SOCIAL SCIENCE RESEARCH:
PRINCIPLES, METHODS, AND PRACTICES

ANOL BHATTACHERJEE

SOCIAL SCIENCE RESEARCH:
PRINCIPLES, METHODS, AND PRACTICES

Anol Bhattacherjee, Ph.D.
University of South Florida
Tampa, Florida, USA
abhatt@usf.edu

Second Edition

Social Science Research: Principles, Methods, and Practices, 2nd edition
By Anol Bhattacherjee
First published 2012

ISBN-13: 978-1475146127
ISBN-10: 1475146124

Preface

This book is designed to introduce doctoral and graduate students to the process of scientific research in the social sciences, business, education, public health, and related disciplines. This book is based on my lecture materials developed over a decade of teaching the doctoral-level class on Research Methods at the University of South Florida. The target audience for this book includes Ph.D. and graduate students, junior researchers, and professors teaching courses on research methods, although senior researchers can also use this book as a handy and compact reference.

The first and most important question potential readers should have about this book is how is it different from other text books on the market? Well, there are four key differences. First, unlike other text books, this book is not just about "research methods" (empirical data collection and analysis) but about the entire "research process" from start to end. Research method is only one phase in that research process, and possibly the easiest and most structured one. Most text books cover research methods in depth, but leave out the more challenging, less structured, and probably more important issues such as theorizing and thinking like a researcher, which are often prerequisites of empirical research. In my experience, most doctoral students become fairly competent at research methods during their Ph.D. years, but struggle to generate interesting or useful research questions or build scientific theories. To address this deficit, I have devoted entire chapters to topics such as "Thinking Like a Researcher" and "Theories in Scientific Research", which are essential skills for a junior researcher.

Second, the book is succinct and compact by design. While writing the book, I decided to focus only on essential concepts, and not fill pages with clutter that can divert the students' attention to less relevant or tangential issues. Most doctoral seminars include a fair complement of readings drawn from the respective discipline. This book is designed to complement those readings by summarizing all important concepts in one compact volume, rather than burden students with a voluminous text on top of their assigned readings.

Third, this book is free in its download version. Not just the current edition but all future editions in perpetuity. The book will also be available in Kindle e-Book, Apple iBook, