-the-mud or overly domineering, it is important to set ground rules for focus groups at the outset of the discussion. Participants should be reminded that they were invited to participate because the researcher wants to hear from all of them. Therefore, the moderator should not let more than one person speak simultaneously and avoid letting just a couple of participants dominate the conversation. One way to do this is to begin the discussion by asking participants to briefly introduce themselves or provide a brief response to an opening question. This introduction will help set the tone of having all group members participate. Also, ask participants to avoid having side conversations; sharing thoughts about or reactions to what is said in the group is essential and should not be limited to only a few group members.

As the focus group gets rolling, the moderator will play a less active role than in a one-on-one interview. There may be times when the conversation stagnates, or the moderator wishes to guide the conversation in another direction. In these instances, moderators need to demonstrate that they have been paying attention to what participants have said. Being prepared to interject statements or questions such as "I would like to hear more about what Sally and Joe think about what Dominick and Ashley have been saying" or "Several of you have mentioned. . . What do others think about this?". It can also help redirect the conversation, shift the focus to participants who have been less active in the group, and serve as a cue to those dominating the conversation that it is time to allow others to speak.

In sum, focus groups are a valuable method for researchers who wish to gather in-depth information about social processes. Focus groups are similar to one-on-one qualitative interviews in many ways. However, they allow researchers to observe group dynamics that cannot be observed in one-on-one interviews. Historically, focus group research was more commonly used by applied researchers than by academics. However, in recent decades social scientists from all domains have discovered the usefulness of focus groups to understand social processes and have begun using this data collection method in their studies.

Key Takeaways

Chapter 10: Interviews
• Discuss the role of the interviewer • Conducting and analyzing a qualitative interview • Conducting and analyzing a quantitative interview • Discuss the strengths and weaknesses for both qualitative and quantitative interviews • Discuss using focus groups in a research project

.

11: Field Research

Introduction

If researchers wanted to know who conducts more of the housework in households, how could they find the answer? One way might be to interview people and ask them. That is exactly what Arlie Hochschild did in her study of the second shift, her term for the work that goes on in the home after the day's work for pay is completed (Hochschild & Machung, 2012). Hochschild interviewed 50 heterosexual, married couples with children to learn about how they did or did not share the second shift work. Many couples reported that they shared the load of the second shift equally, sometimes dividing the house into areas that were "hers" and those that were "his." However, Hochschild was not satisfied with just people's reports of second-shift work. She chose to observe 12 of these couples in their homes and see for herself just how the second shift was shared.

Hochschild discovered that even couples who claimed to share the second shift did not have as equitable a division of duties as they professed. For example, one couple who told Hochschild during their interview that they shared the household work equally had explained that the wife was responsible for the upstairs portion of the house. The husband took responsibility for the downstairs portion. Upon conducting observations in this couple's home, however, Hochschild discovered that the upstairs portion of the house contained all the bedrooms, bathrooms, kitchen, dining room, and living room. At the same time, the downstairs included a storage space and the garage. This division of labor meant that the woman carried the weight of responsibility for the second shift. Without a field research component to her study, Hochschild might never have uncovered these and other truths about couples' behaviors and sharing household duties.

OBJECTIVES
• Define "Field Research" • Describe the strengths and weaknesses of field research • Describe how to get started with field research: choosing a site and role • Describe how to write field notes and then analyze those notes

What is Field Research?

Field research is a qualitative method of data collection to understand, observe, and interact with people in their natural settings. Thus when researchers talk about being in "the field," they mean being out in the real world and involved in the everyday lives of the people they are studying. Field research and participant observation are commonly used terms for this type of project. Field research is an umbrella term that includes the myriad activities that field researchers engage in when collecting data: participate, observe, interview some of the observed people, and analyze documents or artifacts created by the people they observe.

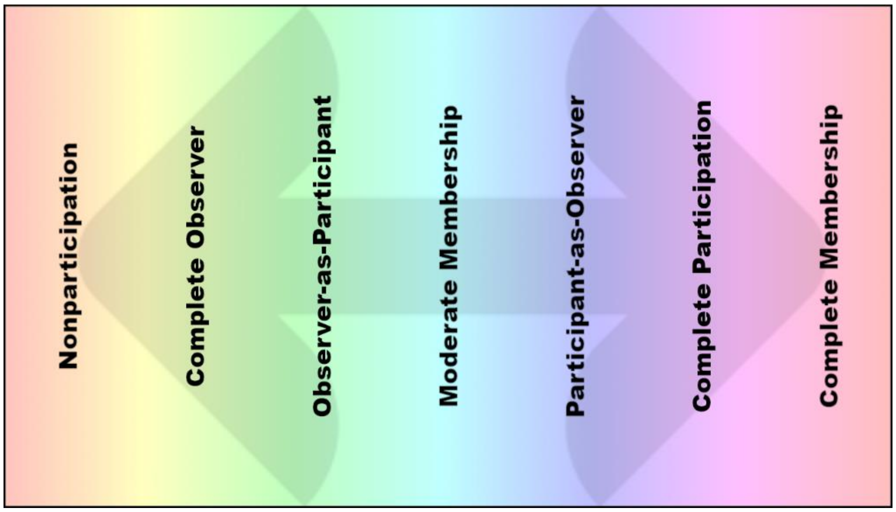


Figure 70: Participant Observation Levels

This chapter focuses only on field research participation and observation because interviews (Chapter 10) and document analysis (Chapter 12) are covered elsewhere. These aspects of field research are usually considered together and are known as participant observation. Like field research, participant observation also has multiple meanings. Researchers conducting participant observation vary in the extent to which they participate or observe (Baker, 2006). While many "participation scales" have been developed, Baker proposes a continuum where "Nonparticipation" lies at one end and "complete membership" lies at the other, as illustrated in Figure 70.

- **Nonparticipation**. Researchers using this method have no involvement with the group being studied. Researchers are not physically present but can observe using an entirely different environment. As an example of this level of involvement, imagine a researcher watching some sort of group interaction from another room on a closed-circuit television system.
- **Complete observer.** Researchers using this method are physically present with the insiders being observed but have no interaction or, at most, superficial interaction. Researchers are only present to listen and observe.
- **Observer-as-participant**. Researchers using this method are engaged in more observation than participation but still interact, maybe brief interviews, with the insiders. Researchers do not

become friends with the insiders and would not "get a beer after work" but would feel comfortable asking them why they were doing some task in a particular manner.

- **Moderate membership**. Researchers using this method attempt to maintain a balance between being an insider and pure observation. They would participate in certain activities but not those that are at the core of insider membership. As an example, researchers observing drug dealers may "hang out" and listen to music with them but would not engage in any sort of illegal activity.
- **Participant-as-observer**. Researchers using this method become more involved with insiders' main activities but still do not fully commit to the members' values and goals. Researchers may develop friendships with the insiders and even participate in social activities, like going to dinner together.
- **Complete participation**. Researchers using this method are said to "go native" with the insiders. They become part of the insiders' group and share all of the goals and norms of the group being studied. This level of involvement can be problematic since researchers may become so ingrained in the group being observed that they can no longer offer unbiased observations. For this reason, most research experts warn that "going native" should be avoided.
- **Complete membership**. Researchers using this method have entirely "gone native" and are part of the group being observed. The main difference between this level and the previous level is that researchers who attain full membership do so intentionally and have no hesitation in being part of the group being observed.

As it might have been imagined based on the examples of the observational roles assumed, field research is well equipped to answer "how" kinds of questions. Whereas survey researchers often aim to answer "why" questions, field researchers ask how processes occur, how those observed in the field interact, and how events unfold.

Field research is a method that anthropologists initially crafted for cultural understanding and interpretation (Wolcott, 1999). Dissatisfied with studying groups of people based solely on secondhand accounts and inspection of artifacts, several anthropologists decided to try living in or near the communities they studied to learn from and about them. Two anthropologists, in particular, Franz Boas (Boas, 1964) and Bronislaw Malinowski (Malinowski, 2014), are credited with developing this method around the turn of the 20th century. Boas lived with native populations in Canada and the American Northwest. Malinowski lived in Papua New Guinea with people who were native to the area. Sociologists picked up on the idea and the benefits of field research. Soon, several sociologists had embraced this new method and adapted field research for their studies of groups. Many of the early field researchers in sociology were former social workers interested in sociological research because they hoped to be social reformers.

Strengths and Weaknesses of Field Research

The following summarizes the strengths and weaknesses of field research covered in this section.

Strengths

- Yields very detailed data.
- Emphasizes the role and relevance of social context.
- Can uncover social facts that may not be immediately obvious or of which research participants may be unaware.

Weaknesses

- May lack breadth; gathering very detailed information means being unable to gather data from a vast number of people or groups.
- May be emotionally taxing.
- Documenting observations may be challenging.

FIGURE 71: FIELD RESEARCH STRENGTHS AND WEAKNESSES

Strengths of Field Research

Field research allows researchers to gain firsthand experience and knowledge about the people, events, and processes they study. No other method offers quite the same kind of closeup lens on everyday life. This close-up on everyday life means that field researchers can obtain detailed data about people and processes, perhaps more detailed than they can obtain using any other method.

Field research is an excellent method for understanding the role of social context in shaping people's lives and experiences. It enables a greater understanding of the intricacies and complexities of daily life. Field research may also uncover elements of people's experiences or group interactions previously unknown. This benefit, in particular, is a unique strength of field research. With other methods, such as interviews and surveys, respondents cannot answer questions outside their experience or provide unfamiliar information. Also, because field research typically occurs over an extended period, social facts that may not be immediately evident to a researcher can be uncovered during a field research project.

Weaknesses of Field Research

The rich data collected during field research comes at a cost. Because a field researcher's focus is detail-oriented, it is by necessity also somewhat narrow. Field researchers cannot gather data from as many individuals as a survey, for example. Indeed, field researchers generally sacrifice breadth in exchange for depth. Related to this point is the fact that field research is highly time-intensive.

Field research can also be emotionally taxing. To a certain extent, interview research requires developing a relationship between a researcher and the participants, but field research requires a much more significant investment in the researcher's life. It may be said that interviews are like casual dating, while field research is like a marriage.

The relationships developed as a field researcher are sustained over a much more extended period than the hour or two it might take to conduct an interview. Not only do the relationships last longer, but they are also more intimate. Several field researchers have documented the complexities of relationships with research participants (Taylor, 2011), (Sanjari, Bahramnezhad, Fomani, Shoghi, & Cheraghi, 2014), (Greene, 2014). On the plus side, these relationships can be enriching (and yield the rich, detailed data noted as a strength in the preceding discussion). Nevertheless, field researchers experience the highs and the lows of daily life and interactions as in any relationship. Moreover, participating in day-to-day life with one's research subjects can result in some tricky ethical quandaries. It can be a challenge if the goal is to observe as "objectively" as possible.

Finally, documentation can be challenging for field researchers. Where survey researchers have the completed questionnaires and interviewers have recordings, field researchers generally have only the documents of their observations. This challenge becomes immediately apparent upon entering the field. It may not be possible to take field notes as they observe, nor will they necessarily know which details to document or which will become the essential details to have noted. Finally, the notes taken after some observation may be incomplete since researchers may not recall everything exactly.

Getting In

When embarking on a field research project, researchers must consider two major decisions: where to observe and what role to take at the field site. The decision about each of these will be shaped by several factors, some of which researchers will control and others they will not. The decisions about where to observe and what role to play will also affect the data they can gather and analyze.

Choosing a Site

The research location may be determined somewhat by the research question, but investigators may not have a well-developed question before they begin observations because this is usually inductive research. In some cases, field researchers do not define a research question until they determine where the data are taking them. Other times, they begin with a research question but remain open to the possibility that their focus may shift as they gather data. In either case, when a site is chosen, many factors must be considered. What do they hope to accomplish with the field research? What is their topical/substantive interest? Where are they likely to observe behavior that has something to do with that topic? How likely is it that they will have access to the locations that are of interest? How much time do they have to conduct participant observations? Are the observations limited to a single location or multiple?

Perhaps the best place to start as researchers identifies a site or sites for their field research is to think about their limitations. One limitation that could shape participant observation is time. Field researchers typically immerse themselves in their research sites for many months, sometimes even years (Davies & Crane, 2010), (Jack, Moult, Anderson, & Dodd, 2010). Researchers must ask themselves if they have several years to conduct research or seek a smaller-scale field research experience? How much time do they have to participate and observe per day? Per week? Identifying the time available helps them determine where and what sort of research sites to choose.

Researchers must also think about where they live and whether travel is an option. Some field researchers move to live with or near their population of interest, but that may not be an option in most cases. Professor Erik Larson's research on variations in economic institutions in a global environment, for example, has taken him across the globe, from Fiji to Ghana to Iceland (Larson, 2010). Sociologist Sara Dorow's research on transnational adoption took her from the United States to China (Dorow, 2006). These are just two of many examples of researchers who have traveled the globe to collect data.

Researchers must also consider how the social location might limit what or where they can study while choosing a site. The "ascribed" aspects of the researcher are involuntary, such as age, race, or mobility. How might the ascribed status of a middle-aged man, for example, shape the researcher's ability to conduct full participation in a study of children's birthday parties? In contrast, the "achieved" aspects of the researcher are under some control, like education level and wealth. For field research, researchers may also choose whether or the extent to which they reveal the achieved aspects of their identities. There are numerous examples of field researchers whose achieved statuses granted them access to field sites into which they might not have otherwise been allowed. For example, a licensed paralegal may gain access to law offices that would not be possible for other people.

The preceding discussion should not be taken to mean that researchers cannot, should not, or do not study those from whom they differ. There have been plenty of successful field studies conducted by researchers who may have looked out of place in the sites they chose to investigate. Teresa Gowan, a self-described “small, white English woman,” conducted field research with homeless men in some of San Francisco’s most notoriously rough neighborhoods (Gowan, 2010). The aim here is not to affirm the socially constructed categories which society emphasizes. Instead, the point is to be aware of which ascribed and achieved aspects of the researcher’s identity may shape decisions about field sites.

Finally, researchers must consider whether the research will be a collaborative or individual project in choosing a research site. Collaborating with others has many benefits, such as covering more ground and collecting more data than working independently. Also, having collaborators in field research projects means having others with whom researchers can share their trials and tribulations in the field. However, collaborative research comes with its own sets of challenges, such as possible personality conflicts among researchers, competing commitments in time and contributions to the project, and differences in methodological or theoretical perspectives. When considering something that is of interest, researchers should consider whether they have possible collaborators and how they could shape the decisions about conducting participant observation.

While this section began by considering the limitations that might shape field site decisions, it is also true to remember the opportunities—social, geographic, and otherwise—that location affords. Perhaps researchers are already members of an organization where they would like to conduct research. Maybe they “know someone who knows someone” who may be able to help access a site. Perhaps they have friends they could stay with, so they could observe participants away from home.

Choosing a site for participation is shaped by all of these factors: the research question and area of interest, a few limitations, some opportunities, and sometimes a bit of being in the right place at the right time.

Choosing a Role

As with choosing a research site, some limitations and opportunities beyond researchers' control might shape the role they take once they begin participant observation. Researchers need to make some deliberate decisions about entering the field and "who" they will be once they are there.

One of the earliest decisions researchers need to make is whether to be overt or covert in the field. As overt researchers, they enter the field with research participants aware that they are the research project subjects. On the other hand, covert researchers enter the field as though they are participants, opting not to reveal that they are also researchers or that the group they have joined is being studied. As it may be imagined, there are strengths and weaknesses to both approaches. A critical point to keep in mind is that whatever decision is made about how they enter the field will affect nearly all subsequent experiences.

Overt researchers may experience some trouble establishing rapport at first. Having an insider at the site who can vouch for the researcher will undoubtedly help. However, the knowledge that subjects are being "watched" will inevitably (and understandably) make some people uncomfortable and possibly cause them to behave differently than they would were they not aware of being research subjects. Because field research is typically a sustained activity over several months or years, participants will likely become more comfortable with the researcher's presence over time. Overt researchers also avoid a variety of moral and ethical dilemmas that they might otherwise face.[30]

Covert researchers can "get in" the site easier but then face other issues. For how long should they conceal their identities? How might participants respond once they discover they have been studied? How will researchers respond if asked to engage in activities they find unsettling, unsafe, or even unethical? Field researcher Richard Mitchell was forced to consider these very questions during his covert research among right-wing survivalists. He was asked to participate in the swapping of violently racist and homophobic stories, an experience over which he later expressed profound grief and deep regret (reported by W. Shaffir and RA Stebbins (Shaffir & Stebbins, 1991)). Beyond their comfort level with deceiving participants and willingness to take risks, deciding whether to enter the field covertly may be made for researchers. Suppose they are researching while associated with any federally funded agency (and even many private entities). In that case, the Institutional Review Board (IRB) probably will have something to say about any planned deception of research subjects. Some IRBs approve deception,

[30] Students interested in this aspect of field research may want to investigate the Hawthorne effect.

but others look warily upon a field researcher engaging in covert participation. The extent to which the research site is a public location, where people may not expect privacy, might also play a role in helping researchers decide whether covert research is a reasonable approach.

Insiders, with whom a researcher may have some prior connection or a closer relationship than with other site participants, are called “key informants,” and they can provide a framework for observations, help “translate” what is observed, and provide important insight into a group’s culture. If possible, having more than one key informant at a site is ideal, as one informant’s perspective may vary from another’s.

Once a decision is made about entering a field site, researchers need to think about the role they will adopt while there. Aside from being overt or covert, they need to determine how close they will be to participants. In the words of Fred Davis, who coined these terms about researchers’ roles, will the researcher be a Martian, a Convert, or a bit of both (Davis F. , 1973)? Davis describes the Martian role as one in which a field researcher stands back a bit, not fully immersed in his subjects' lives, to problematize better, categorize, and see with the eyes of a newcomer what is being observed. From the Martian perspective, a researcher should remain disentangled from too much engagement with participants. The Convert, on the other hand, intentionally dives right into life as a participant. From this perspective, it is through total immersion that understanding is gained.

While Davis’ definition of researcher roles is simple and easy to understand, “Participant Observation Levels” were more thoroughly defined along a continuum from “Nonparticipation” to “Complete Membership” earlier in this chapter. Those planning to engage in field research should carefully evaluate the roles and levels of observation before starting the study.

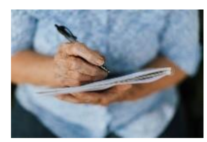
Many of the points made about power and relationships for interviews (Chapter 10) apply to field research as well. The researcher/researched relationship is even more complex in field studies, where interactions with participants last far longer than the hour or two it might take to interview someone. Moreover, the potential for exploitation on the researcher's part is even more significant in field studies as relationships are usually closer, and lines between “research” and personal or off-the-record interaction may get blurred. These precautions should be seriously considered before deciding to embark upon a field research project.

Field Notes

Field notes are an opportunity for a researcher to write poorly and get away with it. While that is said in jest, it does contain a grain of truth. Field notes are a type of writing where researchers should not be going for literary value, making the writing engaging, or even making it readable for anyone other than the researcher. Instead, the aim is to record observations as accurately and quickly as possible. Field notes are the first and necessary step toward developing qualitative analysis. They are also the record that affirms what was observed. In other words, field notes are not to be taken lightly or overlooked as unimportant.

Some say that there are two different kinds of field notes: descriptive and analytic. Though the lines between what counts as "description" and what counts as "analysis" can get fuzzy, the distinction is useful when thinking about writing and interpreting field notes. This section focuses on descriptive field notes, which describe a field researcher's observations as straightforwardly as possible. These notes typically do not contain explanations or comments about those observations; instead, the observations are presented on their own, as clearly as possible. The following section considers the analysis of field notes.

Writing in the Field



Field researchers use a variety of strategies to take notes while in the field. Some research is conducted where sitting with a notebook or tablet is no problem (e. g., observing at a meeting), but this is the exception. More often, field researchers must find creative ways to note their observations while engaged in the field. There are stories about field researchers jotting notes on their hands, keeping tiny notebooks in their pockets, carrying small recorders to make quick observations, and even writing notes on toilet paper during visits to the restroom. With the advent of smartphones, taking notes in the field has become less arduous since it is common to see someone texting or surfing the web from a phone in almost any setting. Figure 72 is an example page from a field notebook found at the United States Library of Congress, https://www.loc.gov/folklife/edresources/ed-trainingdocuments.html.

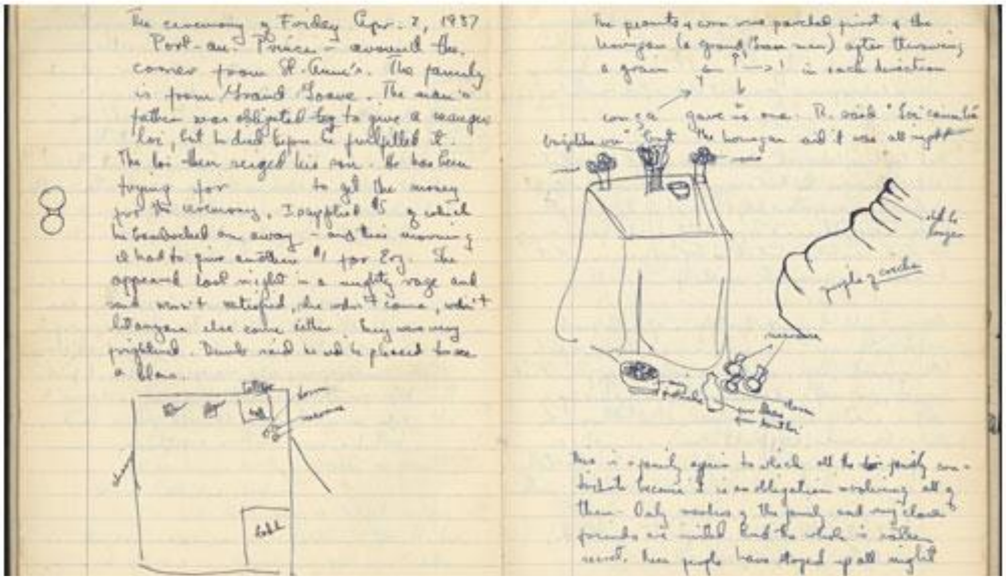


FIGURE 72: EXAMPLE FIELDNOTES

The strategy for recording observations while in the field will be determined chiefly by the chosen site and role. If researchers are in a setting where having a notebook or smartphone in their hands does not look out of place, they should use those tools to take notes. However, they must be careful not to let note-taking distract them from observing what is happening—writing notes while in the field requires a delicate balance between jotting down observations and engaging in the setting. Researchers who are observers will find it easier to balance the note-taking and observation process, but those who are also participants need to be more careful about balance. If researchers happen to be in a location where taking notes “at the moment” would be too obvious, rude, or distracting, they may still be able to jot down a few things very quickly occasionally. They may also need to develop a way of jotting down observations that do not require complete sentences or perhaps even words. Many field researchers develop a version of “shorthand” for notes, using some combination of abbreviations and symbols. Using abbreviations keeps them from taking too much time away from their participation in the field.

As with any other proficiency, writing field notes is a skill that can be improved with practice. Conducting field research and taking field notes are decidedly not informal activities. When engaged in field research, observation is deliberate, not haphazard. That said, for a first-time field researcher, taking field notes can feel like a rather haphazard activity. Understanding when, what, where, and how to write are skills that field researchers develop with experience.

No matter how difficult it can be to write notes while in the field, it is worth the effort. Field researchers rely on the notes they take to develop more complete notes later and eventually develop analysis. There is an old philosophical question: if a tree falls in the woods, but nobody hears it, did it make a sound? While the answer to that question is outside the purview of this book, observations that are not noted may as well have not happened when it comes to field research. Researchers cannot possibly be expected to remember everything they see over the hours, days, months, or years spent collecting data in the field. For this reason, writing notes in the field (to the extent possible) is essential, as is "filling in" those notes as soon as researchers are in a location where they can focus on more formal note-taking.

Here is a summary for writing in the field.

- Must find creative ways to take notes
- Researcher's role shapes the type of field notes recorded
- Writing field notes improves with practice
- Field notes are always worth the effort

FIGURE 73: SUMMARY FOR WRITING IN THE FIELD

Writing Out Of The Field

Immediately upon leaving any observation in the field, researchers should take the time to expand the brief notes taken while in the field. Even if they feel that the notes are complete, researchers can be surprised by how much more they recall once they sit down without distractions and read through their notes. Rereading is also an excellent opportunity to add reflections about the observations.

As they enter their notes into a computer, researchers should "fill in the blanks" and write as much as possible about what was just observed. Even if it seems mundane, it is fair to say that field notes can never contain too much detail. Writing as much as possible also helps researchers avoid generalizing in their field notes. The notes should be specific about observations, so rather than saying that "everyone" talked about something, notes about who was talking, or even that the researcher is unsure precisely who was talking, is helpful. Rather than noting that someone was "angry," it is best to describe how that impression was formed; for example, was that person yelling, red in the face, or shaking a fist?

Researchers must also describe exactly where some activity took place and detail the surroundings (in addition to describing the observed interactions and conversations). Early in a field research project, researchers may focus slightly more on describing the "lay of the land" than later in the project. Early notes might include very detailed descriptions of the locations and the people involved in interactions. It is also reasonably common for researchers to draw a map or, if appropriate, take pictures of the field sites. If observations will be conducted in the same place and with the same people, these descriptive details noted early on will become less noticeable over time, so it is helpful to have some documentation of the researcher's first impression.

For example, the following is an excerpt from Blackstone's first meeting with two key informants in a field research project concerning the breast cancer movement (Blackstone, 2012).

> 1/14/99, 11:00am
>
> Met Jane and Polly at the XX office today. I was scheduled to be there at 10:30 but traffic was so bad due to last night's snow storm that I did not get there until 11:00am. Jane and Polly did not seem bothered by my tardiness (Polly, "We don't keep a time clock around here."). I walked into the building and took the elevator up to the second floor. I was a little unsure about where to go from there so I just walked into the first open door and said, "I'm looking for the XX office." A woman showed me into a large office (long and slightly irregular shape with windows on one wall, a desk and table and many chairs. Also two computers set up on a counter that runs along the wall across from the windows.) Two women were looking at a computer screen that was on the counter. When I walked in I introduced myself and Jane and Polly introduced themselves to me. Both women shook my hand, though Jane was the first to do so and did so with slightly more selfassurance than Polly. Polly told me to hang my coat on one of the "coat racks" and gestured to the many chairs that were around the office. I placed my coat and purse in what I hoped would be the most out of the way location; a corner behind the table.

This excerpt will not win the Pulitzer Prize for its riveting story or prose, but that is not its purpose. Instead, Blackstone's goal was to describe a location and a first impression of the two women who would be likely candidates for key informants. One thing of note is that quotation marks are used every time a person is directly quoted. Including as many direct quotes as possible is a good idea since such quotes support the analytic points made later when describing patterns in the data. Recording direct quotes is another reason that taking notes in the field (to the extent possible) is a good idea. Direct quotes may be challenging to remember hours or even minutes after hearing them. For this reason, researchers may wish to write verbatim quotes while in the field and then take the time to describe the circumstances under which something was said later on when compiling full notes.

Another useful convention is to use punctuation, like all-capital letters or brackets, to distinguish between observations and interpretation. It is not always easy to distinguish between a dispassionate observation and its interpretation, but most researchers attempt to distinguish between these two categories of information in their field notes.

Indeed, the “here is what I thought” portions of a researcher’s field notes may never be used, but those sections can inform the analysis of data. Sometimes, bracketed notes express emotion or complex thoughts or feelings, especially when researchers feel upset or annoyed by something in the field. Because field research requires developing personal relationships with “subjects,” and because interpersonal relationships experience various highs and lows, it is important to express feelings about those relationships in the notes. Writing these more personal reflections may become necessary for analysis later, or they may be cathartic at the moment. They might also reveal biases researchers have about the participants, and honesty is essential about that confounding factor.

Every field researcher’s approach to writing up field notes will vary according to whatever strategy works best for that individual. Where one researcher may use brackets to document personal feelings and reflections on bits of data, others may use the “comments” function in a word processing program or use a different font type, size, or color to distinguish observations from reflections. Others might create two columns for their full field notes, the first containing notes only about what was observed directly and the other containing reactions and impressions. There is no right or wrong way to write field notes as long as a strategy enables researchers to write accurately in as much detail as possible while distinguishing observations from reflections.

Here is a summary for writing out of the field.

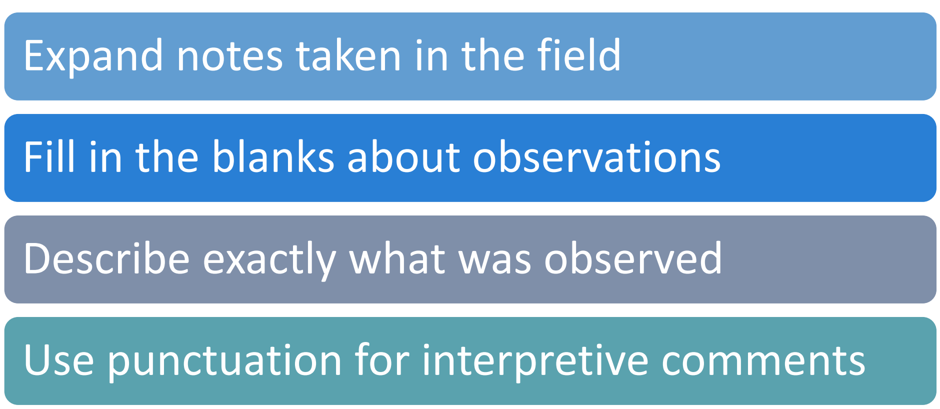


Figure 74: Summary For Writing Out Of The Field

Analysis of Field Research Data

Writing and analyzing field notes involves moving from description to analysis. Some field notes can be primarily descriptive, but some can be more analytic. Analytic field notes are notes that include the researcher's impressions about their observations. Analyzing field note data is a process that occurs over time, beginning when field researchers enter the field and continuing as interactions are happening in the field, as researchers write up descriptive notes, and as they consider what those interactions and descriptive notes mean.

Often field notes will develop from a more descriptive state to an analytic state when the field researchers exit a given observation period. Messy jotted notes or recordings in hand (or in some cases, literally on hand) are entered on a computer to make the notes more readable. Carefully paying attention while in the field is essential; so is what goes on immediately upon exiting the field. Field researchers typically spend several hours keyboarding field notes after each observation has occurred. Keyboarding is often where the analysis of field research data begins. Having time outside of the field to reflect upon their thoughts about what was observed and the meaning of those observations is crucial to developing analysis in field research studies.

Once the analytic field notes have been entered into a computer program, field researchers can begin to look for patterns across the notes by coding the data. Coding involves the iterative process of open and focused coding outlined in Chapter 10. It is crucial that researchers note as much as possible while in the field and as much as can be recalled after leaving the field because they never know what might become critical. Things that seem unimportant at the time may later reveal themselves to have crucial relevance.

Sometimes the analytic process of field researchers and others who conduct inductive analysis is referred to as grounded theory (see Chapter 2). Grounded theory occurs, as the name implies, from the "ground up." It requires that researchers begin with an open-ended and open-minded desire to understand a social situation or setting and involves a systematic process whereby they let the data guide rather than guiding the data by preset hypotheses. The goal when employing a grounded theory approach is, perhaps not surprisingly, to generate theory. Its name implies that discoveries are made from the ground up and that theoretical development is grounded in researchers' empirical observations of a group's tangible experiences.

As exciting as it might sound to generate theory from the ground up, the experience can also be quite intimidating and produce anxiety as the open nature of the process can sometimes feel a little out of control. Without hypotheses to guide their analysis, researchers engaged in grounded theory work may experience some feelings of frustration or angst. The good news is that the process of developing a coherent theory that is grounded in empirical observations can be quite rewarding.

Ethical Issues

Ethics were covered in Chapter 3, but field observations raise several interesting ethical issues that were not considered earlier. While observation may be considered the least obtrusive data-gathering technique, it can include issues of privacy. Interestingly, as late as 1989, a researcher named Jorgensen postulated that observation could not raise privacy issues since there were no human subjects involved. He felt that as long as there was no experimentation going on, then observing was benign. Since then, the United States federal government and all research institutions have specified that observation does, indeed, raise significant privacy concerns. The concerns are even more pronounced when the researcher is engaged in covert (researcher-as-participant) observation, where the researcher’s deception modifies privacy. Spradley (Spradley, 2016) suggested that field researchers follow these guidelines from the American Anthropological Association.

FIGURE 75: ETHICAL ISSUES FOR FIELD OBSERVATIONS

Chatman (Chatman, 1992) suggested that researchers face two different ethical dilemmas: 1) guilty knowledge (where the investigator is privy to confidential information) and 2) dirty hands (where the investigator could correct some wrongdoing but does not). Both of these ethical dilemmas can be very challenging to an investigator. Imagine that one of the participants revealed to the investigator that she was considering suicide during an observation period. What would be the correct course of action? While the strict observation process would perhaps dictate that the investigator does nothing, it would seem human nature to tell someone and try to help the participant. This type of situation creates a very tricky ethical dilemma.

Reliability and Validity

Field observation, like all research, must meet a reasonable reliability and validity standard. Reliability is one goal for both quantitative and qualitative researchers. In quantitative research, reliability is often defined as reproducing the results; however, that is not a reasonable goal for field observation. Instead, it has been suggested that field researchers repeat their observations over varying conditions and several locations.

In qualitative research, like field observation, validity is generally defined as plausible, credible, trustworthy, and defensible. However, a threat to validity is researcher bias. The entire project rests squarely on the researcher, who must determine what to observe, record, and interpret. The researcher could very easily bias the report by, for example, failing to mention anything about prejudice in the workplace. In a nutshell, validity is nothing more than providing an accurate picture of the phenomena under observation. Researchers could triangulate multiple observers or theoretical foundations to improve validity, or they could regularly engage in critical self-reflection.

In short, both reliability and validity are indispensable goals for any research project. However, those terms are defined differently for qualitative projects, like field observation, then for quantitative projects.

A Sampling of Research

Managerial Work

In 1971, Henry Mintzberg wrote a seminal research paper concerning management still cited today (Mintzberg, 1971). His goal was to determine what managers did, and at the time that he conducted his research, there was very little known about management other than anecdotal stories. Mintzberg observed the chief executives of five organizations (a consulting firm, a school system, a technology firm, a consumer goods manufacturer, and a hospital) for one-week periods. He collected both structured (interviews) and unstructured (observations) data organized in three different “records:“

- Chronology record—activity patterns throughout the day

- Mail record—each of 890 physical mail pieces (remember that this was before email, texting, and other electronic communication) was processed. Its purpose, format, sender, and action received were recorded.
- Contact record—each of 368 verbal interactions was noted, and the purpose, medium (telephone, meeting, tour), participants, the form of initiation, and location were recorded.

Mintzberg then provided two sets of conclusions; the first was the characteristics of management work, and the second was the ten roles managers assumed.

CHARACTERISTICS OF MANAGERIAL WORK

1. The manager performs a significant quantity of work at an unrelenting pace. Mintzberg recounted the enormous amount of work expected of managers and stated that "Free time appears to be very rare." Even during off-hours, chief executives spent much time on work-related reading.
2. Managerial activity is characterized by variety, fragmentation, and brevity. Mintzberg indicated that managerial activity had no pattern and was quite fragmented. He also noted that significant activities were interspersed with the trivial in no particular pattern. He stated that "Half of all the activities studied lasted less than nine minutes and only ten percent exceeded on hour's duration."
3. Managers prefer issues that are current, specific, and ad hoc. He indicated that ad hoc reports and uncertain information, like speculation and hearsay, were preferred to historical, specific information. He stated that the managerial environment ". . . breeds, not reflective planners, but adaptable information manipulators . . . "
4. The manager sits between his organization and a network of contacts. Mintzberg found that managers spend a significant amount of time in lateral communication, the external information system. The contacts may be clients, associates, suppliers, peers, trade organizations, and others.
5. The manager demonstrates a strong preference for verbal media. Mintzberg found that managers preferred verbal communications, like telephone, to written forms, like mail. He found that ". . . managers dislike the documented form of communication." Given the age of this study, it is interesting to speculate if managers in the internet age still prefer verbal media, like cell phone calls, to written media, like text messages.

6. Despite the preponderance of obligations, managers appear to be able to control their affairs. Mintzberg noted that other researchers found that managers were little more than puppets reacting to the demands, but his research disagreed with that view. He indicated that managers can define their long-term commitments, like projects and committees, and that managers can exploit situations like ceremonial speeches to lobby for change.

THE MANAGER'S WORK ROLES

Mintzberg categorized 890 pieces of mail and 368 verbal contacts into various roles. These activities involved one of ten roles divided into three basic behaviors: interpersonal contact, processing information, and making decisions.

1. Interpersonal
    - Figurehead. As the legal point of contact in the organization, the manager is a symbol who is obliged to perform many ceremonial duties, sign legal documents, and receive visitors.
    - Leader. The manager's relationship with his subordinates. He must motivate them and develop their working environment. He is also responsible for staffing the organization.
    - Liaison. The lateral communication role of the manager and involves contact with outside peers and experts.
2. Informational
    - Nerve center. The manager serves as the focal point for all non-routine information. He is the information generalist who may not know what an employee is doing daily but knows what the organization is doing at all times. The manager also knows what is going on externally in areas like the competition and the legislative environment.
    - Disseminator. The manager's information must be transmitted to subordinates to be of any value; thus, the manager disseminates information.
    - Spokesman. The manager is obligated to disseminate information to entities outside the organization; thus, he is the official spokesman for the organization.
3. Decisional
    - Entrepreneur. The manager must have a vision of the organization's future and implement processes to move the organization in a defined direction. This work is often done via improvement projects designed to position the organization for future growth and development.

- Disturbance handler. The manager is forced to make corrections as necessary. These disturbances can involve personnel (e. g., a mid-level manager who has been arrested). They can be some systemic issue (e. g. a competitor launching a new product that changes the market).
- Resource allocator. The manager is responsible for having all of the needed resources on hand. The manager also ensures that those resources are allocated appropriately throughout the organization.
- Negotiator. Managers must negotiate with groups who are setting standards for their work, performing support activities, or to whom they wish to sell their goods or services.

IMPROVING THE MANAGER'S EFFECTIVENESS

Mintzberg concludes his report with a brief section where he describes how managers may improve their effectiveness. He states, "As organizations become increasingly large and complex, this burden [of responsibility] increases." Mintzberg suggests that managers can learn more about their roles in the organization to schedule time more efficiently. Managers can recognize that they have much of the information needed to run the company, so they seek better ways to disseminate it. Mintzberg concludes with this thought.

> The ultimate solution to the problem—to the overburdened manager seeking meaningful help—must derive from research. We must observe, describe, and understand the real work of managing; then and only then shall we significantly improve it.

Turkopticon

It is becoming increasingly common to conduct observation via the internet. Researchers find it easy to lurk in various user groups and gather data as people post to those groups. It is also easy for researchers to search for posts about some topic and gather hundreds of comments in just a few minutes. Internet research brings its bag of concerns, including ethics, but places like Facebook and Twitter are proverbial goldmines of data for researchers. Irani and Silberman considered the invisibility of workers at the Amazon Mechanical Turk (AMT), a crowd-sourced worksite, then analyzed those workers' interactions through Turkopticon. On this site, they could share experiences about both good and bad employers. This work was done entirely online and is an excellent example of virtual participant-observation (Irani & Silberman, 2013).

Amazon Mechanical Turk (https://www.mturk.com/) is an online site where employers can hire people from around the world to complete "Human Intelligence Tasks" such as image processing, data verification, and information gathering. Workers are paid per task, often only pennies for simple tasks. For example, one task was to "Clean Up How-To Questions" at a payment of five cents per question. The problem with AMT is that the workers are "invisible" to the employer and can be easily abused. Workers often earn only a few dollars per hour of work, well below the minimum wage in most countries, and non-payment of fees earned is common. The report's author, Lilly Irani, created Turkopticon to provide a platform where AMT workers could rate employers and discuss their work. Irani was a participant-observer with workers on Turkopticon and engaged a few workers in open-ended interviews.

Here are the problems that Irani wanted to address.

- ". . . by hiding workers behind web forms and APIs, AMT helps employers see themselves as builders of innovative technologies, rather than employers unconcerned with working conditions."
- "Once a worker submits work, the employer can choose whether to pay for it. This discretion allows employers to reject work that does not meet their needs, but also enables wage theft."
- ". . . AMT's participation agreement grants employers full intellectual property rights over submissions regardless of rejection, workers have no legal recourse against employers who reject work and then go on to use it."
- ". . . dispute resolution between workers and employers becomes intractable."
- "Dissatisfied workers' within AMT had little option other than to leave the system altogether. Because AMT treats workers interchangeably and because workers are so numerous (tens of thousands by the most conservative estimates), AMT can sustain the loss of workers who do not accept the system's terms."
- "Because Amazon collects money for task volume, Amazon has little reason to prioritize worker needs in a market with a labor surplus."

Irani posted a question on AMT and asked about a "Bill of Rights" for workers. She received 67 responses and noted that these issues were commonly mentioned.

- 35 workers felt that their work was regularly rejected unfairly or arbitrarily
- 26 workers demanded faster payment (Amazon allows employers 30-days to evaluate and pay for work)
- 7 explicitly mentioned a "minimum wage" or "minimum payment" per task
- 14 mentioned "fair" compensation generally

- 8 expressed dissatisfaction with employers' and Amazon's lack of response to their concerns

The remainder of the report concerns the creation and deployment of the Turkopticon and the "bumps" that Irani overcame. Her conclusion includes this statement:

> Turkopticon has succeeded in attracting a growing base of users that sustain it as a platform for an information-sharing community. In part because of its practical embeddedness, it has drawn sustained attention to ethical questions in crowdsourcing over the course of its operation.

Key Takeaways

Chapter 11: Field Research
• "Field Research" was defined as a qualitative method of data collection aimed at understanding, observing, and interacting with people in their natural settings. • The strengths of field research include: it yields detailed data, emphasizes social context, can uncover facts that may not be obvious to the casual observer. • The weaknesses of field research include: the project is ordinarily narrow in scope, it may be emotionally taxing for the researcher, and documentation is challenging. • Selecting a site is an important starting point. Sites can be local but can also be regional, national, or global. Defining the study site is dependent on many factors, not the least of which is the funding available for the study. • The researcher's role in the study can be as distant as a dispassionate observer or as integrated as a group "insider." The selection of the researcher's role will shape the entire project and must be thoughtfully considered at the outset. • Field notes fall into two categories: hastily scribbled notes taken during some activity and more carefully written notes compiled immediately following an activity. Researchers must attempt to keep pure observations separate from their interpretation of those observations. • Analyzing field notes uses a coding process like that used for analyzing interviews.

12: Unobtrusive Research

Introduction

Are female and male athletes at the professional and college levels treated equally? It would be reasonable to think that after 40 years since the passing of Title IX (the civil rights law that prohibits sex discrimination in education, including athletics) that the answer would be an easy yes. Moreover, the growing visibility of women athletes in sports such as golf, basketball, hockey, and tennis would indicate equality. Nevertheless, Professor Michael Messner's (Messner, 2002) unobtrusive research shows otherwise, as does Professors Buysse and Embser-Herbert's (Buysse & Embser-Herbert, 2004) content analysis of photographs in a college athletics media guide. Buysse and Embser-Herbert's unobtrusive research shows that traditional definitions of femininity are fiercely maintained through colleges' visual representations of women athletes as passive and overtly feminine (instead of athletic). Unobtrusive research made it possible to clear up misconceptions about changes for women athletes over the past 40 years.

OBJECTIVES
• Define "Unobtrusive Research." • Describe the strengths and weaknesses of unobtrusive research. • Describe methods used for unobtrusive data collection and analysis. • Describe how data collected by others can be used. • Discuss reliability in unobtrusive research.

What Is Unobtrusive Research?

This chapter explores unobtrusive methods of collecting data, which do not interfere with the subjects under study. Both qualitative and quantitative researchers use unobtrusive research methods. Unobtrusive methods share the unique quality that they do not require researchers to interact with those they are studying. It may seem strange that business, a discipline dedicated to understanding human purchasing behavior, would employ a methodology that requires no interaction with human beings. Nevertheless, humans create plenty of evidence of their behaviors. They:

- write letters to the editor of their local paper;
- create various entertainment sources for themselves such as movies and television shows;
- consume goods;

- lie on the grass in public parks.

All these activities leave something behind—worn paths, trash, recorded shows, and printed papers. These are all potential sources of data for the unobtrusive researcher.

Unobtrusive research methods include content analysis, indirect measures, and using data collected by others. These methods are similar because they do not require direct interaction between researchers and their human subjects, but each has its unique qualities. This chapter also considers how data gathered unobtrusively is analyzed and how reliability can be improved.

Strengths of Unobtrusive Research

Researchers who seek evidence of what people do rather than what they say they do (as in survey and interview research) might consider using unobtrusive methods. Field researchers may also claim this advantage over interview and survey research. However, field researchers cannot be sure about the effect their presence in the field may have on people and their interactions. Like all research projects, unobtrusive research projects face the risk of introducing researcher bias into the work. However, researchers employing unobtrusive methods do not need to be concerned about the effect of the research on their subjects. This effect, known as the Hawthorne effect, is not a concern for unobtrusive researchers because they do not interact directly with their research participants. Mitigating the Hawthorne effect is one of the significant strengths of unobtrusive research.

Another benefit of unobtrusive research is that it can be relatively low-cost compared to some other research methods. Because “participants” are generally inanimate objects instead of human beings, researchers may be able to access data without having to worry about paying participants for their time (though indeed travel to or access to some documents and archives can be costly).

Unobtrusive research is also relatively forgiving. It is far easier to correct mistakes in data collection when conducting unobtrusive research than when using any other method. Imagine the challenge, for example, if researchers realized at the end of conducting 50 in-depth interviews that they had accidentally omitted two critical questions from the interview guide. What options would they have? Re-interview all 50 participants? Try to figure out what respondents might have said based on their other responses? Re-frame the research question? Scratch the project entirely? None of these options are ideal. The same problems arise if a mistake is made in survey research. For field researchers, the consequences of “messing up” during data collection can be even more disastrous. Imagine discovering after tagging along on a political candidate’s campaign that a “do-over” is needed. In this case, that simply is not an option. The campaign is over, and the researcher would need to find a new source of data. Fortunately for unobtrusive researchers, going back to the data source to gather more information or correct some problem in the original data collection is a relatively straightforward prospect.

Finally, unobtrusive research is well suited to studies that focus on processes that occur over time. While longitudinal surveys can preserve long-running processes, they cannot examine processes that occurred decades before data collection began, nor are they the most cost-effective ways to examine wide-ranging processes. On the other hand, unobtrusive methods enable researchers to investigate events and processes that have long since passed. They also do not rely on retrospective accounts, which may be subject to errors in memory, as some longitudinal surveys do.

In summary, these are the strengths of unobtrusive research.

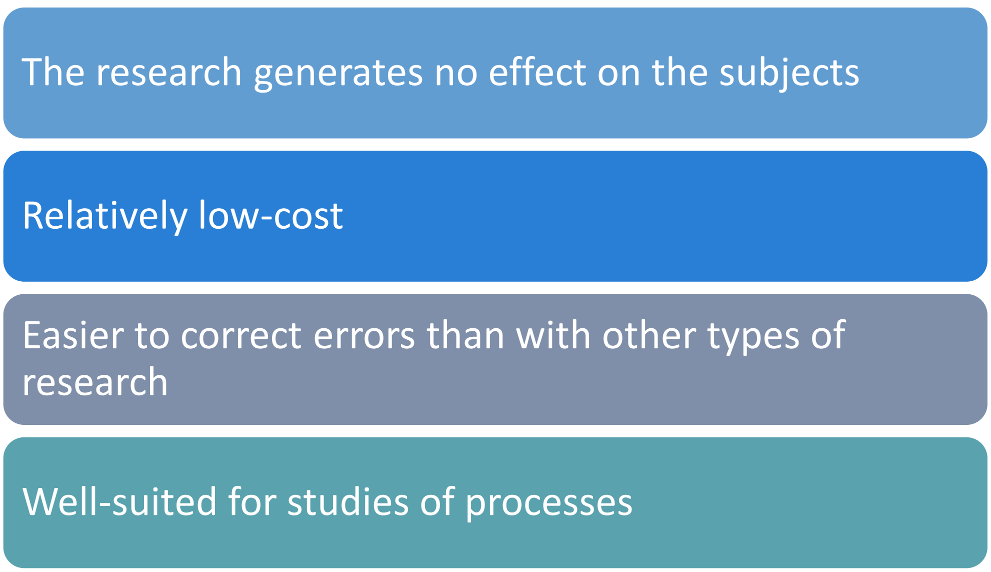


FIGURE 76: STRENGTHS OF UNOBTRUSIVE RESEARCH

Weaknesses of Unobtrusive Research

While there are many benefits to unobtrusive research, this method also comes with a unique set of drawbacks. Because unobtrusive researchers analyze data that may have been created or gathered for purposes entirely different from the researcher's goal, validity problems sometimes arise in such projects. It may also be that data sources do not exist for whatever a researcher is examining. These data challenges mean that unobtrusive researchers may be forced to tweak their original research interests or questions to suit the available data better. Finally, it is not easy to account for context in unobtrusive research projects. For example, in a field research project, the researcher can see what events lead to some occurrence and observe how people respond to that occurrence. It is not easy to ascertain why something occurred in unobtrusive research, but it is easy to know what has occurred.

In summary, these are the weaknesses of unobtrusive research.

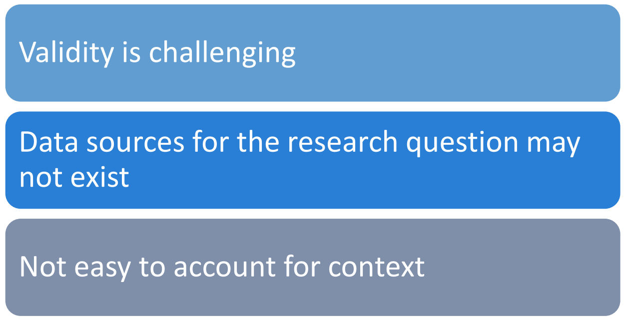


FIGURE 77: WEAKNESSES OF UNOBTRUSIVE RESEARCH

Unobtrusive Data Collection

This section focuses on unobtrusive data collection and what to do with them once collected. There are two main ways of gathering data unobtrusively: conducting a content analysis of existing texts and analyzing physical traces of human behavior, both explored here.

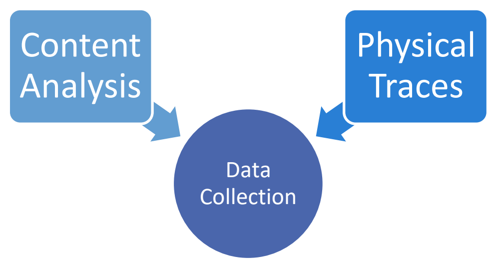


FIGURE 78: TYPES OF UNOBTRUSIVE DATA COLLECTION

Content Analysis

One way of conducting unobtrusive research is to analyze texts, which come in all kinds of formats. Content analysis addresses the questions of "Who says what, to whom, why, how, and with what effect?" (Babbie, The practice of social research (ed.), 2010). Thus, content analysis is a type of unobtrusive research that involves the study of human communications. Another way to think of content analysis is as a way of studying texts and their meaning. This definition of "text" may be more liberal than is found in a dictionary. It includes written documents (e. g., newspapers or letters) and content that was heard (e. g., speeches or other performances). Content analysts might also investigate visual representations of human communication like television shows, photographs, and movies.

Braunsberger and Buckler (Braunsberger & Buckler, 2011) investigated why people participate in consumer boycotts as an example of content analysis. They analyzed comments submitted to an online boycott petition concerning Canadian Seafood. As a result of analyzing a sample of 1200 of the 17,496 boycott pledges, the researchers concluded that 70.1% of the pledges wished it to discontinue its egregious behavior, and 29.67% wanted to send a message that the boycott would impact the company's bottom line. (Note, these two groups overlapped.)

As a second example, Cheyne, Dorfman, and Bukofzer (Cheyne, Dorfman, Bukofzer, & Harris, 2013) analyzed the websites of 16 different kinds of cereal marketed to children like Apple Jacks, Cocoa Puffs, and Lucky Charms. They found that the sites used various progressive levels of presence to encourage children to engage with their products. As measured by traffic data, the more successful sites featured activities with deeper "levels of immersion," like advergames.

Both of these examples used unobtrusive techniques to measure phenomena.

Content analysis is the systematic analysis of the content of a text (e. g., who says what to whom) and is typically conducted as follows.

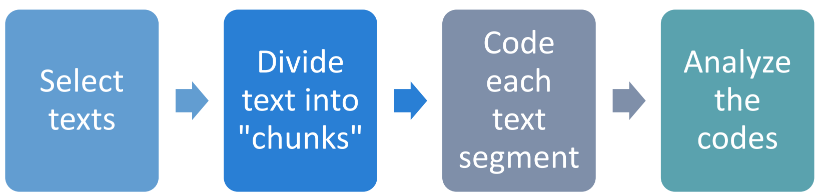


FIGURE 79: CONTENT ANALYSIS PROCESS

1. When there are many texts to analyze (e. g., newspaper stories, and financial reports), the researcher begins by sampling a selected set of texts from the population for analysis. The selection is not random; texts with more pertinent content are specifically chosen.
2. The researcher identifies and applies rules to divide each text into segments or "chunks" that can be treated as separate units of analysis. This process is called unitizing. For example, assumptions, effects, enablers, and barriers in texts may constitute such units.
3. The researcher constructs and applies one or more concepts to each unitized text segment in a coding process. For coding purposes, a scheme is developed from themes discovered as the text is classified.
4. The coded data are analyzed, often quantitatively and qualitatively, to determine which themes occur most frequently, in what contexts, and how they relate to each other.

A simple type of content analysis is sentiment analysis, a technique used to capture people's opinions or attitudes toward an object, person, or phenomenon. Reading messages about a political candidate posted to an online forum and classifying each message as positive, negative, or neutral is an example of such an analysis. In this case, each message represents one unit of analysis. This analysis will help identify whether the sample as a whole is positively or negatively disposed or neutral towards that candidate. Similarly, examining the content of online reviews is another example. Though this analysis can be done manually, for extensive data sets (millions of text records), natural language processing and analytics programs are available to automate the coding process and maintain a record of how sentiments fluctuate with time.

A frequent criticism of content analysis is that it lacks systematic procedures that allow the analysis to be replicated by other researchers. Schilling (Schilling, 2006) addressed this criticism by organizing different content analytic procedures into a spiral model. This model consists of five levels or phases in interpreting text:

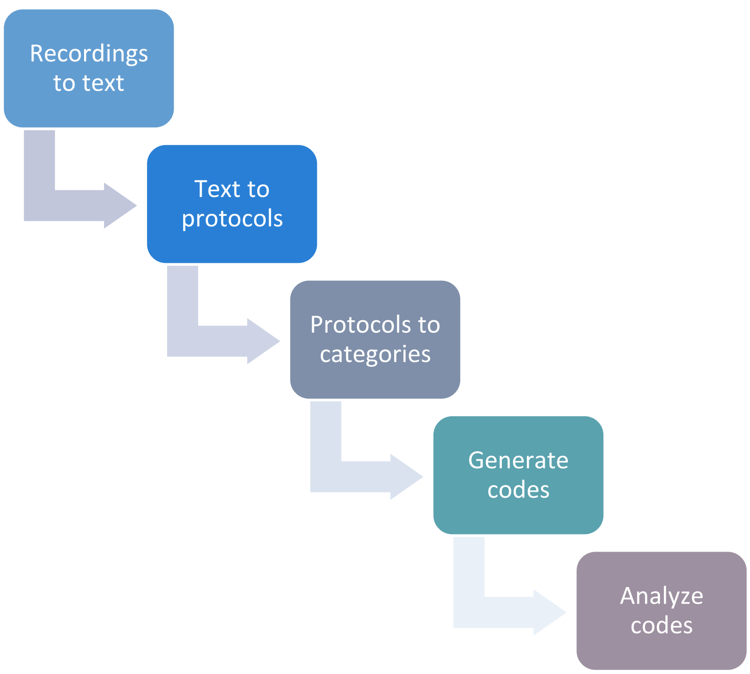


FIGURE 80: ANALYZING CONTENT

1. Convert recorded tapes into raw text data or transcripts for content analysis.
2. Convert raw data into condensed protocols.
3. Convert condensed protocols into a preliminary category system.
4. Use the preliminary category system to generate coded protocols.
5. Analyze coded protocols to generate interpretations about the phenomenon of interest.

Content analysis has several limitations. First, the coding process is restricted to the information available in text form. For instance, if a researcher is interested in studying people's views on capital punishment, no such archive of text documents is available, then the analysis cannot be done. Second, sampling must be done carefully to avoid sampling bias. For instance, if the population is the published research literature on a given topic, then researchers have systematically omitted unpublished research or recent work that is yet to be published.

Hermeneutic Analysis. Hermeneutics is a particular type of content analysis where the researcher attempts to "interpret" the subjective meaning of a given text within its socio-historic context. Unlike grounded theory or content analysis, which ignores the context and meaning of text documents during the coding process, hermeneutic analysis is a genuinely interpretive technique for analyzing qualitative data. This method assumes that written texts narrate an author's experience within a socio-historic context and should be interpreted within that context. Therefore, the researcher continually iterates between a singular interpretation of the text (the part) and a holistic understanding of the context (the whole) to develop a fuller understanding of the phenomenon in its situated context, which German philosopher Martin Heidegger called the *hermeneutic circle*.

More generally, hermeneutics is the study of interpretation and the theory and practice of interpretation. Traditional hermeneutics was derived from religious studies and referred to the interpretation of written texts, especially in religion, literature, and law. In the 20th century, Heidegger suggested that a more direct, non-mediated, and authentic way of understanding social reality is to experience it rather than observing it. The philosophical hermeneutics focus then shifted from interpretation to existential understanding. Heidegger argued that texts are how readers can read about and relive an author's experiences. Contemporary hermeneutics, developed by Heidegger's students, further examined the limits of written texts for communicating social experiences. They proposed a framework of the interpretive process, encompassing all forms of communication, including written, verbal, and non-verbal. They also explored issues that restrict the communicative ability of written texts, such as presumptions, language structures (e. g., grammar and syntax), and semiotics (the study of written signs such as symbolism and analogy).

Hermeneutics is sometimes confused with exegesis, which refers to the interpretation or critical explanation of written text only, especially religious texts. However, this use of hermeneutics improperly restricts it to only written texts.

Following are examples of research projects that used content analysis.

Shen and Bissell (Shen & Bissell, 2013) analyzed the marketing of beauty products on Facebook using content analysis of the product advertising. They found a significant difference in the way beauty products manufacturers and department stores marketed the products. The manufacturers tend to use entertainment like surveys and games in their Facebook ads, while department stores tend to use promotions like coupons and free samples in their ads.

Park (Park & Gretzel, 2007) completed a meta-analysis of published research related to travel destination marketing with websites. Nine success factors were identified for websites that market travel: 1) Information Quality; 2) Ease of Use; 3) Security/Privacy; 4) Visual Appearance; 5) Personalization; 6) Responsiveness; 7) Interactivity; 8) Trust; and, 9) Fulfillment. However, it was also determined that some of the factors were more or less important on travel websites than non-travel, but the importance may shift as web technologies change.

Davis, Piger, and Sedor (Davis, Piger, & Sedor, 2012) completed an analysis of about 23,000 press releases of quarterly earnings statements between 1998 and 2003. They found that "…levels of net optimistic language in earnings press releases are predictive of firm performance in future quarters." In other words, if managers use optimistic language when they release their quarterly earnings reports, it portends future earnings increases.

One thing of note about the above examples is that the data sources represent primary and secondary sources. Primary sources are original research like Shen and Davis, who reported the research results they conducted themselves. On the other hand, secondary sources have already been published and analyzed by others, like Park's published reports about effective website marketing. The distinction between primary and secondary sources is essential for many aspects of business research, but it is imperative to understand when conducting content analysis. While there are certainly instances of content analysis in which secondary sources are analyzed, it is safe to say that it is more common for content analysts to analyze primary sources. Where secondary sources are analyzed, the researcher's focus is usually on how the original analyst or presenter of data reached conclusions or on the choices made in terms of how the data were presented.

Sometimes students new to research methods struggle to grasp the difference between a literature review and content analysis of scholarly literature. A literature review analyzes peer-reviewed articles to understand what is known and unknown about a particular topic. These articles are written by scholars, published in an academic journal, and based on empirical research that has been conducted using accepted data collection and analysis techniques. Scholarly sources are culled to arrive at some conclusion about the overall knowledge concerning a topic. Findings in a literature review are generally taken at face value.

Conversely, a content analysis of scholarly literature would raise questions not raised in a literature review. A content analysis examines scholarly articles to learn something about the authors (e. g., Who publishes what, where?), publication outlets (e. g., How well do different journals represent the diversity of the discipline?), or topics (e. g., How has the popularity of topics shifted over time?).

A content analysis of scholarly articles is a “study of the studies” instead of a “review of studies.” Perhaps, for example, a researcher wishes to know whether more men than women authors are published in the top-ranking journals in the discipline. The researcher could conduct a content analysis of different journals and count authors by gender (though this may be a tricky prospect if relying only on names to indicate gender). Alternatively, perhaps researchers would like to learn whether or how various topics of investigation go in and out of style. They could investigate changes over time in topical coverage in various journals. Content analysts will not summarize the content of articles but instead determine how, why, or by whom particular articles were published.

Content analysis can be qualitative or quantitative, and often researchers will use both strategies to strengthen their investigations. Qualitative content analysis aims to identify themes in the text being analyzed and their underlying meaning. Brown (Brown, 2013) conducted a content analysis of 500 randomly sampled news stories about welfare reform from 1993 to 1997 in California and Arizona. She found that California tended to look at welfare reform as a legal issue while Arizona tended to see it as a racial issue. Quantitative content analysis, on the other hand, involves assigning numerical values to raw data so that it can be analyzed using various statistical procedures. Chavez, Whiteford, and Hoewe (Chavez, Whiteford, & Hoewe, 2010) conducted a quantitative content analysis of United States newspaper reporting about Mexican immigration. They found, for example, that 41.3% of the stories they analyzed were between 501−1000 words long, and the most significant number of stories (50.6%) were about crime.

Physical Traces

Texts are not the only sort of data that researchers can collect unobtrusively. Unobtrusive researchers might also be interested in analyzing the evidence that humans leave behind, telling us something about who they are or what they do. This evidence includes physical traces left by humans and the artifacts that tell us something about their beliefs, values, or norms. Physical traces include worn paths across campus, the materials in a landfill or in someone's trash (a data source Reilly used (Reilly & Wallendorf, 1987)), indentations in furniture, or empty shelves in grocery stores. Examples of material artifacts include video games and video game equipment, sculptures, mementos left on gravestones, housing structures, or even kitchen utensils.

The National Museum of American History in Washington, D.C., has an exhibit displaying chef Julia Child's home kitchen[31], where she filmed many of her famous cooking shows. Seeing the kitchen may help researchers understand how cooking has changed over the decades since Child's shows were on air. For example, they can learn how the layout of kitchens, utensils, and appliances they contain influenced how guests are entertained, how much time is spent preparing meals, and how much time is spent cleaning up afterward. The use of particular kitchen gadgets and utensils might even indicate something about the homeowner's social class. Answers to these questions have a bearing on regular human norms and interactions and are the sorts of questions researchers using unobtrusive methods might be interested in answering.

One challenge with analyzing physical traces and material artifacts is that researchers generally do not have access to the people who left the traces or created the artifacts under analysis. It can be especially tricky to analyze the meanings of these materials if they come from some historical or cultural context other than the researcher's own. Situating the traces or artifacts under analysis in their original contexts and the researcher's own is not always easy. It can lead to problems related to validity and reliability. How can researchers know that they view an object or physical trace how it was intended to be viewed? Do they have the necessary understanding or knowledge about the background of its original creators or users to understand their motivations when they created it?

Imagine an alien trying to understand some aspect of Western human culture simply by examining our artifacts. Cartoonist Mark Parisi demonstrates the misunderstanding that could ensue in his drawing featuring three tiny aliens standing atop a toilet. One alien says, "Since water is the life-blood on this planet, this must be a temple of some sort.. . . Let's stick around and see how they show their respect"

[31] See https://amhistory.si.edu/juliachild/jck/html/textonly/visiting.asp

(1989)[32]. Without a contextual understanding of Western human culture, the aliens have misidentified the purpose of the toilet, and they will be in for quite a surprise when someone shows up to use it!

The point is that while physical traces and material artifacts make excellent sources of data, analyzing their meaning takes more than simply trying to understand them from the researchers' contextual position. They must also be aware of who caused the physical trace or created the artifact, when they created it, why they created, and for whom they created it. Answering these questions will require accessing materials in addition to the traces or artifacts themselves. It may require accessing historical documents or, if a contemporary trace or artifact, another data collection method such as interviews with its creators.

Analysis of Unobtrusive Data

Once the set of texts, physical traces, or artifacts to be analyzed are identified, the next step is to figure out how to proceed with the analysis. This step requires that procedures for coding are developed, the difference between manifest and latent content is understood, and patterns across the coded data are identified.

Coding procedures were introduced to analyze qualitative survey data in Chapter 8. While the coding procedures used for written documents obtained unobtrusively may resemble those used to code interview data, many data sources differ dramatically from written documents or transcripts. For example, how are sculptures, worn paths, or perhaps kitchen utensils coded? The idea of conducting open coding and focused coding on these sources as for a written document seems impossible. So how are patterns across the sculptures or worn paths or utensils identified? One option is to take field notes and then code patterns in those notes. For example, imagine analyzing kitchen utensils. Taking field notes might be a helpful approach for observations of people using utensils on a television program. Keep in mind that if the observation is in person, then the method is no longer unobtrusive.

Instead of observing people in television shows, if the data includes a collection of basic utensils, note-taking may not be the most effective way to record observations. Instead, a code sheet could be developed to record details about the utensils in the sample. A code sheet, sometimes referred to as a tally sheet, is an unobtrusive researcher's instrument for recording observations.

In the example of kitchen utensils, perhaps the research goal is how utensils have changed over time. If researchers had access to sales records for utensils over the past 50 years, then those records could

[32] See https://www.offthemark.com/cartoon/leisure-hobbies/home-garden/2006-05-30

identify the top-selling utensil for each year. Researchers would make some notes about each of the 50 utensils included in the sample to do so. They might note its name, purpose, and perhaps price in current dollar amounts for each top-rated utensil. They might also want to assess how easy or difficult the utensil is to use or some other qualitative assessment about its use or purpose. To rate the difficulty of use, researchers could devise a 5-point scale, with one being very easy to use and five being very difficult to use. They could even record other notes or observations about the utensils that may only come to light after seeing the utensils being used. The following table may be similar to a code sheet developed for a kitchen utensil study.

	1960	1970	1980	1990	2000
Utensil name					
Utensil purpose					
Price (in 1960 dollars)					
Ease of use (1-5 scale)					
Other notes					

TABLE 22: SAMPLE CODE SHEET

The code sheet contains both qualitative and quantitative data. The "ease of use" rating is a quantitative assessment, so statistical analysis of the patterns can be calculated, perhaps noting the mean value on ease of use for each observed decade. Other data are qualitative and would need to be analyzed using open and focused coding to identify patterns. In both cases, whether the data being coded are quantitative or qualitative, the aim is to identify patterns across the coded data.

The "Utensil purpose" row in the sample code sheet provides an opportunity for assessing both manifest and latent content. The manifest content is observed content that is most apparent; it is the surface content. The latent content, on the other hand, is less obvious. Latent content refers to the underlying meaning of the observed surface content. In the example of utensil purpose, a utensil's manifest content may be the stated purpose of the utensil. The latent content may be the researchers' assessment of why that utensil is top rated. Perhaps after coding, the manifest content patterns may emerge that indicate something about the meanings of utensil purpose. Perhaps researchers would conclude that the shift from an emphasis on utensils designed to facilitate entertaining in the 1960s to those designed to maximize efficiency in the 2000s reflects a shift in how people spend time in their homes.

Kathleen Denny's (Denny, 2011) study of scouting manuals offers another excellent example of the differences between manifest and latent content. Denny compared Boy Scout and Girl Scout handbooks to understand gender socializing among scouts. Denny learned from this manifest content that boys are offered more individual-based and more scientific activities by counting activity types described in the manuals. In contrast, girls are offered more group-based and more artistic activities. Denny also analyzed the latent meaning of the messages that scouting handbooks portray about gender; she found that girls were encouraged to become "up-to-date traditional women" while boys were urged to adopt "an assertive heteronormative masculinity."

Analyzing Others' Data

One advantage (or disadvantage, depending on which parts of the research process researchers enjoy most) of unobtrusive research is that researchers may be able to skip the data collection phase altogether. Whether they wish to analyze qualitative data or quantitative data sources, hundreds of free data sets are available to researchers. For example, the United States Census Bureau makes both raw data and reports available from their website. Researchers can find information about population demographics down to the city block level in some cases, economic indicators like income and rent, education levels, country of origin, and a wide variety of other data.

The following list contains only a few of the more commonly-used public data sources for business research. Note that the following URLs are for government (.gov) or education (.edu) domains since those organizations would be more likely to post unbiased data. The only exceptions are two .org sites that belong to the United Nations and the World Bank since those would also post unbiased data.

Public Databases

- **Agency for Healthcare Research and Quality**. This site is a compendium of health systems in the United States. A "health system" is defined as an activity that includes at least one hospital and one group of physicians providing comprehensive care at that hospital. This database provides information about those systems, like the system's name, number of physicians, and number that serves children. https://www.ahrq.gov/chsp/data-resources/compendium.html
- **Bureau of Justice Statistics**. This site contains more than 60 databases covering many aspects of the United States criminal justice system. Included are databases like "Annual Probation Survey and Annual Parole Survey," "Census of Jail Inmates," and "Recidivism of State Prisoners." The site is well-organized, and it is easy to find data of interest. https://www.bjs.gov/index.cfm?ty=dca

- **Bureau of Labor Statistics**. This bureau makes data about labor available to researchers. Included are databases about inflation and prices, employment, unemployment, projections, pay and benefits, spending and time use, productivity, workplace injuries, occupational requirements, and regional and international resources. https://www.bls.gov/
- **Bureau of Economic Analysis**. This site is the Department of Commerce's economic analysis databases. Researchers can find data about gross domestic product, fixed assets, and personal income using this site. https://www.bea.gov/
- **Census Bureau**. The United States Census Bureau has a vast wealth of data about the US population dating back to the 1700s. While most researchers know that the Census Bureau collects data about the number of people who live in a region, they also have data about race, education level, income, and other demographic factors. https://data.census.gov/cedsci/?intcmp=aff_cedsci_banner
- **Data.Gov**. This site is an aggregator of more than 200,000 public data sets. It is well organized and contains data from governmental and educational sources. In general, this should be the first stop for researchers seeking public data for research projects. https://www.data.gov/
- **Department of Agriculture**. The United States Department of Agriculture provides multiple data files and apps, charts, and maps to tell the story of agriculture. Examples of the data files include aquaculture, dairy data, and bioenergy statistics. https://www.ers.usda.gov/
- **Department of Education**. This site contains data about education in the United States, including both K-12 and post-secondary. The data sets available include the National Student Loan Data System, College Scorecard, Integrated Postsecondary Education Data System, and School Survey on Crime and Safety. https://www2.ed.gov/rschstat/landing.jhtml?src=pn
- **Department of Health and Human Services Health Data**. This site includes more than 3,000 data sets related to health and wellness that various government agencies provide. https://healthdata.gov/
- **Dept of Housing and Urban Development**. The HUD posts data sets concerning housing in the United States. These data sets include the American Housing Survey, Fair Market Rents, and Geospatial Data Resources. https://www.huduser.gov/portal/pdrdatas_landing.html
- **United States International Trade Administration Exports**. Data concerning international trade that originates in the United States. The data sets are divided into national, state, and metro sections. https://www.export.gov/Trade-Data-and-Analysis

- **Federal Bureau of Investigation Crime Data**. The FBI makes crime data available for researchers. The various categories are Assaults on Law Enforcement Officers, Police Employee Data, Hate Crime, Human Trafficking, Uniform Crime Reporting Program Participation Data, Cargo Theft, U.S. Territory Data, and Arrest Data. Researchers can also find the "Summary (SRS) Data with Estimates" that includes the data used by the FBI's annual publications. https://crime-data-explorer.fr.cloud.gov/
- **Federal Housing Finance Agency**. This agency tracks data related to housing in the United States. The data sets include the house price index, market data, and the National Mortgage Database. https://www.fhfa.gov/DataTools/Downloads
- **Federal Reserve**. The federal reserve makes data about banking, finance, and exchange rates available for researchers. These data sets include the Survey of Small Business Finances, Mortgage Debt Outstanding, and Industrial Production and Capacity Utilization. https://www.federalreserve.gov/data.htm
- **Foreign Assistance**. The United States offers more than $25 billion in foreign aid to many countries around the world. Data files can be downloaded by country, U.S. agency, or program.https://www.foreignassistance.gov/
- **Harvard Dataverse**. Students at Harvard University conduct thousands of research projects every year. They submit their raw data to the Harvard Dataverse, and those data can be downloaded by researchers anywhere. The site has more than 85,000 datasets organized into 13 different subjects, like business and management, law, and social science. https://dataverse.harvard.edu/
- **Inter-university Consortium for Political and Social Research (ICPSR)**. The University of Michigan has made more than 11,000 social science-related data sets available. As an example of the data available, the "500 Family Study" includes "…in-depth information on middle class, dual-career families living in the United States." The data are divided into four data sets, the Cortisol Data that examines psychological stress, the Experience Sampling Method Data that examines how individuals spend their time, the Parent Data the examines parents' occupations and other information, and the Adolescent Data that examines the family relationships from an adolescent perspective. https://www.icpsr.umich.edu/icpsrweb/ICPSR/
- **National Center for Health Statistics**. This site posts data from the Centers for Disease Control concerning health statistics. It includes data like birth and death rates, the Longitudinal Studies

of Aging, and the National Survey of Children's Health. https://www.cdc.gov/nchs/data_access/ftp_data.htm

- **National Centers for Environmental Information**. This site posts data provided by the National Oceanic and Atmospheric Administration about the environment. It includes historical weather information, satellite radiance data, and paleoclimatology. https://www.ncdc.noaa.gov/data-access
- **United Nations Statistics Division**. The UN provides data sets that include population, national accounts, education, labor, price indices, and many other factors for every nation and geographical area (like Northern Africa). http://data.un.org/
- **World Bank Open Data**. The World Bank posts data related to banking and monetary policy for countries around the world. The data can be browsed by country/region, time, or geospatial values. https://datacatalog.worldbank.org/

Public Document Repositories

The following list contains a few of the many repositories for reports and other published documents.

- **CIA World Factbook**. The CIA Factbooks are detailed reports that would be valuable to anyone who needs background information about a country. https://www.cia.gov/the-world-factbook/
- **Google Scholar**. This resource can be used to search for papers published in thousands of different scholarly journals, along with dissertations and thesis that may not have been published in a journal. It is the "go-to" source for searches for scholarly publications. https://scholar.google.com
- **National Archives**. The National Archives are familiar to people researching their ancestry, but the archives include documentation about businesses, foundations, countries, governmental contracts, and even periods like 1800−1900. https://www.archives.gov/research
- **Public Library of Science**. PLOS is a nonprofit publisher of more than 215,000 peer-reviewed scientific articles in many different fields. https://www.plos.org/
- **US Congress**. This site is for the United States Congress. It includes the text of all Senate and House bills along with a daily digest of congressional activities. https://www.congress.gov/

Keep in mind that the resources mentioned here represent just a snapshot of the many sources of publicly available data that can be easily accessed via the web. Also, it is essential to note that the above data sources are appropriate for United States research, but students interested in other countries will find similar data for most industrialized nations.

Reliability In Unobtrusive Research

This final section of the chapter investigates reliability concerns in unobtrusive research projects that warrant attention (Krippendorff, 2009). The primary concerns have to do with how and by whom the coding of data occurs. Additionally, stability, reproducibility, and accuracy issues present unique problems for reliability in unobtrusive research projects.

Stability refers to the extent to which the results of coding vary across different periods. If stability is a problem, it will reveal itself when the same person codes the same content with different results. Coding is stable when the same content has been coded multiple times by the same person with consistent results. Researchers who discover instability problems in their coding procedures should revise their coding rules, so they are less ambiguous. Ambiguities in the text itself might also contribute to problems of stability. While the original text documents cannot be altered, simply being aware of possible ambiguities in the data may help reduce the likelihood of problems with stability. It is also possible that problems with stability may result from a coding error, such as inadvertently jotting a "1" instead of a "10" on the code sheet.

Reproducibility, or intercoder reliability, is how well the procedure results in the same outcome when different people code the exact text. Cognitive differences among the individuals coding data may result in problems with reproducibility, as could ambiguous coding instructions. Random coding errors might also cause problems. One way of overcoming problems of reproducibility is to have coders code together. Resolving coding ambiguities as a team leads to a shared understanding of how to code various bits of data.

Finally, accuracy refers to the extent to which one's coding procedures correspond to some preexisting standard. This process presumes that a standard coding strategy has already been established for whatever text is being analyzed. It may not be the case that official standards have been set but perusing the prior literature for the collective wisdom on coding in a particular area is time well spent. Scholarship focused on similar data or coding procedures will undoubtedly help clarify and improve the coding procedures.

Key Takeaways

Chapter 12: Unobtrusive Research
• Unobtrusive research uses methods that do not interfere with the subjects under study. • The strengths of unobtrusive research include less bias introduced than in field research, low cost, forgiving of error, and is particularly well-suited for longitudinal types of studies. • The weaknesses of unobtrusive research include a mismatch between the data used and the research goals and the inability to account for context in the data. • Data collection methods include content (and hermeneutic) analysis and indirect measures. • Data collected with unobtrusive methods are analyzed with coding manuals, similar to that used for interviews. • There are hundreds of public databases available for researchers who are willing to use existing data.

13: Interpretive Research

The influence of culture on shopping behavior raises many questions that business owners must consider. Do Christmas decorations in a mall entice shoppers to come in, or do they drive away customers whose religious beliefs do not include Christmas? What role do social conventions like family, schooling, leisure activities, and religion play in determining where to shop and what to buy? Should a merchant cater to the majority or attempt to be neutral in advertising and product selection activities? While it is reasonably easy to market to a homogeneous population (think about the food products sold in a largely Hasidic Jewish area), it is much more difficult to market on a regional or more extensive level. Interpretive research attempts to answer questions about "who lives here," which would be of vital interest to merchants.

Objectives
• Define "Interpretive Research." • Discuss data collection techniques for interpretive research. • Define the various types of interpretive research: case research, action research, and ethnography. • Discuss rigor in interpretive research.

Introduction

Recall that positivist or deductive research methods, such as laboratory experiments and survey research, are explicitly for theory (or hypothesis) testing. In contrast, interpretive research or inductive research methods, such as action research and ethnography, are intended for theory building. Unlike a positivist method, where researchers start with a theory and test theoretical postulates using empirical data, in interpretive methods, researchers start with data and work to derive a theory about the phenomenon of interest from the observed data.

The term interpretive research is often used loosely and synonymously with qualitative research, although the two concepts are different. Interpretive research is based on the assumption that social reality is not singular or objective. Instead, it is shaped by human experiences and social contexts (ontology) and is, therefore, best studied within its socio-historic context by reconciling subjective interpretations of its various participants (epistemology). Because interpretive researchers view social reality as embedded within and impossible to abstract from their social settings, they "interpret" the reality through a "sense-making" process rather than a hypothesis testing process. In contrast, the positivist paradigm assumes that the reality is relatively independent of the context and can be studied in a practical manner using objective techniques such as standardized measures. Whether a researcher should pursue interpretive or positivist research depends on the nature of the phenomenon under consideration and the best way to study it.

Qualitative versus quantitative research refers to considerations about the type of data to collect and analyze. Qualitative research relies mainly on qualitative data, such as interviews and observations, in contrast to quantitative research, which employs quantitative data such as surveys. Hence, qualitative research is not amenable to statistical procedures such as regression analysis but is coded using content analysis. Sometimes, coded qualitative data is tabulated quantitatively as frequencies of observations, but these data are not statistically analyzed. Many interpretive researchers reject this coding approach as a futile effort to seek objectivity in a social phenomenon that is essentially subjective.

Although interpretive research tends to rely heavily on qualitative data, quantitative data may add more precision and a clearer understanding of the phenomenon of interest than qualitative data. For example, Eisenhardt (Eisenhardt, Making fast strategic decisions in high-velocity environments, 1989) studied high-velocity firms. She collected data on how long it took to make strategic decisions, ranging from 1.5 to 18 months, and how many alternatives were considered for each decision. She then surveyed her respondents to capture their perceptions of organizational conflict. The numeric data helped her clearly distinguish the high-speed decision-making firms from the low-speed decision makers without relying on respondents' subjective perceptions. The quantitative data then allowed her to examine the number of decision alternatives considered and the extent of conflict in high-speed versus low-speed firms. Interpretive research should attempt to collect both qualitative and quantitative data about the phenomenon of interest. Joint use of qualitative and quantitative data is often called "mixed-mode designs" and may lead to unique insights prized in the scientific community.

Interpretive research has its roots in anthropology, sociology, psychology, linguistics, and semiotics. This type of research has been available since the early 19th century, long before positivist techniques were developed. Many positivist researchers view interpretive research as erroneous and biased, given the subjective nature of the qualitative data collection and interpretation process employed in such research. However, the failure of many positivist techniques to generate exciting insights or new knowledge has resulted in a resurgence of interest in interpretive research since the 1970s, albeit with exacting methods and stringent criteria to ensure the reliability and validity of interpretive inferences.

Distinctions From Positivist Research

In addition to fundamental differences in philosophy, interpretive and positivist research differ in several other ways. Interpretive research employs a directed sampling strategy, where study sites, respondents, or cases are selected based on two considerations:

- Does the sample fit the phenomenon being studied?
- Does the sample possess characteristics that make it uniquely suited for the study?

In contrast, positivist research employs random sampling (or a variation of this technique), where cases are chosen randomly from a studied population. Hence, convenience samples and small samples are considered acceptable in interpretive research as long as they fit the purpose of the study but not in positivist research.

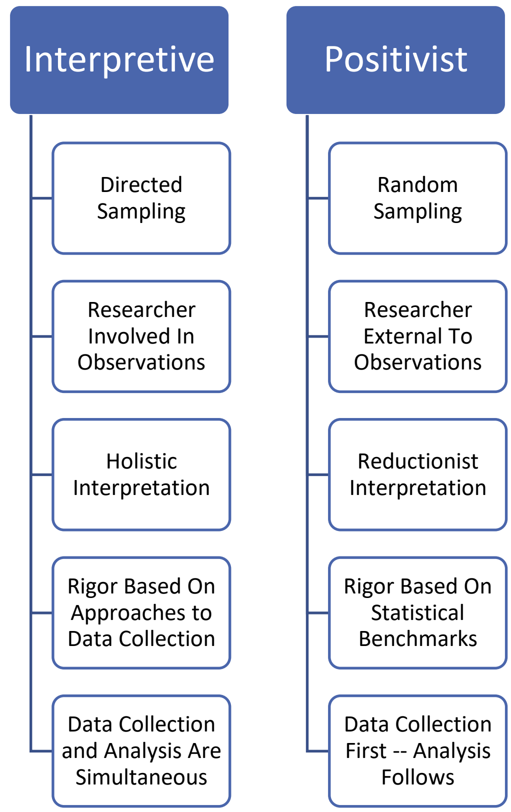


FIGURE 81: INTERPRETIVE VS. POSITIVIST RESEARCH

Second, the role of researchers receives critical attention in interpretive research but not positivist. Researchers are considered part of the social phenomenon studied in some methods such as ethnography, action research, and participant observation. Their specific role and involvement in the research process must be made clear during data analysis. In other methods, such as case research, researchers must take a “neutral” or unbiased stance during the data collection and analysis processes and ensure that their personal biases or preconceptions do not taint the nature of subjective inferences derived from interpretive research. In positivist research, however, researchers are considered external to and independent of the research context and are not presumed to bias the data collection and analytic procedures.

Third, interpretive analysis is holistic and contextual rather than reductionist and isolationist. Interpretive conclusions tend to focus on language, signs, and meanings from the participants' perspective in the social phenomenon, in contrast to statistical techniques employed heavily in positivist research. Rigor in interpretive research is evaluated in systematic and transparent data collection and analysis approaches rather than statistical benchmarks for construct validity or significance testing.

Lastly, data collection and analysis can proceed simultaneously and iteratively in interpretive research. For instance, researchers may conduct an interview and code it before proceeding to the following interview. Simultaneous analysis helps researchers correct potential flaws in the interview protocol or adjust it to better capture the phenomenon of interest. Researchers may even change their original research questions if they realize that their original questions are unlikely to generate new or valuable insights. Flexibility is a valuable but often understated benefit of interpretive research. It is not available in positivist research where the project cannot be modified once data collection has started without redoing the entire project.

Benefits And Challenges Of Interpretive Research

Interpretive research has several unique advantages.

- It is well-suited for exploring hidden reasons behind the complex, interrelated, or multifaceted social processes, such as inter-firm relationships or inter-office politics, where quantitative evidence may be biased, inaccurate, or otherwise difficult to obtain.
- It is often helpful for theory construction in areas with no or insufficient a priori theory.
- It is also appropriate for studying context-specific, unique, or idiosyncratic events or processes.
- Interpretive research can also help uncover exciting and relevant research questions and issues for follow-up research.

At the same time, interpretive research has its own set of challenges.

- This type of research tends to be more time- and resource-intensive than positivist research in data collection and analytic efforts. Too little data can lead to false or premature assumptions, while too much data may not be effectively processed.
- Interpretive research requires well-trained researchers capable of seeing and interpreting complex social phenomena from the participants' perspectives and reconciling them without injecting their personal biases into the inferences.
- All participants or data sources may not be equally credible, unbiased, or knowledgeable about the phenomenon of interest. They may have undisclosed agendas that may lead to misleading or false impressions. Inadequate trust between participants and the researcher may hinder

complete and honest self-representation by participants, and trust-building takes time. It is the job of the interpretive researcher to "see through the smoke" (hidden or biased agendas) and understand the true nature of the problem.

- Inferences drawn from interpretive research are heavily contextualized and not replicable or generalizable.
- Interpretive research may sometimes fail to answer the research questions of interest or predict future behaviors.

Characteristics of Interpretive Research

All interpretive research adheres to a standard set of principles.

- **Naturalistic inquiry**. Social phenomena must be studied within their natural setting. Because interpretive research assumes that social phenomena are situated within and cannot be isolated from their social context, interpretations of such phenomena must be grounded within their socio-historical context. This requirement implies that contextual variables should be observed and considered in seeking explanations of a phenomenon of interest, even though context-sensitivity may limit the generalizability of inferences.
- **Researcher as an instrument**. Researchers are often embedded within the social context they study and are considered part of the data collection instrument. They must use their observational skills, trust the participants, and extract correct information. Further, their insights, knowledge, and experiences of the social context are critical to accurately interpreting the phenomenon of interest. At the same time, researchers must be fully aware of their personal biases and preconceptions and not let such biases interfere with their ability to present a fair and accurate portrayal of the phenomenon.
- **Interpretive analysis**. Observations must be interpreted through the eyes of the participants embedded in the social context. Interpretation must occur at two levels. The first level involves viewing or experiencing the phenomenon from the subjective perspectives of the social participants. The second level is to understand the meaning of the participants' experiences to provide a "thick description" or a rich narrative story of the phenomenon of interest that can communicate why participants acted the way they did.
- **Use of expressive language**. Documenting participants' verbal and nonverbal language and analyzing such language are integral components of interpretive analysis. The study must ensure that the story is viewed through the eyes of a person and not a machine. It must depict the

emotions and experiences of that person so readers can understand and relate to that person. The use of imageries, metaphors, sarcasm and other figures of speech is widespread in interpretive analysis.

- **Temporal nature**. Interpretive research is often not concerned with searching for specific answers but with understanding or "making sense of" a dynamic social process as it unfolds over time. Hence, the researcher must be immersed at the study site for an extended period to capture the entire evolution of the phenomenon of interest.
- **Hermeneutic circle**. Interpretation is an iterative process of moving back and forth from observations (text) to the entirety of the social phenomenon (context) to reconcile their apparent discord and construct a theory with the participants' diverse viewpoints and experiences. Such iterations between the generating understanding and recording observations of phenomena must continue until "theoretical saturation" is reached, where any additional iteration does not yield more insight.

Interpretive Data Collection

Data are collected in interpretive research using a variety of techniques. The most frequently used technique is interviewing (See Chapter 10 for more information about interviews). A second technique is observing. Two types of observation are commonly used. Direct observation is where researchers are a neutral and passive external observer, as in case research. Participant observation is where researchers are active participants, as in action research. A third technique is reviewing documents, where external and internal documents, such as memos and newspaper articles, may be used for further insight or corroborate other evidence.

Interpretive Research Designs

Case research.

Case research is a method of intensively studying a phenomenon over time within its natural setting at one or a few sites. Multiple data collection methods, such as interviews, observations, prerecorded documents, and secondary data may be employed, and inferences about the phenomenon of interest tend to be rich, detailed, and contextualized. Case research can be employed in a positivist manner for theory testing or in an interpretive manner for theory building. This method is more prevalent in business research than in other social science disciplines.

CASE RESEARCH VS. CASE DESCRIPTION

It is essential to recognize that case research is different from case descriptions such as Harvard case studies discussed in business classes. Case descriptions typically describe an organizational problem in rich detail to stimulate classroom discussion and critical thinking among students or analyze how well an organization handled a specific problem. Case research is a formal research technique that involves a specific method intended to derive explanations of organizational phenomena. Often, a case research project generates a case description.

Case research has several unique strengths over competing research methods such as experiments and survey research.

- Case research can be used for either theory-building or theory testing, while positivist methods can only be used for theory testing. In interpretive case research, the constructs of interest need not be known in advance but may emerge from the data as the research progresses.
- The research questions can be modified during the research process if the original questions prove less relevant or salient. Reframing the research question is not possible in any positivist method after the data are collected.
- Case research can help derive a richer, more contextualized, and more authentic interpretation of the phenomenon of interest than most other research methods by its ability to capture a rich array of contextual data.
- The phenomenon of interest can be studied from the perspectives of multiple participants and using multiple levels of analysis (e. g., individual and organizational).

At the same time, case research also has some inherent weaknesses.

- Because it involves no experimental control, the internal validity of inferences remains weak. Of course, this is a common problem for all research methods except experiments. However, as described later, controls may be addressed in case research using "natural controls."
- The quality of inferences derived from case research depends heavily on the integrative powers of the researcher. An experienced researcher may see concepts and patterns in case data that a novice researcher may miss. Hence, the findings are sometimes criticized as being subjective.
- Because the inferences are heavily contextualized, it may be challenging to generalize inferences from case research to other contexts or other organizations.

Case research is a complex research method that requires advanced research skills on the researcher's part and is often prone to error. Benbasat (Benbasat, Goldstein, & Mead, 1987) describes five problems frequently encountered in case research studies.

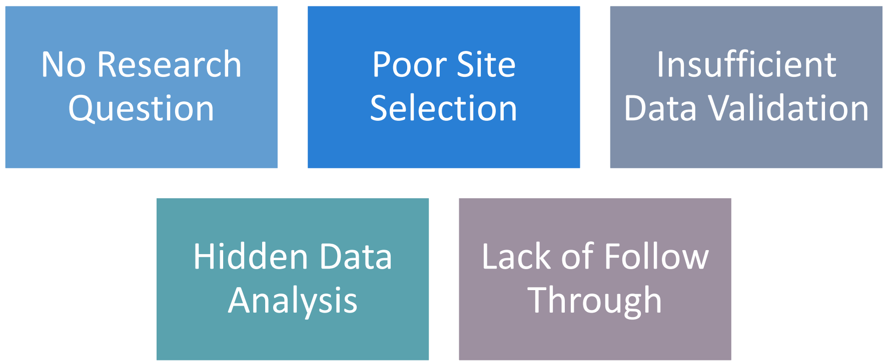


FIGURE 82: PROBLEMS WITH CASE RESEARCH

- Many case research studies start without specific research questions and therefore end up without having any specific answers or insightful inferences.
- Case sites are often chosen based on access and convenience rather than fit with the research questions and may not address those questions.
- Researchers often do not validate or triangulate data collected using multiple means, leading to biased interpretation based on responses from biased interviewees.
- Many studies provide few details on how data were collected (e.g., the interview questions posed and the documents examined) or analyzed, which may raise doubts about the reliability of the inferences.
- Despite its strength as a longitudinal research method, many case research studies do not follow through a phenomenon longitudinally and present only a cross-sectional view of temporal organizational processes.

Several critical decisions must be made by a researcher when considering a case research method.

- Is this a suitable method for the research questions being studied? The case research method is particularly appropriate for 1) exploratory studies designed to discover new constructs in areas where theory building is at the formative stages, 2) studies where the experiences of participants and context of actions are critical, and 3) for studies aimed at an understanding of complex, temporal processes (why and how) rather than their causes (what). This method is also well-suited for studying complex organizational processes that involve multiple participants and interacting sequences of events, such as organizational change and large-scale technology implementation projects.

- What is the appropriate unit of analysis for a case research study? Since case research can simultaneously examine multiple units of analysis, researchers must decide whether they wish to study a phenomenon at the individual, group, and organizational level or multiple levels. For instance, a study of group decision making or group work may combine individual-level constructs such as individual participation in group activities with group-level constructs, such as group cohesion and group leadership. Thus, case studies can potentially derive a richer understanding than that can be achieved from a single level of analysis.
- Should researchers employ a single-case or multiple-case design? The single-case design is appropriate:
    - at the outset of theory generation
    - if the situation is unique or extreme
    - if it is revelatory (i. e., the situation was previously inaccessible for scientific investigation)
    - if it represents a critical or contrary case for testing a well-formulated theory

  The multiple-case design is appropriate:
    - for theory testing
    - establishing the generalizability of inferences
    - developing nuanced interpretations of a phenomenon
- What sites should be chosen for case research? Given the contextualized nature of inferences derived from case research, site selection is a particularly critical issue because selecting the wrong site may lead to the wrong inferences. If the research goal is to test theories or examine the generalizability of inferences, then dissimilar case sites should be selected to increase variance in observations. For instance, if the research goal is to understand the process of technology implementation in firms, a mix of large, mid-sized, and small firms should be selected to examine whether the technology implementation process differs with firm size. Site selection should not be opportunistic or based on convenience, but rather based on the fit with research questions through a process called “theoretical sampling.”

- How can research rigor be improved? Yin (Yin, 2017) recommends using multiple case sites with replication logic, viewing each case site as similar to one experimental study, and then following rules of scientific rigor similar to that used in positivist research.
- What techniques of data collection should be used in case research? Although interviews are the most popular data collection technique for case research, interview data can be supplemented with techniques like direct observation, documentation, archival records, and physical artifacts. Furthermore, the researcher should triangulate or validate observed data by comparing responses between interviewees.

Most case research studies tend to be interpretive. Interpretive case research is an inductive technique where evidence collected from one or more case sites is systematically analyzed and synthesized to allow concepts and patterns to emerge to build new theories or expand existing ones. In any given case research project, some of the following stages may need to be rearranged or modified; however, sampling, data collection, and analysis techniques should match traditional positivist methods as much as possible.

- **Define research questions**. Like any other scientific research, case research must define research questions theoretically and practically and identify some intuitive expectations about possible answers to those research questions or preliminary constructs to guide initial case design. In positivist case research, the research questions are based on existing theory, but in interpretive case research, no prior theories are available but are developed as the project unfolds. As a result, in interpretive case research studies, research questions may be changed as the study develops, contrasting with positivist case research.

- **Select case sites**. Researchers should use a process of "theoretical sampling" (not random sampling) to identify case sites. In this approach, case sites are chosen based on theoretical, rather than statistical, considerations, such as replicating previous cases, extending preliminary theories, or filling anticipated categories or polar types. Care should be taken to ensure that the selected sites fit the nature of research questions, minimize extraneous variance or noise due to firm size, industry effects, and maximize variance in the dependent variables of interest. For example, suppose the research goal is to examine how some firms innovate better than others. In that case, researchers should select firms of similar size within the same industry to reduce industry or size effects and select some more innovative and some less innovative firms to increase variation in firm innovation. Instead of cold-calling or writing to a potential site, it is better to contact someone at the executive level inside each firm who has the authority to approve the project or someone who can identify a person of authority. During initial conversations, researchers should describe the project, including items like:

- potential benefits to the site
- how the collected data will be collected and used
- who is involved in the data collection
- the desired interviewees
- the amount of time, effort, and expense required of the sponsoring organization

Researchers must also assure confidentiality, privacy, and anonymity of both the firm and the individual respondents.

- **Create instruments and protocols**. Since interviews are the primary mode of data collection in case research, an instrument, or list of questions to be asked, should guide the interview process. Questions may be open-ended (unstructured), closed-ended (structured), or a combination of both. The interview protocol must be strictly followed, and the interviewer must not change the order of questions or skip any questions during the interview process. However, some deviations are allowed to probe further into respondent's comments that are ambiguous or interesting. The interviewer must maintain a neutral tone, not lead respondents in any specific direction, and agree or disagree with any response. In addition, additional sources of data, such as internal documents and memorandums, annual reports, financial statements, newspaper articles, and direct observations, should be sought to supplement and validate interview data.

- **Select respondents**. Select interview respondents at different organizational levels, departments, and positions to obtain divergent perspectives on the topic of interest. A random sampling of interviewees is preferable; however, a snowball sample is acceptable, as long as a diversity of perspectives is represented in the sample. Interviewees must be selected based on their involvement with the phenomenon under investigation and their ability and willingness to answer the researcher's questions accurately and adequately, not on convenience or access.

- **Start data collection**. It is usually a good idea to electronically record interviews for future reference. However, such recording must only be done with the interviewee's consent. Even when interviews are being recorded, the interviewer should take notes to capture essential comments or critical observations, behavioral responses (e.g., respondent's body language), and the researcher's impressions about the respondent and his/her comments. After each interview is completed, the entire interview should be transcribed verbatim into a text document for analysis.

- **Conduct within-case data analysis**. Data analysis may follow or overlap with data collection. Overlapping data collection and analysis has the advantage of adjusting the data collection process based on themes emerging from data analysis or further probe into these themes. Data analysis is done in two stages. In the first stage (within-case analysis), the researcher should examine emergent concepts separately at each case site and patterns between these concepts to generate an initial theory of the problem of interest. The researcher can review interview data subjectively to "make sense" of the research problem using personal observations or experience at the case site. Alternatively, a coding strategy like a grounded theory approach, using open coding, axial coding, and selective coding, may derive a chain of evidence and inferences. Homegrown techniques, such as a graphical representation of data (e. g., network diagram) or sequence analysis (for longitudinal data), may also be used. Note that there is no predefined way of analyzing the various types of case data. The analytic data techniques may need to be modified to fit the nature of the research project.

- **Conduct cross-case analysis**. Multi-site case research requires cross-case analysis as the second stage of data analysis. Researchers should look for similar concepts and patterns between different case sites in such analysis, ignoring contextual differences that may lead to quirky conclusions. Such patterns may be used to validate the initial theory or refine it (by adding or dropping concepts and relationships) to develop a more inclusive and generalizable theory. This analysis may take several forms. For instance, researchers may select categories (e. g., firm size, or industry) and look for within-group similarities and between-group differences (e. g., high versus low performers or innovators versus laggards).

- **Build and test hypotheses**. Based on emergent concepts and themes that are generalizable across case sites, tentative hypotheses are constructed. These hypotheses should be compared iteratively with empirical evidence to see if they fit the observed data, and if not, the constructs or relationships should be refined. Also, researchers should compare the emergent constructs and hypotheses with those reported in the prior literature to make a case for their internal validity and generalizability. Conflicting findings must not be rejected but instead reconciled using creative thinking to generate greater insight into the emergent theory. When further iterations between theory and data yield no new insights or changes in the existing theory, "theoretical saturation" is reached, and the theory-building process is complete.

- **Write the case research report**. In writing the report, researchers should clearly describe the process used for sampling, data collection, data analysis, and hypotheses development so readers can independently assess the reasonableness, strength, and consistency of the reported inferences. A high level of clarity in research methods is needed to ensure that the researcher's preconceptions do not bias the findings.

Action research.

Action research is a qualitative but positivist research design aimed at theory testing rather than theory building. This interactive design assumes complex social phenomena are best understood by introducing changes, interventions, or "actions" into those phenomena and observing the outcomes of such actions on the phenomena of interest. The researcher is usually a consultant who initiates an action in response to a social problem. Then, the researcher examines the action's influence on the phenomenon and generates insights into the relationship between the action and phenomenon. An example of action may be an organizational change program (e. g., introducing new technology) initiated to improve an organization's performance. The researcher's choice of actions must be based on theory, explaining why and how such actions may yield the desired change. The theory is validated by how the chosen action is successful in remedying the targeted problem. Simultaneous problem solving and insight generation is the central feature that distinguishes action research from other research methods (which may not involve problem-solving) and consulting (which may not involve insight generation). Hence, action research is an excellent method for bridging research and practice.

There are several variations of the action research method. The most popular of these methods is participatory action research, designed by Evered (Susman & Evered, 1978). This method follows an action research cycle consisting of five phases (see Figure 83).

1. **Diagnosing**. Identify and define a problem in its social context.
2. **Action planning**. Identify and evaluate alternative solutions to the problem and decide on a future course of action (based on theoretical rationale).
3. **Action taking**. Implementation of the planned course of action.
4. **Evaluating**. Examining the extent to which the initiated action successfully resolves the original problem, i. e., whether theorized effects are indeed realized in practice.
5. **Learning.** Use the experiences and feedback to generate insights about the problem and suggest future modifications or improvements to the action.

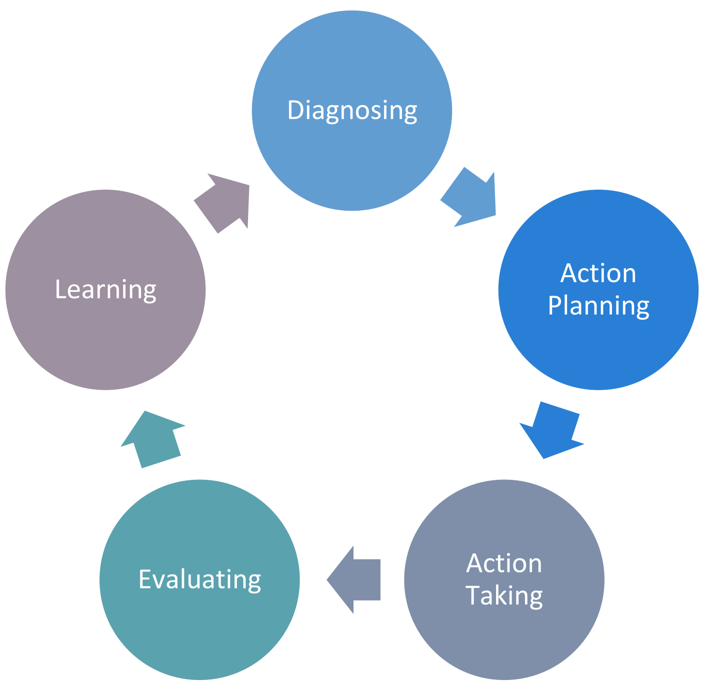


FIGURE 83: ACTION RESEARCH CYCLE

Based on action evaluation and learning, the action may be modified or adjusted to address the problem better, and the action research cycle is repeated with the modified action sequence. It is suggested that the entire action research cycle be traversed at least twice so that learning from the first cycle can be implemented in the second cycle. The primary mode of data collection is participant observation, although other techniques such as interviews and documentary evidence may corroborate the researcher's observations.

Ethnography

The ethnographic research method, mainly derived from anthropology, emphasizes studying a phenomenon within the context of its culture. Researchers must be deeply immersed in the social culture over an extended period, usually eight months to two years. They should engage, observe, and record the daily life of the studied culture and its social participants within their natural setting. The primary mode of data collection is participant observation, and data analysis involves a "sense-making" approach. In addition, researchers must take extensive field notes and narrate their experiences in descriptive detail so that readers may experience the same culture. In this method, researchers have two roles: rely on their unique knowledge and engagement to generate insights (theory), and convince the scientific community of the trans-situational nature of the studied phenomenon.

The classic example of ethnographic research is Jane Goodall's (pictured on the right) study of primate behaviors. She lived with chimpanzees in their natural habitat at Gombe National Park in Tanzania, observed their behaviors, interacted with them, and shared their lives. During that process, she chronicled how chimpanzees seek food and shelter, how they socialize with each other, their communication patterns, their mating behaviors, and other activities. A more contemporary example of ethnographic research is Myra Bluebond-Langer's (Bluebond-Langner, 2000) study of decision making in families with children suffering from life-threatening illnesses. She chronicled the physical, psychological, environmental, ethical, legal, and cultural issues that influence the familys' decision-making. The researcher followed the experiences of approximately 80 children with incurable illnesses and their families for over two years. Data collection involved participant observation and formal/informal conversations with children, their parents and relatives, and health care providers to document their lived experience.

Phenomenology

Phenomenology is a research method that emphasizes the study of conscious experiences to understand the reality around us. It is based on the ideas of German philosopher Edmund Husserl in the early 20th century, who believed that human experience is the source of all knowledge. Phenomenology is concerned with the systematic reflection and analysis of phenomena associated with conscious experiences, such as human judgment, perceptions, and actions, with the goal of 1) appreciating and describing social reality from the diverse subjective perspectives of the participants involved and 2) understanding the symbolic meanings ("deep structure") underlying these subjective experiences. Phenomenological inquiry requires that researchers eliminate any prior assumptions and personal biases, empathize with the participant's situation, and tune into existential dimensions of that situation to fully understand the deep structures that drive conscious thinking and behavior of the studied participants.

Some researchers view phenomenology as a philosophy rather than as a research method. In response to this criticism, Giorgi and Giorgi (Giorgi & Giorgi, 2003) developed an existential-phenomenological research method to guide studies in this area. This method can be grouped into data collection and data analysis phases.

- Data Collection
    - Interview participants regarding the phenomenon of interest
    - Transcribe interviews for analysis
- Data Analysis
    - Read transcripts to get a sense of the whole
    - Identify units of significance (establish parts)
    - Assign values to units of significance by reliving participants' subjective experience
    - Develop themes to tie together the units of significance into layered meanings
    - Identify and reconcile "deep structure."

In the data collection phase, participants in a social phenomenon are interviewed to capture their subjective experiences and perspectives regarding the phenomenon under investigation. Examples of questions that may be asked include "can you describe a typical day" or "can you describe that particular incident in more detail?" These interviews are recorded and transcribed for further analysis.

During data analysis, researchers read the transcripts to 1) get a sense of the whole and 2) establish "units of significance" that can faithfully represent participants' subjective experiences. Examples of such units of significance are concepts such as "felt-space" and "felt-time," which are then used to document participants' psychological experiences. For instance, did participants feel safe, independent, trapped, or joyous when experiencing a phenomenon ("felt-space")? Did they feel that their experience was pressured, slow, or discontinuous ("felt-time")? Phenomenological analysis should consider the participants' temporal landscape (i.e., their sense of the past, present, and future). Researchers must transpose themselves in an imaginary sense in the participant's situation (i.e., temporarily live the participant's life). The participants' lived experience is described in the form of a narrative or using emergent themes. The analysis then delves into these themes to identify multiple layers of meaning while retaining the fragility and ambiguity of subjects' lived experiences.

Rigor In Interpretive Research

While positivist research employs a "reductionist" approach by simplifying social reality into theories and laws, interpretive research attempts to interpret social reality through the subjective viewpoints of the embedded participants within the context where the reality is situated. These interpretations are heavily contextualized and are naturally less generalizable to other contexts. However, because the interpretive analysis is subjective, it is often considered less rigorous by many researchers who embrace functionalism. Interpretive research is based on a set of philosophical assumptions that are unlike those of functionalism. Therefore, traditional notions of rigor, such as reliability, internal validity, and generalizability, do not apply as they would in a positivist study. However, Lincoln (Lincoln, 1985) provides an alternative set of criteria that can be used to judge the rigor of interpretive research.

- **Dependability.** Interpretive research can be viewed as dependable or authentic if two researchers assessing the same phenomenon using the same set of evidence independently arrive at the same conclusions or the same researcher observing the same or a similar phenomenon at different times arrives at similar conclusions. This concept is similar to that of reliability in positivist research. The agreement between two independent researchers is similar to the notion of inter-rater reliability and agreement between two observations of the same phenomenon by the same researcher akin to test-retest reliability. To ensure dependability, interpretive researchers must provide adequate details about their phenomenon of interest and the social context in which it is embedded to allow readers to authenticate their interpretive inferences independently.

- **Credibility**. Interpretive research can be considered credible if readers find its inferences to be believable. This concept is akin to that of internal validity in functionalistic research. The credibility of interpretive research can be improved by providing evidence of the researcher's extended engagement in the field and demonstrating data triangulation across subjects or data collection techniques. Credibility is enhanced by maintaining meticulous data management and analytic procedures,

such as verbatim transcription of interviews, accurate records of contacts and interviews, and clear notes on theoretical and methodological decisions, allowing an independent audit of data collection and analysis.

- **Confirmability**. Confirmability refers to the extent to which the findings reported in interpretive research can be independently confirmed by others (typically, participants). Confirmability is similar to the notion of objectivity in functionalistic research. Since interpretive research rejects the notion of an objective reality, confirmability is demonstrated in terms of "inter-subjectivity," i. e., if the study's participants agree with the inferences derived by the researcher. For instance, if a study's participants generally agree with the inferences drawn by a researcher about a phenomenon of interest (based on a review of the research paper or report), then the findings can be viewed as confirmable.

- **Transferability**. Transferability in interpretive research refers to the extent to which the findings can be generalized to other settings. This idea is similar to that of external validity in functionalistic research. The researcher must provide rich, detailed descriptions of the research context ("thick description") and thoroughly describe the structures, assumptions, and processes revealed from the data so that readers can independently assess whether and to what extent are the reported findings transferable to other settings.

Key Takeaways

Chapter 13: Interpretive Research
• Define "Interpretive Research." • Discuss data collection techniques for interpretive research. • Define the various types of interpretive research: case research, action research, and ethnography. • Discuss rigor in interpretive research.

Mixed Methods

All quantitative and qualitative research methods have specific strengths and weaknesses. Mixed methods attempt to use more than one research method on a given project to utilize the strengths of each method while mitigating their weaknesses.

14: Mixed Methods

Often, researchers are not comfortable taking a single approach to a question and suspect that there is more to be discovered by looking at the question from multiple perspectives. In these cases, researchers may choose to use a second or even third approach to the research design to triangulate a more satisfactory explanation for the question. Like two or paths merging into one, researchers can merge several methods into a single project.

OBJECTIVES

- Describe the strengths and weaknesses of quantitative research methods.
- Describe descriptive and inferential techniques.
- Describe the strengths and weaknesses of qualitative research methods.
- Define grounded theory and describe how grounded theory is developed.
- Compare and contrast quantitative and qualitative methods.
- Define mixed methods.
- Describe sequential explanatory, sequential exploratory, and convergent parallel research methods.

Introduction

There are two ways to approach a project: quantitative research and qualitative research. Quantitative research projects gather numeric data and analyze those data with statistical tools. Qualitative research projects gather non-numeric data and analyze those data with non-mathematical tools. However, combining both types of analysis in a single research project is possible, a mixed method. This chapter first briefly revisits quantitative and qualitative methods and then considers the process used to combine those methods.

Figure 84 is a summary of the quantitative and qualitative research paradigms.

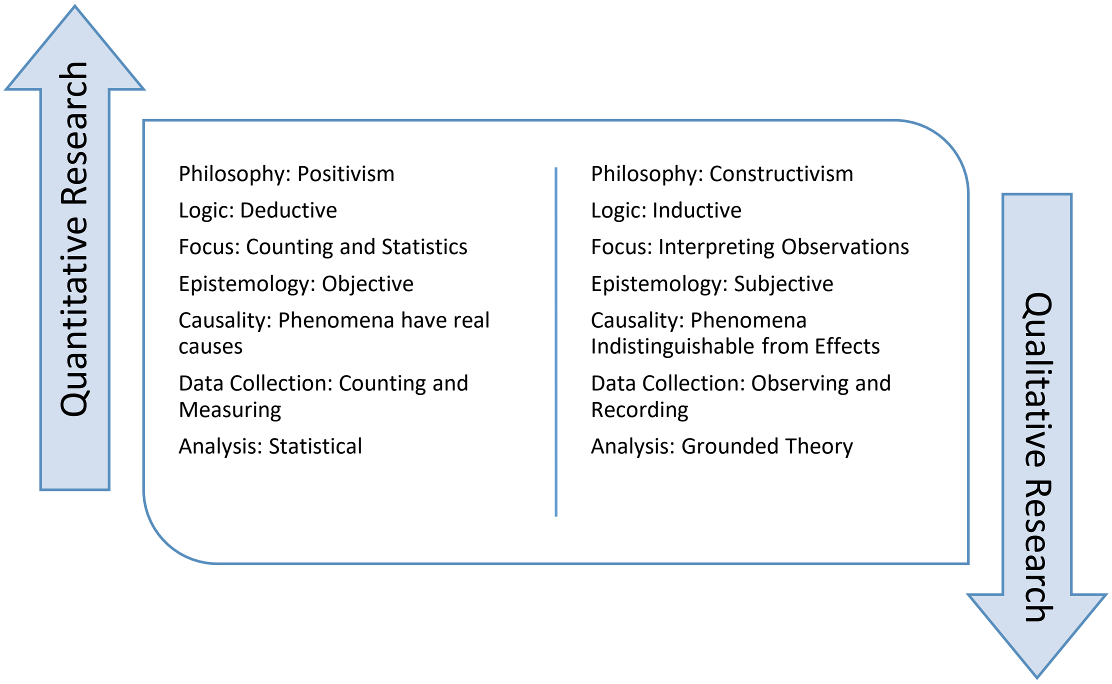


FIGURE 84: QUANTITATIVE VS. QUALITATIVE METHODS

Quantitative Analysis

Numeric data collected in a research project can be analyzed quantitatively using statistical tools in two different ways.[33]

- Descriptive analysis refers to statistically describing, aggregating, and correlating variables.
- Inferential analysis refers to the statistical testing of hypotheses.

Descriptive

Univariate analysis, or analysis of a single variable, refers to a set of statistical techniques that can describe the general properties of one variable. Univariate statistics include 1) frequency distribution, 2) central tendency, and 3) dispersion.

Frequency distribution. The frequency distribution of a variable summarizes the frequency that individual values are found in a variable. For instance, it is easy to measure how often customers in a grocery store purchase types of products, like “produce,” “dairy,” and “meat.” If the number (or percentage) of observations within each category are counted and displayed in a table, it would be called a frequency distribution, as seen in Figure 85. A frequency distribution can also be depicted as a

[33] Chapter 6 more thoroughly describes the statistical techniques summarized in this section.

bar chart, as shown in Figure 86. The horizontal axis represents the number of purchases in each category, and the vertical axis representing the categories.

Item	Number
Produce	374
Dairy	291
Meat	187

FIGURE 85: SAMPLE FREQUENCY TABLE

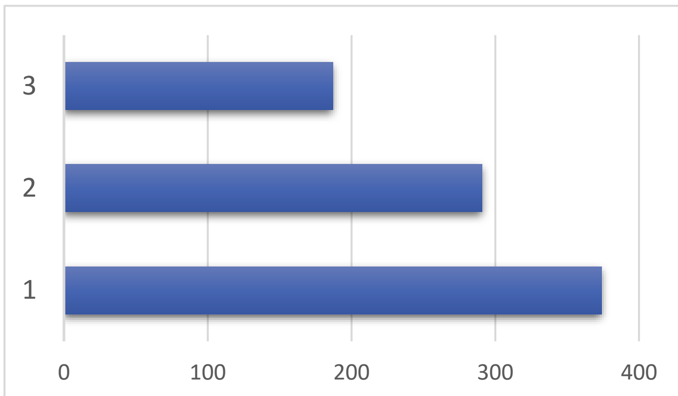


FIGURE 86: SAMPLE BAR CHART

Central tendency. It is often desirable to find the center of a data set, and there are three commonly-used calculations: mean, median, and mode.

- The arithmetic mean (often called the "mean") is the "average" of all values in a given distribution. This measure is the calculation most school children are taught. It is found by adding all of the values together and dividing by the number of values.
- The median is the middle value in a distribution. This measure is computed by placing all of the values in order and selecting the one in the middle. If there are two middle values (if there is an even number of values), then the median is the mean of the two middle values.
- The mode is the most frequently occurring value in a distribution of values. Mode is usually only used for categorical data rather than numeric. For example, if an item on a survey asked whether respondents rented, leased, or owned their office building, it would not make sense to find an "average" for those values; instead, the most frequently selected response would be reported as the mode.

Dispersion. This term refers to the way values are spread around the center of the data. Two common measures of dispersion are the range and standard deviation.

- The range is the difference between the highest and lowest values in a distribution. The range is sensitive to the presence of outliers, which makes its use problematic. For instance, imagine that the values of the houses in a neighborhood were listed, and they were all between $100K and $200K except for one large house worth $300K. That one outlier value would make the range much more extensive than expected.
- The standard deviation is calculated by finding each value's distance from the mean of the data set (its deviation) then finding the mean of all of those deviations. While the calculation is somewhat complex, all statistical software packages can calculate the standard deviation. Because of how it is calculated, the standard deviation is not sensitive to outliers, so it is often used to indicate the data dispersion.

Bivariate analysis, or analysis of two variables, refers to statistical techniques that describe the relationship between two variables. The most common bivariate statistic is a correlation, which is a number between −1.00 and +1.00, denoting the strength and direction of the relationship between two variables. For example, consider a data set containing selected specifications found in the 1974 Motor Trend magazine for 32 automobiles (1973–74 models)[34]. The first few items in that data set are shown in Figure 87.

Name	Mpg	Cyl	Disp	Hp	Wt	Qsec
Mazda RX4	21.0	6	160	110	2.620	16.46
Datsun 710	22.8	4	108	93	2.320	18.61
Hornet 4 Drive	21.4	6	258	110	3.215	19.44
Valiant	18.1	6	225	105	2.460	20.22

FIGURE 87: SAMPLE OF MOTOR TREND CAR DATA

Two of the variables in this data set are disp (engine displacement) and qsec (quarter-mile time, in seconds). It would be normal to expect that the greater the engine displacement (that is, the larger the engine), the faster the automobile would finish a quarter-mile track. It turns out that the correlation between displacement and quarter-mile time is −0.43. The negative sign indicates that the relationship is negative; as the displacement increases, the time decreases. The magnitude of the correlation indicates that this is not a particularly strong relationship so that some small engines can finish the

[34] The Motor Trend data was originally published in a report by Henderson and Velleman in *Biometrics* (Henderson & Velleman, 1981).

quarter-mile track as quickly as larger engines. A researcher would want to know why and one of the first confounding factors to consider would be the weight of the automobile. Are automobiles with larger engines also heavier and, thus, slower through the quarter-mile?

Researchers can use a powerful visual aid to compare the correlations between numerous variables in a single data set using a correlation plot. Figure 88 shows a correlation plot for each of the Motor Trend variables listed in Figure 87.

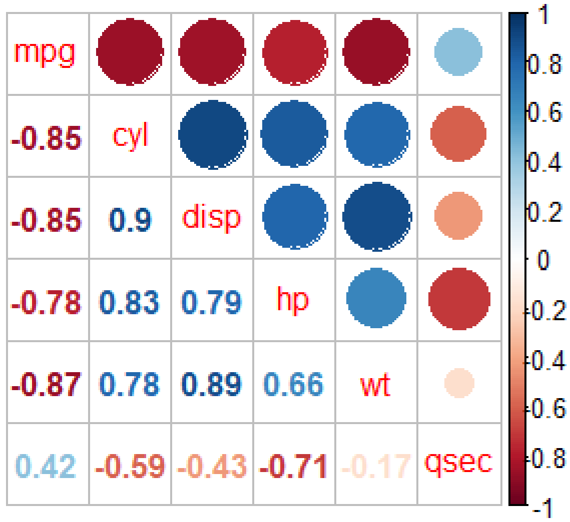


FIGURE 88: CORRELATION PLOT

In the correlation plot, the names of the variables are listed diagonally from the top left to the bottom right. In the lower half of the plot, the correlations are reported numerically. Thus, the correlation between mpg and cyl is −0.85. Those correlations are also color-coded using the scale found on the right side of the plot. Since −0.85 is a relatively strong negative correlation, it is printed in a dark red color. The top half of the plot shows the correlations using circles where both color and size indicate the strength and direction of the correlation. Thus, the correlation between wt and qsec is very weak since the size of the circle is small, and it is negative since the color is pale pink. Researchers can quickly locate strong positive and negative correlations using a correlation plot, like mpg and wt (strong negative) or disp and wt (strongly positive).

Another helpful tool is a scatter plot. Consider Figure 89, which shows the relationship between the waiting time and eruption time for the Old Faithful geyser in Yellowstone Park[35]. The plot clearly shows

[35] These data were first published by Härdle in *Smoothing Techniques with Implementation in S* (Härdle, 2012).

that the longer people have to wait for an eruption (the time on the X-Axis increases), the longer the eruption will last (the time on the Y-Axis increases). The scatter plot also shows two clear groups of points, so it would be reasonable to conclude that there are "short" eruptions and "long" eruptions.

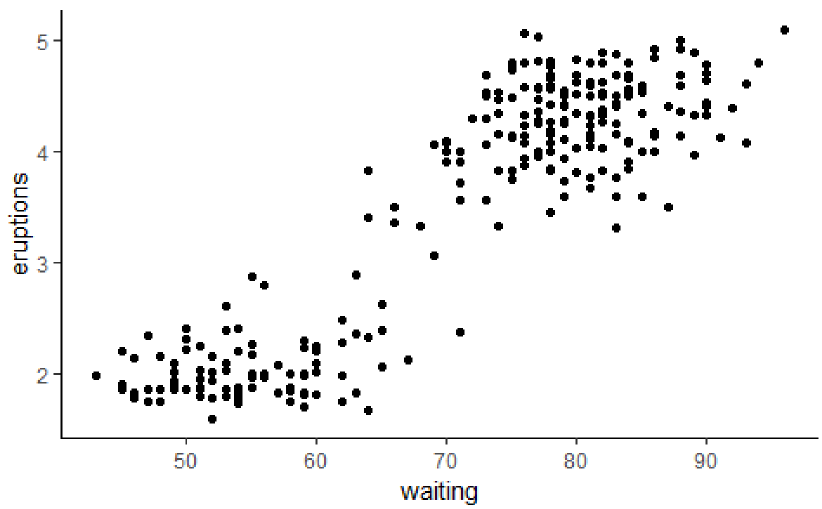


FIGURE 89: SCATTER PLOT

Inferential Quantitative Analysis

Inferential statistics are procedures used to reach conclusions about the associations between variables. They differ from descriptive statistics in that they are explicitly designed to test a hypothesis. Numerous statistical procedures fall into this category, but all are supported by statistical software like *R*. This chapter provides only a short primer on the most basic and frequent inferential procedures. First, though, it is essential to define the concept of hypothesis.

A hypothesis is a proposition put forth to explain some observed phenomena. Often, a hypothesis also functions in a predictive manner and can be tested by scientific methods. A hypothesis typically has these characteristics:

- **Clear**. The hypothesis must be stated in clear and precise language.
- **Testable**. A hypothesis must be testable and have a way to determine if the hypothesis is true or false.
- **Consistent**. A hypothesis should be consistent with known facts or an established body of literature.
- **Timely**. A hypothesis must be able to be confirmed (or rejected) in a reasonable time frame.

Research projects start when a researcher comes up with a notion about a topic of interest. For example, a researcher might assume that ads in a local newspaper are more effective than those on a local radio station. This assumption would generate a research project to test the effectiveness of ads in both media. The project would start with a *Null Hypothesis*, which states that the initial assumption is wrong; the status quo is correct. Then, the researcher would construct an *Alternative Hypothesis* to support the assumption. For example, the researcher may propose these null and alternative hypotheses.

- **Null Hypothesis**. The type of media does not change the effectiveness of an ad.
- **Alternative Hypothesis**. Ads placed in a local newspaper are more effective than those placed on a local radio station.

It is impossible to prove an alternative hypothesis definitively since it is impossible to analyze all potentially relevant data. In the case of the advertising media mentioned above, it would not be possible to prove the alternative hypothesis for every possible combination of newspaper and radio ads, in all possible markets, in all possible seasons, for all possible products. Typically, the veracity of the alternative hypothesis is determined by rejecting (or not) the null hypothesis. For example, if the researcher ran ads in a newspaper and on the radio and the newspaper ads seemed to generate more business, would that verify the hypothesis? Is the result nothing more than chance?

Statistically, the probability that a conclusion is caused by chance is called the *p-value* (for "probability value"). Researchers determine the maximum acceptable probability that the project's conclusion is chance at the project's onset. Most business and marketing research uses a *p-value* of 0.05 (or 5%) as the cutoff. Thus, a calculated *p-value* less than 0.05 indicates that the observed results are not just chance, and the null hypothesis can be rejected. However, if the *p-value* exceeds 0.05, then the null hypothesis cannot be rejected. All statistics programs, like *R*, automatically calculate a *p-value* as part of the output for many statistical tests.

One of the most straightforward inferential analyses is comparing treatment and control groups. For example, determining whether students enrolled in an "enhanced" mathematics program perform better than those in a traditional program. In this case, the measured outcome variable could be a standardized test score following each mathematics course. The analytic technique for comparing these types of scores is a t-test[36].

The t-test examines whether the means of two groups are statistically different from each other (non-directional or two-tailed test) or whether one group has a greater or lesser mean than the other (directional or one-tailed test). In the mathematics example, if the goal is to examine whether students in the enhanced mathematics program perform better than those in a traditional program, it would be a one-tailed test. The hypothesis can be stated as:

- **Null**. The enhanced program scores are not greater than those from the traditional program.
- **Alternate**. The enhanced program scores are greater than those from the traditional program.

Note that a statistical significance test aims to reject the null hypothesis—in other words, to show that there is a significant difference between the two groups being compared. If a t-test is conducted on the outcome scores for the two mathematics programs, it will produce a *p-value* (among other outcomes). If that value is less than 0.05, researchers can assume that the means are genuinely different and reject the null hypothesis.

Extending the mathematics program example, imagine that the hypothesis is changed a bit such that the amount of instructional time, either three or six hours/week, is also considered. This change creates a 2x2 factorial design, with the two factors being program type (enhanced vs. traditional) and instructional time (three vs. six hours/week). This type of design helps researchers estimate the independent effect of each factor, called main effects, and the combined effect of both factors; the interaction effect. This type of factorial design can be analyzed using a two-way ANOVA, but, again, researchers would look for a *p-value* less than 0.05 to indicate which factors were significant.

A few other statistical tests of quantitative data are commonly found in the literature.

- **Factor analysis** is a data reduction technique used to statistically aggregate a large number of observed measures (items) into a smaller set of unobserved (latent) variables called factors

[36] The t-test was introduced in 1908 by William Sealy Gosset, a chemist working for the Guiness Brewery in Dublin, Ireland, to monitor the quality of stout — a dark beer popular with nineteenth-century porters in London. Because his employer did not want to reveal the fact that it was using statistics for quality control, Gosset published the test in Biometrika using his pen name “Student.” The test involved calculating the value of “t,” which was a letter used frequently to denote the difference between two groups. Hence, the name “Student’s t-test.”

based on their underlying bi-variate correlation patterns. For example, perhaps a researcher could aggregate income, home value, and tax bracket into a single factor named "wealth."

- **Discriminant analysis** is a classification technique that places a given observation in nominal categories based on a linear combination of predictor variables. It is popular in marketing applications, such as classifying customers or products based on salient attributes identified from large-scale surveys.
- **Logistic regression** is a model in which the outcome variable is binary (zero or one) and is presumed to follow a logistic distribution. The regression analysis aims to predict the probability of the successful outcome by fitting data into a logistic curve. An example is predicting the probability of heart attack within a specific period, based on predictors such as age, body mass index, and exercise regimen. Logistic regression is popular in the medical sciences.
- **Probit regression** is a model in which the outcome variable can have only two values, like "yes-no" on a credit application. The goal of the regression is to predict the probability of each outcome. This type of analysis is a widespread technique in actuarial science, financial services, insurance, and other industries for applications such as creditworthiness based on a person's credit rating, salary, debt, and other information from a loan application.
- **Path analysis** is a technique for analyzing directional relationships among a set of variables. It allows for examining complex models where the dependent variable in one equation is the independent variable in another equation and is widely used in contemporary social science research. For example, perhaps a project is designed to show that they are happier because their disposable income increases as people age. One study links disposable income to age. The study results are used in a follow-on study to determine a relationship between income and happiness.
- **Time series analysis** is a technique for analyzing time-series data or variables that continually change with time. Examples of applications include forecasting stock market fluctuations and urban crime rates. This technique is popular in econometrics, mathematical finance, and signal processing.

Qualitative Analysis

Qualitative analysis is the analysis of qualitative data such as text data from interview transcripts. Unlike quantitative analysis, which is statistics-driven and largely independent of the researcher, qualitative analysis heavily depends on the researcher's analytic skills and knowledge of the social context from where the data are collected. The qualitative analysis emphasizes "sense-making" or understanding a phenomenon rather than predicting or explaining. A creative and investigative mindset is needed for qualitative analysis along with a set of analytic strategies.

Grounded Theory

How is a vast set of qualitative data acquired through participant observation, in-depth interviews, focus groups, or other narratives analyzed? One technique is grounded theory, a technique of interpreting data about a social phenomenon to build theories. The technique was developed by Glaser and Strauss (Glaser & Strauss, 1967) in their method of constant comparative analysis of grounded theory research. The technique was refined by Corbin and Strauss (Corbin & Strauss, 1990) to illustrate specific coding techniques—a process of classifying and categorizing text data segments into a set of codes, categories, and relationships. The interpretations are "grounded in" (or based on) observed empirical data, hence the name. Strauss and Corbin (Strauss & Corbin, Basics of qualitative research: Procedures and techniques for developing grounded theory, 1998) describe the following three coding techniques for analyzing text.

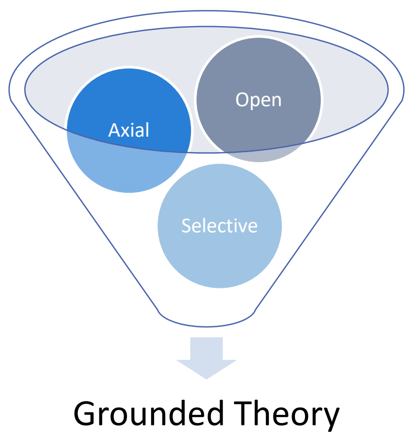


FIGURE 90: CODING TECHNIQUES FOR GROUNDED THEORY

- **Open coding** is designed to identify key ideas in the textual data related to the phenomenon of interest. The researcher examines the raw textual data line by line to identify events, incidents, ideas, actions, perceptions, and interactions of relevance, which are coded as "concepts." Each concept is linked to specific portions of the text (coding unit) for later validation. Some concepts may be simple, clear, and unambiguous, while others may be complex, ambiguous, and viewed differently by different participants. The coding unit may vary with the concepts being extracted. Simple concepts such as "organizational size" may include just a few words of text, while complex ones such as “organizational mission” may span several pages. Concepts can be named using the researcher’s naming convention or standardized labels taken from the research literature. Once a basic set of concepts are identified, they can code the remainder of the data while looking for new concepts. While coding, it is essential to identify the recognizable characteristics of each concept, such as its size, color, or level (e. g., high or low), so that similar concepts can later be grouped.

- **Axial coding** is used to group concepts into higher-order categories. While concepts may be context-specific, categories tend to be broad and generalizable and ultimately evolve into constructs in a grounded theory. Categories are needed to reduce the number of concepts the researcher must work with and build a “big picture” of the issues salient to understand the phenomenon under consideration. The existing literature can name these categories, mainly if the research aims to extend current theories. The characteristics (properties) and dimensions of each category should be identified. The dimension is the value of one of the characteristics along a continuum. For example, a “communication media” category may have a characteristic called “speed,” which can be dimensionalized as fast, medium, or slow. The relationships between categories may be evident or subtle. In the latter instance, researchers may use a coding scheme to understand which categories represent conditions (the circumstances in which the phenomenon is embedded), actions/interactions (the responses of individuals to events under these conditions), and consequences (the outcomes of actions/ interactions). As conditions, actions/interactions, and consequences are identified, theoretical propositions start to emerge, and researchers can start explaining why a phenomenon occurs, under what conditions, and with what consequences.
- **Selective coding** involves identifying a significant category or core variable systematically, then logically relating this main category to other categories. The main category can evolve from existing categories or a higher-order category that subsumes previously coded categories. New data are selectively sampled to validate the main category and its relationships to other categories (i. e., the tentative theory). Selective coding limits the range of analysis and makes it move faster than open and axial coding.

During the coding process, the coder is alert for categories that may emerge from new data related to the phenomenon of interest. Hence, open, axial, and selective coding often proceed simultaneously. Coding of new data and theory refinement continues until theoretical saturation is reached, i. e. when additional data does not yield more than marginal change in the core categories or the relationships.

The “constant comparison” process implies continuous rearrangement, aggregation, and refinement of categories, relationships, and interpretations based on the increasing depth of understanding and an iterative interplay of the following four stages of activities:

1. Comparing incidents/texts assigned to each category (to validate the category)
2. Integrating categories and their properties
3. Delimiting the theory (focusing on the core concepts and ignoring less relevant concepts)
4. Writing theory.

After a grounded theory is generated, it must be refined for internal consistency and logic. Researchers must ensure that the central construct has the stated characteristics and dimensions, and if not, the data analysis may be repeated. Researchers must then ensure that the characteristics and dimensions of all categories show variation. For example, if behavior frequency is one such category, then the data must provide evidence of frequent performers and infrequent performers of the focal behavior. Finally, the theory must be validated by comparing it with raw data. If the theory contradicts empirical evidence, then the coding process may be repeated to reconcile such contradictions or unexplained variations.

Quantitative Vs. Qualitative

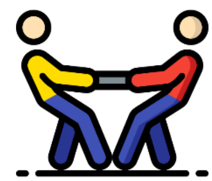
Given their differences, it may come as no surprise that quantitative and qualitative research do not coexist in complete harmony. Some quantitative researchers criticize qualitative methods because they lack objectivity, are challenging to evaluate in terms of reliability and validity, and do not allow generalization to people or situations other than those studied. At the same time, some qualitative researchers criticize quantitative methods because they overlook the richness of human behavior and experience and instead answer simple questions about easily quantifiable variables.

In general, however, qualitative researchers are well aware of objectivity, reliability, validity, and generalizability. They have developed several frameworks for addressing these issues. Also, in general, quantitative researchers are well aware of the issue of oversimplification. They do not believe that all human behavior and experience can be adequately described in terms of a small number of variables and their statistical relationships. Instead, they use simplification as a strategy for uncovering general principles of human behavior.

Combining Quantitative And Qualitative

Quantitative research based on positivism has historically been the cornerstone of business and marketing research. Purists call for researchers to "eliminate their biases, remain emotionally detached and uninvolved with the objects of study and test or empirically justify their stated hypotheses" (Baker, 2006).

Qualitative research is based on interpretivism and its practitioners, "contend that multiple-constructed realities abound, that time- and context-free generalizations are neither desirable nor possible, that research is value-bound, that it is impossible to differentiate fully causes and effects, that logic flows from specific to general and that knower and known cannot be separated because the subjective knower is the only source of reality" (Johnson & Onwuegbuzie, 2004).



In the 1980s and 1990s, researchers began to call for a "truce" in the "Paradigm Wars" between quantitative and qualitative methods. Many prominent authors and researchers felt that the two research methodologies are compatible; they may be combined in a single research project. The proverbial pendulum has swung the other direction, and many researchers now believe that there is no central problem area that should be studied exclusively with one research method. They believe that quantitative research answers the "if" question while qualitative research answers the "how or why." This research paradigm is a mixed-methods approach, though other phrases, like "multi-modal," are occasionally used.

Mixed-method research offers powerful tools to investigate complex systems and processes in business, marketing, and economics. This method includes all phases of a research project, including philosophical assumptions, research questions, design, collection, analysis, integration, and presentation of data and results[37].

The nature of the research question guides the selection of the method. Researchers in business-related fields use a quantitative methodology to study and answer research questions on causality, generalization, and magnitude of effect. The qualitative methodology is the choice of researchers who seek to answer research questions that explore how or why a given phenomenon occurs, develop a theory, or describe the subjectivity of an individual experience.

Mixed-method research attempts to utilize the strengths of each of the two approaches, quantitative and qualitative, and, for this reason, it is increasingly used to address contemporary research problems.

---

[37] The material in this section is adapted from *Mixed-Method Research in the Health Sciences* (Lorenzini, 2017).

An indication of the increased interest in this method was the publication of guidelines on mixed-methods research in various fields, like information systems (Venkatesh, Brown, & Bala, 2013) and health sciences (Creswell, Fetters, & Ivankova, Designing a mixed methods study in primary care, 2004).

Over the years, several definitions of mixed methods have emerged, but researchers seem to be focused on defining mixed methods by their characteristics, which are as follows.

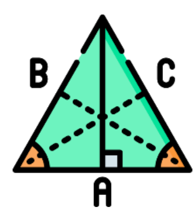
- The collection and analysis of both quantitative and qualitative data take place.
- Rigorous procedures are used to carry out quantitative and qualitative research.
- There is integration or a combination of results.
- Procedures are developed in which data collection, analysis, and integration take place.

It is often argued that the quantitative approach cannot capture what is understood about the context where the study took place, but qualitative research compensates for these weaknesses. On the other hand, qualitative research is seen as deficient for several reasons:

- The personal interpretations made by the researcher;
- The bias that is created from the researcher's interpretations;
- The small number of participants;
- The difficulty of generalizing the results.

Using mixed methods, researchers can use all available tools from both methods rather than confining themselves to strategies associated with either quantitative or qualitative research.

Triangulation



One of the most commonly-used mixed-methods tools is called "triangulation." This term comes from old navigation methods where ships' captains could pinpoint their location by "shooting an azimuth" to two different visible points and then drawing intersecting lines on a map. In mixed methods research, triangulation is the process of using multiple sources to pinpoint a solution. Tashakkori (Tashakkori, Teddlie, & Teddlie, 1998) discusses four types of triangulation: data triangulation (the use of a variety of data sources in a study), investigator triangulation (the use of several different researchers), theory triangulation (the use of multiple perspectives to interpret the results of a study), and methodological triangulation (the use of multiple methods to study a research problem). Triangulation was the solution that broke the "Paradigm Wars" and permitted mixed methods research to begin to be accepted by scholars.

Steps For Design

While every research project is different, Venkatesh (Venkatesh, Brown, & Bala, 2013) created a six-step process for mixed methods research. While this process may need to be modified for specific projects, it is an excellent general-purpose start.

1. **Decide on the appropriateness of a mixed-methods approach.** Not every research project will benefit from a mixed-methods approach. Consider things like the research question, purpose of the research project, and the selection of available paradigms. Keep in mind that a mixed-methods project is more time and resource-intensive and could be too much for a simple project.
2. **Develop strategies for the design.** There are three overall mixed method strategies (covered below), and once the decision is made to use a mixed-methods procedure, it is time to determine the strategy to use. This decision is informed by what approaches of the project (qualitative or quantitative) will be completed first, the priority of approaches, how the data and results will be mixed, and the amount of time and expertise available.
3. **Develop strategies for collecting and analyzing data.** Because the project will be completed in two phases (qualitative and quantitative), it is essential to plan to collect and analyze data. Mistakes made in the first collection phase may not become evident until the second phase, which could sink the entire project.
4. **Draw meta-inferences from the results.** This step is similar to developing a theory that explains the data and analysis. The information about grounded theory (above) may help with this step of the process.
5. **Assess the quality of the meta-inferences.** The meta-inferences may be strong or weak. If they are weak, they need to be reconsidered, or maybe the entire research project needs to be redone with appropriate changes to the protocol.
6. **Discuss potential threats and remedies.** All research projects develop threats, and the researcher needs to approach those threats and propose remedies for them honestly.

Strategies

Mixed-methods projects typically fall into three general strategies, listed below and more fully described later in this chapter[38].

- Sequential Explanatory
- Sequential Exploratory
- Convergent Parallel (Triangulation)

The specific type of strategy best for any given research project depends on the following four factors.

1. The theoretical perspective of the researcher
    - Explicit—-Does the researcher base the research project directly on a theory?
    - Implicit—-Does the researcher only indirectly use theory as a foundation for the research project?
2. Priority of strategy
    - Qualitative—Is the qualitative portion of the research project more critical?
    - Quantitative—Is the quantitative portion of the research project more critical?
    - Equal—Are the two types of research, quantitative/qualitative, equally used in the research project?
3. The sequence of data collection implementation
    - Is the qualitative part completed first, followed by the quantitative part?
    - Is the quantitative part completed first, followed by the qualitative part?
    - Are both parts completed simultaneously?
4. The point at which the data are integrated
    - At data collection
    - At data analysis
    - At data interpretation
    - With some combination

[38] Tashakkori (Tashakkori, Teddlie, & Teddlie, 1998) identifies eight different strategies while Bell (Bell & Bryman, 2015) lists nine strategies. The three listed in this chapter are from Creswell (Creswell, Research design: qualitative, quantitative, and mixed methods approaches, 2014) and seem to be those most commonly cited in the literature.

SEQUENTIAL EXPLANATORY STRATEGY

The sequential explanatory strategy is used when a researcher already has a theory to explain some phenomena and collects data to explain certain theory facets. Figure 91 illustrates the process of a sequential explanatory strategy.

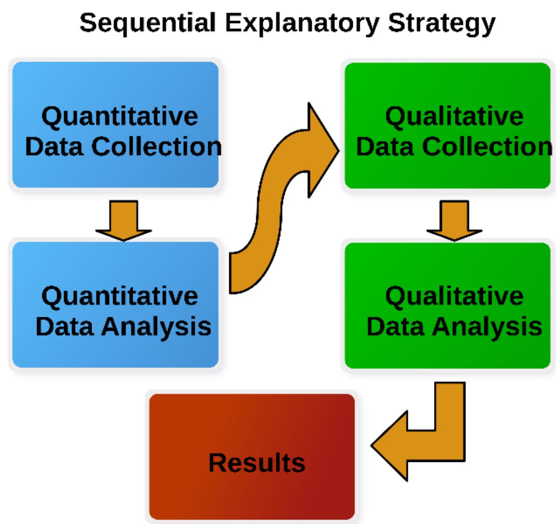


FIGURE 91: SEQUENTIAL EXPLANATORY STRATEGY

The collection and analysis of quantitative data are followed by collecting and analyzing qualitative data where equal priority is given to the two phases. The data are integrated during the results phase. The primary focus of this strategy is to explain quantitative results by exploring those results in more detail or helping explain unexpected results (e.g., using follow-up interviews to understand the results of a quantitative study better).

- **Strength**. This strategy is relatively straightforward due to clear, distinct stages, and it is easier to describe than concurrent strategies.
- **Weakness**. This strategy is time-consuming, especially when both phases are given equal consideration and priority.

SEQUENTIAL EXPLORATORY STRATEGY

The sequential exploratory strategy is used when a researcher seeks to develop a theory related to observations. Figure 92 illustrates the process of a sequential exploratory strategy.

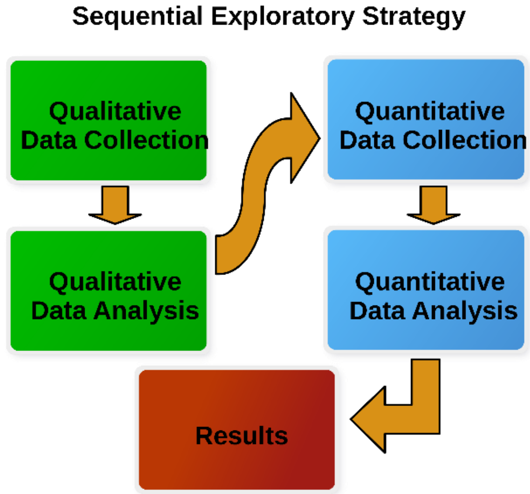


FIGURE 92: SEQUENTIAL EXPLORATORY STRATEGY

The collection and analysis of qualitative data are followed by collecting and analyzing quantitative data where equal priority is given to the two phases. However, priority can be given to either as the project unfolds. Data are integrated during the results phase. This strategy is used primarily to explore a phenomenon by testing the elements of a theory, generalizing qualitative findings, and developing instrumentation.

- **Strength**. This strategy is relatively straightforward due to clear, distinct stages, and it is easier to describe than concurrent strategies.
- **Weakness**. This strategy is time-consuming, especially when both phases are given equal consideration and priority.

CONVERGENT PARALLEL (TRIANGULATION)

The convergent parallel strategy, which is often called "triangulation," is used when a researcher seeks to validate a research project by converging two or more research processes on a single observation. Figure 93 illustrates the process of a convergent parallel strategy.

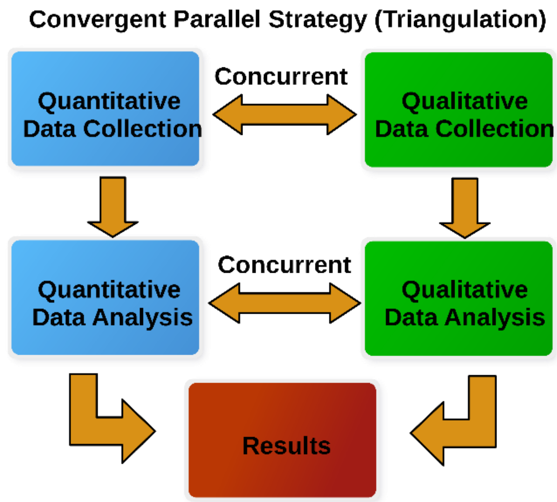


FIGURE 93: CONVERGENT PARALLEL STRATEGY (TRIANGULATION)

There are two concurrent data collection phases where priority should be equal but can be shifted to either approach as necessary. Data are integrated during all phases. The results note any convergence that strengthens knowledge claims or, conversely, a lack of convergence that would disprove the knowledge claims. It is primarily used for confirmation, corroboration, or cross-validation within a single study.

- **Strength**. This strategy is familiar to many researchers. There is a shorter data collection time when compared to sequential methods. It offsets weaknesses inherent to one design by using both.
- **Weakness**. This strategy requires a great deal of expertise and effort to study the phenomenon under consideration using two methods. It may be challenging to compare two types of data as well as resolve discrepancies if they arise.

Mixed-Method Strengths and Weaknesses

Like any research paradigm, mixed methods have both strengths and weaknesses[39].

- Strengths of the mixed-method design:
  - Words, photos, and narratives can be used to add meaning to numbers, while numbers can add precision to words, photos, and narratives
  - It can handle a broader range of research questions because researchers are not limited to one research design

[39] The strengths and weaknesses in this chapter are based on "Demystifying Mixed Methods Research Design: A Review of the Literature" (Caruth, 2013).

- It can present a more robust conclusion
- Offer enhanced validity through triangulation (cross-validation)
- It can add insight and understanding that might be missed when only a single research design is used
- It can increase the capability to generalize the results compared to using only qualitative study designs.

- Weaknesses of the mixed-method design:
  - It can be difficult for a single researcher, mainly when the two designs are best used concurrently; in this case, the study might require a research team
  - It can be more time consuming and expensive when concurrency is involved
  - Requires that the researcher(s) learn multiple methods to combine them knowledgeably, defend the use of multiple methods, and utilize them professionally.
  - Potential for conflict because methodological purists maintain that researchers should work within either a quantitative or a qualitative research design, never mixing the two designs in a single study

Key Takeaways

Chapter 14: Mixed Methods
• Describe the strengths and weaknesses of quantitative research methods. • Describe descriptive and inferential techniques. • Describe the strengths and weaknesses of qualitative research methods. • Define grounded theory and describe how grounded theory is developed. • Compare and contrast quantitative and qualitative methods. • Define mixed methods. • Describe sequential explanatory, sequential exploratory, and convergent parallel research methods.

Reporting

After a research project is completed, the investigator must report the project results, often in writing and orally. This chapter concerns the reporting process.

15: Presenting Research

No conductor would rehearse an entire orchestra on a masterpiece like Beethoven's Ninth Symphony and then present that music to an empty theater. That would be analogous to a researcher working months, sometimes years, on a project that brings a sharp focus to a corner of the researcher's world and then failing to share the results of that research with others. Research is usually presented in both written and oral formats, and, additionally, there are several levels of formality within each format. For example, an oral presentation can be a formal talk, a round table discussion, or a poster session. This chapter concerns presenting research in both formats and includes practical information to help researchers write or present their work.

OBJECTIVES
• Describe what and with whom to share research. • State oral presentation tips. • State written presentation tips. • Describe the process to disseminate findings.

Introduction

Most researchers hope that their work will have relevance to others besides themselves[40]. As such, research is in some ways a public activity. While an individual may conduct the work in a private setting, the knowledge gained from that work should be shared with peers and other parties who may have an interest. Understanding how to share research is an essential aspect of the research process.

What And With Whom To Share

When preparing to share work with others, researchers must decide what to share, with whom to share it, and in what format(s) to share it. This section considers the "what" and "with whom" aspects, while later sections cover the various formats and mechanisms through which research is shared.

Because conducting research is a scholarly pursuit and researchers generally aim to understand business and economics questions, it is crucial that all aspects of research, the good, the bad, and the ugly, are shared. Doing so helps ensure that others will understand, build from, and effectively critique the work.

[40] Some of the material in this chapter was adopted from *Business Communication for Success* (McLean, 2012)

It is also important to share all aspects of a research project for ethical reasons and permit other researchers to replicate it. The following questions will aid researchers in preparing to share research with others.

1. Why was the research conducted?
2. How was the research conducted?
3. For whom was the research conducted?
4. What conclusions can be reasonably drawn from this research?
5. How could the research have been improved?

Answering these questions help researchers be honest with themselves and the readers about their interest, investments, or biases concerning the work. The third question helps identify the major stakeholders, like funders, research participants, or others. This group shares something in common with the research project (e. g., members of the community or social group who were involved in the research). These groups may be interested in the research outcome but may also be a source of bias. The last two questions help identify the strengths and weaknesses of the research project and could point the way for future projects.

While the research project findings would never be altered for different audiences, understanding the audience helps frame the research report meaningfully. For example, the report for a project about the spending habits of elderly pensioners may be much different if rendered for a group of business owners, a governmental committee on aging, the funding agency, and a community meeting. In all cases, researchers would share the significant findings, but the method of presentation and level of detail would vary by audience.

It would be expected that the most detail, including data collection method, sampling, and analytic strategy, would be shared with colleagues and the funding agency. In addition, the funding agency may want information about the exact timeline for the project, along with any bureaucratic hiccups encountered. With a community meeting, though, a more succinct summary of the essential findings using less technical jargon would be appropriate.

Oral Presentations

Settings

Researchers frequently make presentations to their peers in settings like conferences or departmental meetings. These presentations are excellent means for feedback and help researchers prepare to document and publish their work. Presentations take several forms, like:

- formal talks, either as part of a panel at a professional conference or to some other group of peers or other interested parties;
- less formal round-table discussions, which is another standard professional conference format;
- posters displayed in some specially designated areas.

Formal Talk

When preparing a formal talk, researchers need to get details well in advance about the time limit for the presentation, requirements for questions from the audience, and whether visual aids, such as PowerPoint slides, are expected. At conferences, the typical formal talk is usually expected to last between 15 and 20 minutes. Once researchers start talking about something as important as their research, it is common for them to become so engrossed that they forget to watch the clock and then find themselves running short of time. To avoid this all-too-common occurrence, presenters must practice in advance and time those practice sessions.

One common mistake made during formal presentations of research work is setting up the problem the research addresses. Audience members are usually more interested to hear about the researcher's work than to hear the results of a long list of previous related studies. While written reports must discuss related previous studies, presentations must use the precious time available to highlight the current research project.

Another mistake is to read the research paper verbatim. Nothing will bore an audience more quickly than hearing a presenter drone on while reading aloud. Finally, a presentation should highlight only the key points of the study, which, generally, include the research question, methodological approach, significant findings, and a few final takeaways.

Round Table Presentation

In less formal round table presentations, the aim is usually to help stimulate a conversation about a topic. Typically, several research projects are presented, so the time available for each usually is shorter than in a formal presentation. Also, round table presentations always include time for a conversation following the presentations. Round tables are handy when a research project is in the early stages of development. For example, perhaps a researcher has conducted a pilot study and is interested in ideas concerning the next steps. A round table is also an excellent place for a preview of potential objections reviewers may raise regarding the project's approach or conclusions. Finally, round tables are great places to network and meet other scholars who share common interests and are engaged in similar research projects.

Poster Presentation

A poster presentation is a visual representation of a research project. Often, poster sessions are tables lined up in a conference area where researchers have a visual display on the table and stand nearby to answer questions. A poster should not be just pages from the report pasted onto a poster board; instead, researchers tell the "story" of the work in graphs, charts, tables, and other images. Bulleted points are acceptable as long as the people walking past can quickly read and grasp the central argument and findings. Posters, like round tables, can be helpful at the early stages of a research project because they are designed to encourage the audience to engage in conversation about the research. It is unnecessary to share every detail of a research project in a poster; the point is to share highlights and discuss the details with interested people.

Types of Oral Presentations

Most books about oral presentations divide them into several broad types.

- **Speech to Inform**. Increase the audience's knowledge, teach about a topic or issue, and share the speaker's expertise.
- **Speech to Demonstrate**. Show the audience how to use, operate, or do something.
- **Speech to Persuade**. Influence the audience by presenting arguments intended to change attitudes, beliefs, or values.
- **Speech to Entertain**. Amuse the audience in a relatively light-hearted speech that may have a serious point or goal.
- **Ceremonial Speech**. Perform a ritual function, such as give a toast at a wedding reception or a eulogy at a funeral.

Often, presentations are stressful since most people do not like speaking in public. However, the following tips may help.

- **Perfection is not required.** Letting go of perfection can be the most challenging guideline for speakers to apply to themselves. It is human nature to compare ourselves to others. It seems odd, but most people can forgive another researcher for the occasional slip or “umm” during a speech but then turn right around and chastise themselves for making the same error. Everyone has both strengths and weaknesses, and researchers must learn that an essential first step is knowing where to improve. The old saying is that Rome was not built in a day, and good speakers are not developed overnight. No one indeed wants to see a researcher fail during a presentation, so audience members are generally very forgiving of minor speaking faults.
- **Take the time to prepare and get organized.** Researchers know the topic for the speech (usually a research report) and speak to inform or persuade audience members to consider an idea. One of the best ways to build confidence is to know the material being presented “inside out.”
- **Public speaking is not unlike participating in a conversation.** In regular conversations, researchers do not give a second thought to the process of saying something and then waiting for a reply. A public speech follows a similar pattern, but the reply is generally in the form of non-verbal body language.

There are sure to be various obstacles that arise during a presentation.

- **Language**. Researchers work in fields where there are acronyms and insider jargon. As long as a presenter is 100% certain that everyone in the audience understands the terms, it is acceptable to use them, but speakers should defer to standard terms if there is any doubt. As an example, soldiers may understand a sentence like “I left the CHU in the COP and was heading to the DFAC when a fast mover dropped a JDAM in the village.”[41] Only people with military experience would understand that sentence, so it should only be used with that audience.
- **Culture**. Everyone’s culture is different, and a speaker must understand that the audience’s culture may be different enough that communication becomes challenging.
- **Role**. The speaker and audience may play very different roles in an organization, and those roles may create barriers to communication. For example, researchers presenting to a room full of

[41] ”I left the Containerized Housing Unit in the Combat Outpost and was heading to the Dining Facility when a fighter jet dropped a Joint Direct Attack Munition (missle) in the village.”

corporate executives would face a much different communications problem than if the room was filled with line workers. Influential speakers must understand the roles of the audience members and then speak in an understandable way to those people.

- **Goal**. Each of the audience members will have a goal in mind when attending the presentation. Occasionally, the goal may only be that they were assigned to attend the presentation. However, researchers will usually present at a conference where the audience members choose which sessions to attend. As much as possible, the speaker should attempt to understand the audience members' goals and then address those goals.
- **Ethnocentrism**. One obstacle to avoid is for the researcher, or audience members, to feel that they are somehow superior to everyone else. While that can come from ethnicity, it is also a result of scholarly "snobbish" behavior, prejudice, or even stereotypes.

Visual Aids

Nearly all research presentations include some sort of visual aid. It is easier for a speaker to use a graph or chart than verbally describe some data relationship. When preparing visual aids, keep the following in mind.

- **Big**. Visual aids need to be large enough to be seen from the back row of the auditorium. Adjusting the size may be easy for graphs and charts since those visuals can be made more prominent on the screen. However, it may be impossible for a physical artifact, so its utility may be questioned.
- **Clear**. The visual needs to convey whatever message is intended.
- **Simple**. There is an old rule of thumb about visual aids: 6 X 6, which means no more than six lines of text and six words per line.
- **Consistent**. All visuals should use a consistent style so audience members do not have to learn to read each new graphic but can focus instead on the presented information.

Color is a powerful communication tool, but speakers must be careful with color. First, keep in mind that some audience members will not distinguish between two or more colors in the visual aid, so never use color as the sole information source. Also, avoid using too many colors on one chart; a few well-placed colors are always more potent than many colors sprinkled seemingly at random around the visual.

The following tips may help with visual aids.

- Keep visual aids simple.

- Use one key idea per slide.
- Avoid clutter, noise, and overwhelming slides.
- Use large, bold fonts that the audience can read from at least twenty feet from the screen.
- Use contrasting colors to create a dynamic effect.
- Use analogous colors to unify the presentation.
- Use clip art with permission and sparingly.
- Edit and proofread each slide with care and caution.
- Use copies of the visuals as handouts after the presentation.
- Check the presentation room beforehand.
- Have a backup plan in case technology fails, such as providing printed visuals.

Written Presentations

Written reports that other scholars will read generally follow a format required by the publication journal. However, most scholarly reports include an abstract, an introduction, a literature review, a discussion of research methodology, a presentation of findings, and some concluding remarks and discussion about implications of the work. These sections are more fully described later in this chapter. Reports for scholarly consumption also contain a list of references, and many include tables or charts that visually represent some component of the findings. Reading published research in business or economics is an excellent way to understand the core components of scholarly research reports and learn how to write those components.

Reports written for public consumption differ from those written for scholarly consumption. As noted elsewhere in this chapter, knowing the audience is crucial when preparing a written report. Whomever the audience, it is crucial to keep in mind that scientific evidence is being reported. Writers must take their roles as business researchers seriously and be mindful of their place among peers in the discipline. Findings must be presented as clearly and honestly as possible. Give scholars who have come before appropriate recognition, even if current research questions their work. Typically, research writers will never meet the readers face-to-face, but it is beneficial to imagine what readers would ask and provide a detailed response in the written report.

Finally, it is imperative not to commit plagiarism in a research report. Presenting someone else's words or ideas as if they are the researcher's own is among the most egregious transgressions a scholar can commit. Indeed, plagiarism has ended many careers and students' opportunities to pursue degrees (Maurer, Kappe, & Zaka, 2006).

Closely related to plagiarism is libel, the written form of defamation or a false statement that damages a reputation. Suppose false statements that harm a defamed person are published, including publication in a digital or online environment, the author of that statement may be sued for libel. If the person defamed is a public figure, they must prove malice or the intention to harm, but if the victim is a private person, libel applies even if the offense cannot be proven to be malicious.

Writing Style

Writing falls into one of three styles, colloquial, casual, and formal.

COLLOQUIAL STYLE

Colloquial language is an informal, conversational style of writing. It differs from standard business English in that it often uses colorful expressions, slang, and regional phrases. As a result, it can be challenging for an English learner or a person from a different region to understand. Sometimes colloquialism takes the form of a word difference; for example, the difference between a “Coke,” a “tonic,” a “pop,” and a “soda pop” primarily depends on where a person lives. Colloquial phrases can also take the form of a saying. For example, in certain parts of the United States, the phrases “dumb as a box of rocks” and “sharp as a tack” refer to a person’s intelligence, but this may not be obvious for a person not intimate with English.

Colloquial writing may be permissible and even preferable in some business contexts. For example, a marketing letter describing a folksy product such as a wood stove or an old-fashioned corn popper might use a colloquial style to create a feeling of relaxing at home with loved ones. Still, it is essential to consider how colloquial language appears to the audience. Will the meaning of the chosen words be apparent to a reader from a different part of the country? Will a folksy tone sound like the writer is “talking down” to the audience? A final point to remember is that colloquial style is not an excuse for using sexist, racist, profane, or otherwise offensive expressions.

CASUAL STYLE

Casual language style involves everyday words and expressions in a familiar group context, such as conversations with family or close friends. Casual communication is the written equivalent of wearing casual attire, like a t-shirt and jeans. When writing for business, a casual style is usually out of place; instead, a respectful, professional tone represents both the researcher and the company well.

FORMAL STYLE

Formal language is communication that focuses on professional expression with attention to roles, protocol, and appearance. It is characterized by professional vocabulary and syntax. That is, writers using a formal style tend to use a more sophisticated vocabulary, a greater variety of words, and more words with multiple syllables, not to throw big words around, but to enhance the formal mood of the document. Formal writers also tend to use more complex syntax, resulting in longer sentences containing more subordinate clauses.

The appropriate style for a particular business document may be very formal or less so. When an employee replies to a supervisor's email, the exchange may be informal in that it is fluid and relaxed, without much forethought or fanfare. However, it will still reflect the formality of the business environment. The employee is careful to use an informative subject line, a semi-formal salutation ("Hi Mr. Smith" is typical in e-mails), and a brief discussion about the topic at hand. Probably, the employee will also check grammar and spelling before clicking "send."

A formal document such as a proposal or an annual report will involve a great deal of planning and preparation, and its style may not be fluid or relaxed. Instead, it may use distinct language to emphasize the prestige and professionalism of the company. For example, imagine a marketing letter printed on company letterhead and mailed to a hundred sales prospects. Naturally, the letter should represent the company in a positive light and may include a sentence like "The Widget 300 is our premium offering; we have designed it for ease of movement and efficiency of use, with your success foremost in our mind." However, in an e-mail or a tweet, an informal sentence may be used, "W300 is a great stapler."

Report Format

Formal research reports tend to follow a format that has evolved over many years.

Title Page. The report's title is a description of the research findings in a few words. The title page also includes all the researchers' names involved in the project and the institution that sponsored the research project.

Abstract. The abstract is a very brief synopsis of the research findings. The publisher determines the exact size of the abstract, but they generally tend to be about 250 words in length. While the exact content of the abstract is up to the writer, typically, they contain the research problem being investigated, the major hypotheses, the population and sample, a note about the research design, and the significant results. Most researchers wait to write the abstract last and often find that this is one of the most challenging parts of the report. It is tough to distill a research project that may have spanned several years into just 250 words. In the words of Pascal, who wrote in 1657, "I have made this longer than usual because I have not had time to make it shorter."[42]

Table of Contents. A Table of Contents is not typically used, especially for shorter reports. The publisher determines if a Table of Contents is required.

Introduction. The introduction is a short description of the research project, its process, and the anticipated result. The research hypotheses are often included in the introduction. Finally, a few brief details of the methods and results are often found in the introduction.

Literature Review. The literature review usually is one of the more extended parts of the report. It surveys all of the existing related research to position the current project in the universe of prior research. Some literature reviews are structured chronologically, while others are structured thematically. Regardless, there is a clear indication of where the current research project "fits" with existing research.

Methodology. This section is a description of the method used during the research project. It would include the following items:

- Information on the population of interest, how the sampling frame was developed, the method used to select a sample, and the sample itself;
- The study constructs and measures used to operationalize them;
- A description of any instruments used and examples of items from the instruments (typically, the entire instrument used is added as an appendix);
- What type of research design was used, i. e.experimental, survey, qualitative, or other designs;
- The procedure used to design and execute the project along with analyzing the result;
- A description of the data gathered and how those data were analyzed;
- A brief discussion on the reliability and validity of the research project;
- A statement about the generalizability of the study.

This section should be written thoroughly enough that another researcher could duplicate the project if desired.

[42] According to the "Quote Investigator," this is the first known usage of this phrase, though it is used frequently in later literature. Found on September 9, 2019, at https://quoteinvestigator.com/2012/04/28/shorter-letter/#note-3700-1

Results. This section details the research project findings; however, the interpretation of the results is typically saved for the “discussion” section of the report.

Discussion. This section concerns the research project's relevance and how it fits with other research identified in the literature review. This section also details the project’s weaknesses and offers suggestions about how those could have been overcome.

Conclusion. This section is a brief final summary of the entire research project, including the significant findings. This section also suggests future research that would complement the current project.

Appendices. While appendices are not commonly used for brief reports, they are included at this point in the report when needed.

Bibliography. All references used in the report are fully cited here. A reader could find all original research reports mentioned in the literature review or elsewhere in the report if desired.

While there are almost as many formal research report formats as publishers or universities that process those reports, the above is a general-purpose guideline of the types of sections often needed for publication.

Disseminating Findings

The dissemination of research findings involves careful planning, thought, consideration of target audiences, and the best way to communicate with those audiences. Writing up results from a research project and having others take notice are two entirely different propositions. The general rule of thumb is that people will not notice unless they are encouraged to do so. To murder the classic line from the film *Field of Dreams*, just because it is built, does not mean they will come.

Disseminating research findings requires determining who the audience is, where they are located, and how to reach them. When considering who the audience is, think about who is likely to be interested in the research project. The audience might include those who do not express enthusiastic interest but might benefit from an awareness of the research. Of course, research participants are likely to have some interest in what was discovered in the research. Other scholars who study similar topics are another obvious audience for the work. Perhaps some policymakers should take note of the work. Organizations that do work in an area related to the topic of the research project are another possibility. Finally, all inquisitive and engaged members of the public represent a possible audience for the work.

Where the audience is located should be pretty obvious once the composition of that audience is determined. The research participants are known since they were part of the study. Interested scholars can be found at professional conferences and via professional organizations' newsletters and scholarly journals. Policymakers include state and federal representatives who, at least in theory, should be available to hear a constituent speak on matters of policy interest. Organizations that work in an area related to the research topic can be found with a simple web search. Finally, disseminating findings to the general public could take many forms: a letter to the editor of a local newspaper, a blog, or even a Facebook post.

Finally, determining how to reach the target audience will vary depending on which specific audience is interested. The norms of the audience should determine the strategy. For example, scholarly journals provide author submission instructions for researchers wishing to disseminate their work via that journal. The same is true for newspaper editorials, where the newspaper's website may contain details about how to format and submit letters to the editor. To reach out to political representatives, a call to their offices or a simple web search should provide information about how to proceed.

Researchers who have conducted high-quality research and have findings that are likely to be of interest to any constituents besides themselves would have a duty as scholars to share those findings.

Key Takeaways

Chapter 15: Presenting Research
• Describe what and with whom to share research. • State oral presentation tips. • State written presentation tips. • Describe the process to disseminate findings.

Glossary

Action Research. A research method where actions are taken during the research project to correct problems rather than developing a theory. Action research is commonly found in education. A teacher may "try out" some new teaching method to improve a lesson rather than developing a theory or publishing a peer-reviewed paper.

ANCOVA. The "Analysis of Covariance" is used to test the effect of a categorical variable on a continuous, independent variable. The categorical variable is often a treatment in an experimental design; for example, perhaps product packaging color (categorical) is changed. The number of sales (continuous) is analyzed to see if the color affects the sales.

ANOVA. A test that is used to analyze the difference in three or more groups of normally distributed samples. See T-Test.

Applied Research. Research that is intended to be applied to a situation rather than further the knowledge of some topic. For example, if a researcher completes a project designed to increase bottled water sales in a small town, it would be considered applied research. See Basic Research.

Basic Research. Research that is intended to further the knowledge of some topic rather than apply to a specific situation. For example, if a researcher completes a project designed to refine some aspect of the Law of Supply and Demand, it would be considered basic research. See Applied Research.

Bias. An undesired over- or under-estimate of the value of a population's parameter. Bias has many potential sources, including sampling error, measurement error, and missing data. On a survey question, bias tends to elicit a particular response which would skew the data collected.

Binary. A binary scale is used to measure nominal data with only two values: true/false or yes/no.

Bivariate. A type of analysis involving two variables. Examples of bivariate analysis include finding a correlation and regression. See Univariate.

Boundary Condition. The assumptions about the "who, when, and where" in theory. Boundary conditions govern how a theory can be applied or not applied.

Chi-Square. The chi-square test is used to determine the difference between the actual and expected result of a non-parametric (usually nominal) variable in a research project.

Concurrent Validity. The degree that a measure relates to an outcome that is presumed to co-occur. For example, a new employee attitude test results would be the same as an older test if those tests have high concurrent validity. See Predictive Validity.

Construct. The characteristic of a person or organization to be assessed. Typically, a construct is not directly measurable, so various indirect indicators must be used. Examples of constructs include employee attitude, happiness, or self-esteem. Usually, researchers operationalize a concept by defining it in measurable terms, and that becomes a construct.

Construct Validity. The degree to which a test measures what it claims to measure. For example, if a research project purports to investigate one aspect of the local farmers' market, does the project research that aspect? Construct validity is sometimes thought to be the overarching type of validity since research projects that do not address the construct of interest can have no other validity. See Validity.

Constructivism. A philosophical stance that reality is a construct of the human mind and is, therefore, subjective. Typically, qualitative research methods are used by researchers who are constructivists.

Content Validity. A determination of whether a measure correctly assesses the construct's content. For example, if a research project is attempting to determine what drives sales but only measures the price of the merchandise being sold and ignores factors like advertising and competition would call into question the content validity of the study. See Validity.

Continuous Data. Continuous data are quantitative data that can represent any measured value, including fractions and decimals. In mathematics terms, continuous data are members of the Real Number System. See Quantitative Data.

CPI. Consumer Price Index.

Covariate. A covariate is a variable that is potentially predictive of the outcome of other variables in an experimental design. Covariates may be of interest due to their predictive nature, or they may be confounding variables that need to be controlled.

Convergent Validity. The closeness that two measures relate to, or "converge on," a single construct. For example, if a research project measures the number of sales of carbonated drinks, fruit juices, and bottled water in a store, it would be expected that those would converge on a construct of "drink sales." See Discriminant Validity.

Correlation. A correlation is a relationship between two variables. Correlations are typically defined statistically as a value between -1.00 and +1.00. Correlations should not be confused with causation but only indicate that two variables seem to vary together.

Criterion Validity. The degree to which a measure is related to an outcome. See Validity.

Cross-Sectional. A type of research that is conducted at a single point in time that crosses multiple analytical units. This type of study is most often a survey but could be used by other research methods. For example, a survey of several different small business owners in a single city would be cross-sectional. See Longitudinal.

Database. A database is a collection of data that is organized to be easily managed. While the internal structure can be complex, it is typically represented as tables with data in rows and columns, like a spread sheet.

Deductive Research. A research methodology that works from a general theory to specific observations. This type of research is sometimes called "theory-testing."

Dependent Variable. Dependent variables are the outcomes for an observation. For example, if a medical researcher conducts an experiment where a drug is administered, and then the patient's blood pressure is measured, the blood pressure reading is the dependent variable; that is, the blood pressure depends on the drug being administered. See Independent Variable.

Descriptive Research. Research that is designed to describe observed phenomena. The goal is to improve understanding rather than explore new ideas. See Exploratory Research.

Discrete Data. Discrete data are a type of quantitative data that can be counted with integers. See Quantitative Data.

Discriminant Validity. The degree that a measure does *not* measure, or "discriminates between," one of two competing constructs. For example, a measure of the sale of toiletries in a department store would not be related to the construct of "drink sales." See Convergent Validity.

Epistemology. A branch of philosophy that is concerned with the sources of knowledge.

Ethnography. A research method where culture is studied in depth. Typically, researchers "join" a culture and observe social interactions from within. For example, a researcher who lives in a commune for several years and then writes about social interactions observed is conducting ethnographic research.

Excess Kurtosis. Excess kurtosis is a measure of the "tailedness" of a normal distribution. Greater excess kurtosis values indicate longer "tails" (and a "sharper" appearance) on the graph of the distribution.

Explanatory Power. A theory or hypothesis has explanatory power if it accurately predicts phenomena. The strength of the prediction can be statistically measured in quantitative research projects by calculating variance in regression analysis.

Explanatory Research. Research that is designed to explain observed phenomena or processes. See Exploratory Research.

Exploratory Research. Research that explores data to find new ideas.

External Validity. The degree to which a research project's results can be applied outside the context of the study. For example, if a research project that studied manufacturing firms in the midwest could be applied to firms in the south, that study would have high external validity. See Validity.

Face Validity. A determination of whether an indicator is a reasonable measure of an underlying construct "on its face." For example, is the amount of money spent on live theater tickets a measure of social class? See Validity.

Factorial Design. An experimental design where several factors are studied to determine which has the most significant influence on the subject of interest. For example, the sales of a product could be analyzed as a factor of time of day, location of the store, price, and other factors.

Falsifiability. A theory or hypothesis must be disprovable; that is, there must be a way to prove it wrong using evidence.

Functionalism. A belief in the practical application of a theory. Functionalism is more concerned with how a theory can be used in the real world than researching to increase understanding.

Grounded Theory. A theory based on observation rather than experimentation. Thus, the theory's strength depends on the researcher's skill and may not be repeatable by a different researcher or at a different time.

Guttman. The Guttman scale uses a series of questions with increasing intensity to determine how strongly respondents believe some proposition.

Hawthorne. This effect was described in the 1950s when Henry A. Landsberger observed workers in the Hawthorne Works electric company. He noticed that when workers thought they were being observed, they tended to work harder and perform better. Thus, the Hawthorne effect alters peoples' behavior when they think they are being observed.

Hermeneutics. The study of the methodology of interpreting texts. This method was initially applied to Biblical studies but now includes most humanities like law, history, and philosophy.

Hypothesis. A proposed explanation for an observed phenomenon. Often, a hypothesis that may be based on incomplete information is the starting point for further investigation. For example, if a merchant notices that eye-level shelves tend to need restocking more frequently, a hypothesis may be proposed that shoppers purchase goods from eye-level shelves first. The plural form is hypotheses.

Idiographic. An observed phenomenon explains only a single case and does not apply to a broader population. See Nomothetic.

Independent Variable. Independent variables are those that create an observed effect. For example, if a farmer conducts an experiment where different types of fertilizer are applied to two fields to see which is more effective, then the type of fertilizer is the independent variable; that is, the type of fertilizer creates the observed effect. See Dependent Variable.

Inductive Research. A research methodology that works from specific observations to a general theory. This type of research is sometimes called "theory-building."

Internal Validity. The degree to which a research project avoids confounding multiple variables within the study. A project with high internal validity facilitates selecting one explanation over an alternate since confounding variables are controlled. See Validity.

Interpretive Research. Interpretive research explores an observed phenomenon within its social context to discern people's ascribed meaning. This type of research is firmly grounded in constructivism, where it is believed that reality is shaped by perception rather than a knowable "truth."

Interpretivism. A research method that relies on observation and techniques like interviews to understand phenomena.

Interval Data. Interval data are a type of quantitative data measured along a scale where each point is an equal distance from the last. It is possible to compare the distance between two points on an interval scale; for example, the difference between 90 and 100 degrees is the same as the difference between 40 and 50 degrees. However, since an interval scale does not have a zero point, 100 degrees is not twice as hot as 50 degrees. See Quantitative Data.

IRB. Institutional Review Board.

Likert. The Likert scale is one of the most commonly used instruments for measuring attitudes and opinions. It consists of a statement followed by, typically, five selections: Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree.

Logic. A systematic set of principles that validates a theory. Logic is a mode of reasoning that can link observations to explanations.

Logical Consistency. A theory or hypothesis is logically consistent when all constructs, propositions, boundary conditions, and assumptions are congruous.

Longitudinal. A type of research that is conducted over a long period. This type of research most often uses a survey but could be applied to other research methods. For example, repeated surveys over five years of small business owners in a single city would be longitudinal. See Cross-Sectional.

Meta-Analysis. This research method examines data collected from several studies of the same subject to detect trends or overall observations.

Mixed Methods. This research method uses both quantitative and qualitative research processes to understand the phenomenon under consideration better.

Model. A model represents all or part of a system that is constructed to study that system. For example, meteorologists create elaborate models to predict the path of a hurricane.

Nonequivalent Groups Design. A quasi-experimental design where pre-test/post-test measures are made, but the subjects have not been randomly assigned to treatment and control groups.

Nonequivalent Dependent Variable Design. A type of single-group quasi-experimental design where the treatment outcome is measured on one variable but a second related variable is used as a control if the treatment is not expected to influence that variable. For example, the flow of automobile traffic in one intersection may be influenced by a change in the traffic light timing. However, the flow in a similar nearby intersection should not change.

Nominal Data. Nominal data are a type of grouped qualitative data but with no order implied in the grouping. As an example, the gender of survey respondents is nominal data. See Qualitative Data.

Nomothetic. An explanation for an observed phenomenon that is applicable across a broad population rather than a single example. See Idiographic.

Non-Probability Sampling. A type of sampling that does not involve a random selection from the population. This type of sampling is called non-probability since some population members have no probability of being selected. See Probability Sampling.

Nonparametric. Nonparametric data do not conform to a distribution, are skewed, or are qualitative. Statistical tests that work with nonparametric data are generally less powerful and predictive than tests that work with parametric data. See Parametric.

Normal Distribution. A listing of all possible values in a data set plus the number of times each value appears is called a "distribution." In many, perhaps most, business research projects, a distribution exhibits a bell shape when plotted on a graph where the values in the middle of the range are more frequent than values at the extremes of the range. This distribution is called a "normal distribution" because it is frequent.

Objectivism. A philosophical stance that there exists an objective reality can be studied and understood.

Ontology. The branch of philosophy that is concerned with the nature of reality.

Operationalization. The process of designing precise measures for abstract theoretical constructs.

Ordinal Data. Ordinal data are a type of qualitative data that are grouped where the groupings have an implied order. For example, the "satisfaction" rating on a customer survey typically permits respondents to choose from several levels where one level is somehow better than another. See Qualitative Data.

P-Value. The probability that some finding reflects a null hypothesis. In most business research, a p-value of less than 0.05 (or 5%) is desired to infer that the null hypothesis can be rejected.

Paradigm. A pattern or model of how things work in the world. See Theory.

Parametric. Parametric data are data that conform to a distribution, usually a normal distribution. Statistical tests that work with parametric data are generally much more powerful and predictive than tests that work with nonparametric data. See Nonparametric.

Parsimony. A fundamental aspect of research states that if two or more competing explanations are considered, the simplest must be accepted. Thus, a researcher would state that the pyramids were built by humans using available technology rather than aliens in spaceships.

Population. A set of similar items or events of interest to a researcher. For example, the set of small business owners in the United States would be a population. See Sample.

Positivism. This philosophical system posits that any justifiable assertion can be scientifically verified using statistics and logic. Thus, positivism rejects concepts like metaphysics and theism.

Positivist. A researcher who uses positivist techniques on research projects. See Positivism.

Post-Modernism. A philosophical reaction to the assumptions and values of the "modern" period (roughly the 17th to 19th century). Post-modernists believe that rather than an objective reality independent of humans, there is a subjective interpretation of reality, so there is no such thing as a single "Truth."

Pragmatism. An approach to research that values practical application over theory-building.

Precision. Research projects must precisely focus on one aspect of a problem, or they will become so broad that their value will be diminished.

Predictive Validity. The degree to which a measure predicts an outcome. For example, does an increase in beer sales (a measure) predict increased potato chip sales? See Concurrent Validity.

Probability Sampling. A type of sampling that involves a random selection from a population. It is called probability sampling since every member of the population has a probability of being selected.

This type of sampling is frequently called "random sampling" since members of the population are chosen randomly. See Non-probability Sampling.

Proposition. This statement expresses a judgment or opinion.

Qualitative Data. Qualitative data approximates or describes attributes that cannot be directly measured, like employee morale, customer relationships, and management effectiveness. Typically, qualitative data answer questions like "why" and "how come." See Quantitative Data.

Qualitative Research. Qualitative research typically intends to explore observed phenomena to develop hypotheses and dive deep into a problem. Qualitative data collection involves semi-structured activities like focus groups and ethnographies. See Quantitative Research.

Quantitative Data. Quantitative data are numeric measurements of attributes, like the number of employees, the median value of housing, and total revenue. Quantitative data are gathered and analyzed using statistical methods. See Qualitative Data.

Quantitative Research. Quantitative research typically uses numerical data and statistical analysis to find patterns and generalize results to a large population. Quantitative data collection involves structured activities like surveys, interviews, and systematic observations. See Qualitative Research.

Questionnaire. A type of survey research tool comprised of a written set of questions. Questionnaires are typically self-administered. That is, they are sent to respondents and completed without assistance.

Radical Humanism. Humanism is a philosophical and ethical stance that emphasizes the value of human beings. It prefers critical thinking and evidence over dogma and superstition. Radical humanists believe that the world is constantly changing, in sometimes radical ways, with few predictable patterns. Research often involves subjectively interpreting evidence like interviews and focus groups.

Radical Structure. A structuralist believes that the world can be studied objectively and understood mathematically and scientifically without subjective interpretation. Radical structuralists believe that the world is constantly changing, in sometimes radical ways, with few predictable patterns. Research often involves objectively interpreting evidence like direct measurements of populations.

Ratio Data. Ratio data are a type of quantitative data measured along a scale where each point is an equal distance from the next, and there is a zero point. An example of ratio data is people's heights, measured along a uniform scale, e.g., inches or centimeters. Because there is an actual zero point, it is possible to determine that one person is twice as tall as someone else. See Quantitative Data.

Realism. A philosophical position that the world exists apart from human interpretation and understanding. A realist believes that research must be objective and not dependent upon the interpretation of the researcher.

Regression-Discontinuity Design. A quasi-experimental design where subjects are assigned to a treatment or control group based on a cutoff score on a pre-program measure.

Reliability. A descriptor for the consistency of a concept's measure. It is desirable to achieve the same or nearly the same values for each sampling. If the mean age of people in one sample is 30 but 50 in another, it would indicate a problem with the data reliability. See Validity.

Replicability. A research project must be able to be replicated by other researchers or at other times in order to be considered sound.

Sample. A subset of a population from which data are drawn to make inferences about the entire population. See Population.

Sampling Frame. A subset of a sample that is accessible to the researcher. For example, if the sample is high school students, then the sampling frame could be the students in a specific high school or city.

Semantic Differential. The Semantic Differential scale determines attitudes or opinions using a sliding scale of values between two opposite pairs of adjectives. For example, respondents can be asked to choose some value between "1-Dislike" and "5-Like" for a specific snack sample in a store.

Semiotics. The study of signs and symbols and their use or interpretation.

SES. Socio-Economic Status.

Skew. Skew is an asymmetry in a distribution, so a graph appears distorted. A positive skew creates a longer tail on the right side of the graph.

Statistical Conclusion Validity. The degree to which the conclusions found in a research project are correct. Studies with high statistical conclusion validity minimize the two types of statistical errors: Type I (finding a correlation when there is none) and Type II (failing to find a correlation when one exists). See Validity.

Survey. A research method uses standardized questionnaires or interviews to collect data about people and their preferences, thoughts, and behaviors systematically.

T-Test. A test that is used to analyze the difference in two groups of normally distributed samples. See ANOVA.

Theory. A system of ideas that is intended to explain phenomena. Theories that scientists accept have been repeatedly tested and can be used to make accurate predictions. Unlike common usage, a scientific theory is a tested, falsifiable explanation for phenomena. See Paradigm.

Translational Validity. The degree to which a research project has measured a construct. Translational validity is divided into two types: face and content. See Validity.

Univariate. A type of analysis involving a single variable. Univariate analysis findings include the central measure, standard deviation, and frequency distributions. Graphic tools include box plots for continuous data and bar plots for discrete data. See Bivariate.

Validity. A descriptor of whether a research project is measuring the variable under question. For example, if a project hypothesis is that older men tend to tip more than younger men, then the study's validity would be questioned if the researcher only sampled men under 40. See Reliability.

Variable. In scientific research, a variable is a measurable representation of an abstract construct. For example, Intelligence Quotient (IQ) is a construct that cannot be directly measured. However, variables like verbal and mathematical acuity can be measured and are assumed to be a proxy for IQ.

YDS. Youth Development Study.

Bibliography

Akerlof, G. A. (1978). The market for “lemons”: Quality uncertainty and the market mechanism. In *Uncertainty in Economics* (pp. 235–251). Elsevier.

Allcott, H., & Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review, 104*, 3003–37.

Anyan, F. (2013). The Influence of Power Shifts in Data Collection and Analysis Stages: A Focus on Qualitative Research Interview. *Qualitative Report, 18*, 36.

Arnett, J. J. (2008). The neglected 95%: why American psychology needs to become less American. *American Psychologist, 63*, 602.

Association, M. R. (2018). *Code of Marketing Research Standards*. Retrieved from Code of Marketing Research Standards: https://www.insightsassociation.org/issues-policies/mra-code-marketing-research-standards

Babbie, E. (2010). The practice of social research (ed.). *Wadsworth: Nelson Education Ltd*.

Babbie, E., & Wagenaar, T. (2010). Unobtrusive research. *The practice of social research, 320*.

Bacharach, S. B. (1989). Organizational theories: Some criteria for evaluation. *Academy of management review, 14*, 496–515.

Baker, L. (2006). Observation: A complex research method. *Library trends, 55*, 171–189.

Bansal, P., & Roth, K. (2000). Why companies go green: A model of ecological responsiveness. *Academy of management journal, 43*, 717–736.

Bearden, W. O., Hardesty, D. M., & Rose, R. L. (2001). Consumer self-confidence: Refinements in conceptualization and measurement. *Journal of consumer research, 28*, 121–134.

Bell, E., & Bryman, A. (2015). *Business research methods.* Oxford university press.

Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The case research strategy in studies of information systems. *MIS quarterly*, 369–386.

Berger, J., & Iyengar, R. (2013). Communication channels and word of mouth: How the medium shapes the message. *Journal of consumer research, 40*, 567–579.

Bhattacherjee, A. (2012). Social science research: Principles, methods, and practices.

Binder, M., & Edwards, J. S. (2010). Using grounded theory method for theory building in operations management research. *International Journal of Operations & Production Management*.

Biomedical, N. C., & Behavioral Research, M. (1978). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research.* Superintendent of Documents.

Blackstone, A. (2012). *Principles of sociological inquiry–qualitative and quantitative methods.* The Saylor Foundation.

Blee, K. M. (2005). Racial violence in the United States. *Ethnic and Racial Studies, 28*, 599–619.

Bluebond-Langner, M. (2000). *In the shadow of illness: Parents and siblings of the chronically ill child.* Princeton University Press.

Boas, F. (1964). *The Central Eskimo* (Vol. 196). University of Nebraska Press.

Bobbitt-Zeher, D., & Downey, D. B. (2013). Number of siblings and friendship nominations among adolescents. *Journal of Family Issues, 34*, 1175–1193.

Braunsberger, K., & Buckler, B. (2011). What motivates consumers to participate in boycotts: Lessons from the ongoing Canadian seafood boycott. *Journal of Business Research, 64*, 96–102.

Brewer, J., & Hunter, A. (1989). *Multimethod research: a synthesis of styles.* Sage Publications, Inc.

Bridge, G., & Dowling, R. (2001). Microgeographies of retailing and gentrification. *Australian Geographer, 32*, 93–107.

Bright, A. D. (2000). The role of social marketing in leisure and recreation management. *Journal of leisure research, 32*, 12–17.

Brown, H. E. (2013). Race, legality, and the social policy consequences of anti-immigration mobilization. *American Sociological Review, 78*, 290–314.

Buchanan, J., & Jones, M. L. (2010). The efficacy of utilising Nvivo for interview data from the electronic gaming industry in two jurisdictions.

Burrell, G., & Morgan, G. (2017). *Sociological paradigms and organisational analysis: Elements of the sociology of corporate life.* Routledge.

Buysse, J. A., & Embser-Herbert, M. S. (2004). Constructions of gender in sport: An analysis of intercollegiate media guide cover photographs. *Gender & Society, 18*, 66–81.

Cai, H., Chen, Y., & Fang, H. (2009). Observational learning: Evidence from a randomized natural field experiment. *American Economic Review, 99*, 864–82.

Caruth, G. D. (2013). Demystifying Mixed Methods Research Design: A Review of the Literature. *Online Submission, 3*, 112–122.

Chatman, E. A. (1992). *The information world of retired women.* Greenwood Publishing Group.

Chavez, M., Whiteford, S., & Hoewe, J. (2010). Reporting on immigration: A content analysis of major US newspapers' coverage of Mexican immigration. *Norteamérica, 5*, 111–125.

Cheyne, A. D., Dorfman, L., Bukofzer, E., & Harris, J. L. (2013). Marketing sugary cereals to children in the digital age: a content analysis of 17 child-targeted websites. *Journal of health communication, 18*, 563–582.

Cho, J., & Trent, A. (2006). Validity in qualitative research revisited. *Qualitative research, 6*, 319–340.

Chubin, D. E. (1985). Research malpractice. *BioScience, 35*, 80–89.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology, 13*, 3–21.

Council, K. (2019). *Ethical Issues in Research*. Retrieved from Ethical Issues in Research: https://www.kirklees.gov.uk/involve/document/EthicalIssues.pdf

Council, K. (2019). *Research with Children and Young People*. Retrieved from Research with Children and Young People: https://www.kirklees.gov.uk/involve/document/ResearchChildrenYoungPeople.pdf

Creswell, J. W. (2014). Research design: qualitative, quantitative, and mixed methods approaches.

Creswell, J. W., Fetters, M. D., & Ivankova, N. V. (2004). Designing a mixed methods study in primary care. *The Annals of Family Medicine, 2*, 7–12.

Crook, T. R., Todd, S. Y., Combs, J. G., Woehr, D. J., & Ketchen Jr, D. J. (2011). Does human capital matter? A meta-analysis of the relationship between human capital and firm performance. *Journal of applied psychology, 96*, 443.

Cummings, A. D. (1992). The Exxon Valdez Oil Spill and the Confidentiality of Natural Resource Damage Assessment Data. *Ecology LQ, 19*, 363.

Davies, I. A., & Crane, A. (2010). Corporate social responsibility in small-and medium-size enterprises: investigating employee engagement in fair trade companies. *Business Ethics: A European Review, 19*, 126–139.

Davis, A. K., Piger, J. M., & Sedor, L. M. (2012). Beyond the numbers: Measuring the information content of earnings press release language. *Contemporary Accounting Research, 29*, 845–868.

Davis, F. (1973). The Martian and the convert: Ontological polarities in social research. *Urban Life and Culture, 2*, 333–343.

de Oña, J., de Oña, R., & López, G. (2016). Transit service quality analysis using cluster analysis and decision trees: a step forward to personalized marketing in public transportation. *Transportation, 43*, 725–747.

De Vries, R., Anderson, M. S., & Martinson, B. C. (2006). Normal misbehavior: Scientists talk about the ethics of research. *Journal of Empirical Research on Human Research Ethics, 1*, 43–50.

Delaney, J. T., & Huselid, M. A. (1996). The impact of human resource management practices on perceptions of organizational performance. *Academy of Management journal, 39*, 949–969.

Denison, D. R. (1990). *Corporate culture and organizational effectiveness.* John Wiley & Sons.

Denny, K. E. (2011). Gender in context, content, and approach: Comparing gender messages in Girl Scout and Boy Scout handbooks. *Gender & Society, 25*, 27–47.

Dictionaries, O. L. (2018). *Definition of Ethics*. Retrieved from Definition of Ethics: https://en.oxforddictionaries.com/definition/ethics

Dillman, D. A. (2011). *Mail and Internet surveys: The tailored design method–2007 Update with new Internet, visual, and mixed-mode guide.* John Wiley & Sons.

Dobrow Riza, S., Ganzach, Y., & Liu, Y. (2015). Time and job satisfaction: A longitudinal study of the differential roles of age and tenure. *Journal of management*.

Donmoyer, R. (2001). Paradigm talk reconsidered. *Handbook of research on teaching, 4*, 174–197.

Dorow, S. (2006). Racialized Choices: Chinese adoption and the white noise of blackness. *Critical Sociology, 32*, 357–379.

Drost, E. A., & others. (2011). Validity and reliability in social science research. *Education Research and perspectives, 38*, 105.

Duhigg, C. (2012). How companies learn your secrets. *The New York Times, 16*, 2012.

Dutra, L. M., & Glantz, S. A. (2014). Electronic cigarettes and conventional cigarette use among US adolescents: a cross-sectional study. *JAMA pediatrics, 168*, 610–617.

Eisenhardt, K. M. (1989). Making fast strategic decisions in high-velocity environments. *Academy of Management journal, 32*, 543–576.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *The Academy of Management Journal, 50*, 25–32.

Ellwood, D., Kane, T. J., & others. (2000). Who is getting a college education? Family background and the growing gaps in enrollment. *Securing the future: Investing in children from birth to college*, 283–324.

Esterberg, K. G. (2002). Qualitative methods in social research.

Evans, D., Smith, M., & Willen, L. (2005). SECRET.

Faden, R. R., & Beauchamp, T. L. (1986). *A history and theory of informed consent.* Oxford University Press.

Gainsbury, S., & Blaszczynski, A. (2011). The appropriateness of using laboratories and student participants in gambling research. *Journal of Gambling Studies, 27*, 83–97.

Gallopel-Morvan, K., Gabriel, P., Le Gall-Ely, M., Rieunier, S., & Urien, B. (2011). The use of visual warnings in social marketing: The case of tobacco. *Journal of Business Research, 64*, 7–11.

Giorgi, A., & Giorgi, B. (2003). *Phenomenology.* Sage Publications, Inc.

Glaser, B., & Strauss, A. (1967). The discovery of grounded theory. *London: Weidenfeld and Nicholson, 24*, 288–304.

Gluck, J., Schaub, F., Friedman, A., Habib, H., Sadeh, N., Cranor, L. F., & Agarwal, Y. (2016). How short is too short? Implications of length and framing on the effectiveness of privacy notices. *12th Symposium on Usable Privacy and Security (SOUPS)*, (pp. 321–340).

Gowan, T. (2010). *Hobos, hustlers, and backsliders: Homeless in San Francisco.* U of Minnesota Press.

Green, R. M. (2006). Direct-to-consumer advertising and pharmaceutical ethics: The case of Vioxx. *Hofstra L. Rev., 35*, 749.

Greene, M. J. (2014). On the inside looking in: Methodological insights and challenges in conducting qualitative insider research. *The qualitative report, 19*, 1–13.

Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational behavior and human performance, 16*, 250–279.

Härdle, W. (2012). *Smoothing techniques: with implementation in S.* Springer Science & Business Media.

Hechavarria, D. M., & Reynolds, P. D. (2009). Cultural norms & business start-ups: the impact of national values on opportunity and necessity entrepreneurs. *International Entrepreneurship and Management Journal, 5*, 417.

Henderson, H. V., & Velleman, P. F. (1981). Building multiple regression models interactively. *Biometrics*, 391–411.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature, 466*, 29.

Hill, K. (2012). How Target figured out a teen girl was pregnant before her father did. *Forbes, Inc*.

Hochschild, A., & Machung, A. (2012). *The second shift: Working families and the revolution at home.* Penguin.

Hom, P. W., & Griffeth, R. W. (1995). *Employee turnover.* South-Western Pub.

Huck, J. F., & McEwen, T. (1991). Competencies needed for small business success: perceptions. *Journal of Small Business Management, 29*, 90.

Hughes, J. E., Knittel, C. R., & Sperling, D. (2006). Evidence of a shift in the short-run price elasticity of gasoline demand. *Evidence of a shift in the short-run price elasticity of gasoline demand*.

Huhtala, M., Kaptein, M., & Feldt, T. (2016). How perceived changes in the ethical culture of organizations influence the well-being of managers: A two-year longitudinal study. *European Journal of Work and Organizational Psychology, 25*, 335–352.

Human, U. S., & Services, H. (2018). *Protection of Human Subjects*. Retrieved from Protection of Human Subjects: https://www.ecfr.gov/cgibin/retrieveECFR?gp=&SID=83cd09e1c0f5c6937cd9d7513160fc3f&pitd=20180719&n=pt45.1.46&r=PART&ty=HTML

Human, U. S., & Services, H. (2018). *Special Protections for Children as Research Subjects*. Retrieved from Special Protections for Children as Research Subjects: https://www.hhs.gov/ohrp/regulations-and-policy/guidance/special-protections-for-children/index.html

Humphreys, L. (1970). Tearoom trade. *Trans-action, 7*, 10–25.

Irani, L. C., & Silberman, M. (2013). Turkopticon: Interrupting worker invisibility in amazon mechanical turk. *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 611–620). ACM.

Israel, G. D. (1992). Determining sample size.

Jack, S., Moult, S., Anderson, A. R., & Dodd, S. (2010). An entrepreneurial network evolving: Patterns of change. *International Small Business Journal, 28*, 315–337.

Johnson, J. S., & Sohi, R. S. (2016). Understanding and resolving major contractual breaches in buyer–seller relationships: a grounded theory approach. *Journal of the Academy of Marketing Science, 44*, 185–205.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational researcher, 33*, 14–26.

Jørgensen, M. B., Villadsen, E., Burr, H., Punnett, L., & Holtermann, A. (2016). Does employee participation in workplace health promotion depend on the working environment? A cross-sectional study of Danish workers. *BMJ open, 6*, e010516.

Kaplan, A. (2017). *The conduct of inquiry: Methodology for behavioural science.* Routledge.

Keeter, S., Dimock, M., & Christian, L. (2008). Calling cell phones in ’08 pre-election polls. *The Pew Research Center for the People & the Press. Retrieved December, 18*, 2008.

Kimmel, M. S. (2008). The perilous world where boys become men. *HarperCollins*.

Kirkpatrick, D. D. (2005). Conservatives pick soft target: A cartoon sponge. *The New York Times, 20.*

Kotler, P., & Roberto, E. L. (1989). Social marketing. Strategies for changing public behavior.

Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences, 111*, 8788–8790.

Krippendorff, K. (2009). Testing the reliability of content analysis data. *The content analysis reader*, 350–357.

Kuruvilla, S., & Ranganathan, A. (2008). Economic development strategies and macro-and micro-level human resource policies: The case of India's “outsourcing” industry. *ILR Review, 62*, 39–72.

Lambe, C. J., Wittmann, C. M., & Spekman, R. E. (2001). Social exchange theory and research on business-to-business relational exchange. *Journal of Business-to-Business Marketing, 8*, 1–36.

Larson, E. W. (2010). Time and the constitution of markets: internal dynamics and external relations of stock exchanges in Fiji, Ghana and Iceland. *Economy and Society, 39*, 460–487.

Lincoln, Y. S. (1985). Naturalistic inquiry. *The Blackwell Encyclopedia of Sociology*.

Lofland, J. (1995). Analytic ethnography: Features, failings, and futures. *Journal of contemporary ethnography, 24*, 30–67.

Lorenzini, E. (2017). Mixed-method research in the health sciences. *Revista Cuidarte, 8*, 1549–1560.

Macera, J. L. (2016). A theory of thriving for small business owners: A grounded theory study. *A theory of thriving for small business owners: A grounded theory study*.

Malinowski, B. (2014). *Argonauts of the western Pacific.* Routledge.

Markus, M. L. (1987). Toward a "critical mass" theory of interactive media: Universal access, interdependence and diffusion. *Communication research, 14*, 491–511.

Marshall, D., O'Donohoe, S., & Kline, S. (2007). Families, food, and pester power: beyond the blame game? *Journal of Consumer Behaviour: An International Research Review, 6*, 164–181.

Marx, K. (1971). Contribution to the Critique of Hegel's Philosophy of Law. *Introduction". Marx and Engels Collected Works, 3*.

Maurer, H. A., Kappe, F., & Zaka, B. (2006). Plagiarism-A survey. *J. UCS, 12*, 1050–1084.

McElroy, J. C., & Morrow, P. C. (2010). Employee reactions to office redesign: A naturally occurring quasi-field experiment in a multi-generational setting. *Human relations, 63*, 609–636.

McLean, S. (2012). *Communication for Business Success.*

Meier, B. (2005). Maker of heart device kept flaw from doctors. *New York Times*.

Merton, R. K., & Kendall, P. L. (1946). The focused interview. *American journal of Sociology, 51*, 541–557.

Messner, M. A. (2002). *Taking the field: Women, men, and sports* (Vol. 4). U of Minnesota Press.

Milgram, S. (1963). Behavioral study of obedience. *The Journal of abnormal and social psychology, 67*, 371.

Minnesota, U. o. (2018). *Youth Development Study*. Retrieved from Youth Development Study: https://cla.umn.edu/sociology/graduate/collaboration-opportunities/youth-development-study

Mintzberg, H. (1971). Managerial work: Analysis from observation. *Management science, 18*, B–97.

Morgan, D. L. (1996). Focus groups. *Annual review of sociology, 22*, 129–152.

Mosse, G. L. (1996). *The image of man.* New York/Oxford.

Musalem, A., Aburto, L., & Bosch, M. (2018). Market basket analysis insights to support category management. *European Journal of Marketing*.

Parboteeah, K. P., Paik, Y., & Cullen, J. B. (2009). Religious groups and work values: A focus on Buddhism, Christianity, Hinduism, and Islam. *International Journal of Cross Cultural Management, 9*, 51–67.

Park, Y. A., & Gretzel, U. (2007). Success factors for destination marketing web sites: A qualitative meta-analysis. *Journal of travel research, 46*, 46–63.

Percheski, C. (2008). Opting out? Cohort differences in professional women's employment rates from 1960 to 2005. *American sociological review, 73*, 497–517.

Politics, C. N. (2016). *Election 2016*. Retrieved from Election 2016: https://www.cnn.com/election/2016/results/exit-polls

Ponterotto, J. G. (2005). Qualitative research in counseling psychology: A primer on research paradigms and philosophy of science. *Journal of counseling psychology, 52*, 126.

Price, P. C., Chiang, I.-C. A., & Jhangiani, R. (2015). *Research methods in psychology.* BCcampus, BC Open Textbook Project.

Raval, K. M. (2012). Data Mining Techniques. *International Journal of Advanced Research in Computer Science and Software Engineering, 2*.

Reich, R. B. (2016). *Saving capitalism: For the many, not the few.* Vintage.

Reilly, M. D., & Wallendorf, M. (1987). A comparison of group differences in food consumption using household refuse. *Journal of Consumer Research, 14*, 289–294.

Reverby, S. M. (2009). *Examining Tuskegee: The infamous syphilis study and its legacy.* Univ of North Carolina Press.

Richards, K. A., & Hemphill, M. A. (2018). A practical guide to collaborative qualitative data analysis. *Journal of Teaching in Physical Education, 37*, 225–231.

Riedel, S. (2005). Edward Jenner and the history of smallpox and vaccination. *Baylor University Medical Center Proceedings. 18*, pp. 21–25. Taylor & Francis.

Rindfuss, R. R., Choe, M. K., Tsuya, N. O., Bumpass, L. L., & Tamaki, E. (2015). Do low survey response rates bias results? Evidence from Japan. *Demographic Research, 32*, 797–828.

Roper, S., & Parker, C. (2013). Doing well by doing good: A quantitative investigation of the litter effect. *Journal of Business Research, 66*, 2262–2268.

Sanjari, M., Bahramnezhad, F., Fomani, F. K., Shoghi, M., & Cheraghi, M. A. (2014). Ethical challenges of researchers in qualitative studies: the necessity to develop a specific guideline. *Journal of medical ethics and history of medicine, 7*.

Scarce, R. (2016). *Eco-warriors: Understanding the radical environmental movement.* Routledge.

Schilling, J. (2006). On the pragmatics of qualitative assessment. *European journal of psychological assessment, 22*, 28–37.

Schuler, R. S., & Rogovsky, N. (1998). Understanding compensation practice variations across firms: The impact of national culture. *Journal of International Business Studies, 29*, 159–177.

Shaffir, W., & Stebbins, R. A. (1991). An inside view of qualitative research. *Fieldwork, Experiencing*.

Sharma, S. (2000). Managerial interpretations and organizational context as predictors of corporate choice of environmental strategy. *Academy of Management journal, 43*, 681–697.

Sheehan, J. G. (2007). Fraud, conflict of interest, and other enforcement issues in clinical research. *Cleveland Clinic journal of medicine, 74*, S63.

Shen, B., & Bissell, K. (2013). Social media, social me: A content analysis of beauty companies' use of Facebook in marketing and branding. *Journal of Promotion Management, 19*, 629–651.

Sherman, L. W., & Berk, R. A. (1984). The specific deterrent effects of arrest for domestic assault. *American sociological review*, 261–272.

Shinebourne, P. (2009). Using Q method in qualitative research. *International Journal of Qualitative Methods, 8*, 93–97.

Sia, S. K., & Bhardwaj, G. (2009). Employees' perception of diversity climate: role of psychological contract. *Journal of the Indian Academy of Applied Psychology, 35*, 305–312.

Snopes. (2005). *Grandma's Cooking Secret*. Retrieved from Grandma's Cooking Secret: https://www.snopes.com/fact-check/grandmas-cooking-secret/

Spradley, J. (2016). P. 1980. Participant observation. *Wadsworth, Belmont, USA*.

Steinfield, C. W., & Fulk, J. (1990). The theory imperative. *Organizations and communication technology*, 13–25.

Stevens, S. S., & others. (1946). On the theory of scales of measurement.

Strauss, A., & Corbin, J. (1994). Grounded theory methodology. *Handbook of qualitative research, 17*, 273–85.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Procedures and techniques for developing grounded theory.* Thousand Oaks, CA: Sage.

Sullivan, D. M. (2014). From food desert to food mirage: Race, social class, and food shopping in a gentrifying neighborhood. *Advances in Applied Sociology*.

Susman, G. I., & Evered, R. D. (1978). An assessment of the scientific merits of action research. *Administrative science quarterly, 23*, 582–603.

Sutton, R. I., & Staw, B. M. (1995). What theory is not. *Administrative science quarterly*, 371–384.

Sylvetsky, A. C., Hennink, M., Comeau, D., Welsh, J. A., Hardy, T., Matzigkeit, L., . . . Vos, M. B. (2013). Youth understanding of healthy eating and obesity: a focus group study. *Journal of obesity*.

Tashakkori, A., Teddlie, C., & Teddlie, C. B. (1998). *Mixed methodology: Combining qualitative and quantitative approaches* (Vol. 46). Sage.

Taylor, J. (2011). The intimate insider: negotiating the ethics of friendship when doing insider research. *Qualitative research, 11*, 3–22.

Teddlie, C., & Tashakkori, A. (2006). A general typology of research designs featuring mixed methods. *Research in the Schools, 13*, 12–28.

Titlebaum, P., & Lawrence, H. (2016). Perceived motivations for corporate suite ownership in the 'big four' leagues.

Twain, M. (2014). *Following the Equator (With Original Illustrations).* Simon and Schuster.

Van der Heyden, M. A., van de Derks Ven, T., & Opthof, T. (2009). Fraud and misconduct in science: the stem cell seduction. *Netherlands Heart Journal, 17*, 25–29.

Venkatesh, V., Brown, S. A., & Bala, H. (2013). Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. *MIS quarterly, 37*.

Vogel, P., Greiser, T., & Mattfeld, D. C. (2011). Understanding bike-sharing systems using data mining: Exploring activity patterns. *Procedia-Social and Behavioral Sciences, 20*, 514–523.

Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of management review, 14*, 490–495.

Wolcott, H. F. (1999). *Ethnography: A way of seeing.* Rowman Altamira.

Wright, D. K., & Hinson, M. D. (2015). Examining social and emerging media use in public relations practice: A ten-year longitudinal analysis. *Public relations journal, 9*, 1–26.

Wright, G. (2015). An empirical examination of the relationship between nonresponse rate and nonresponse bias. *Statistical Journal of the IAOS, 31*, 305–315.

Wutich, A., Lant, T., White, D. D., Larson, K. L., & Gartin, M. (2010). Comparing focus group and individual responses on sensitive topics: a study of water decision makers in a desert city. *Field Methods, 22*, 88–110.

Yin, R. K. (2017). *Case study research and applications: Design and methods.* Sage publications.

Yusuf, A. (1995). Critical success factors for small business: Perceptions of South Pacific entrepreneurs. *Journal of small business management, 33*, 68.

Zimbardo, P. G., Maslach, C., & Haney, C. (2000). Reflections on the Stanford prison experiment: Genesis, transformations, consequences. *Obedience to authority: Current perspectives on the Milgram paradigm*, 193–237.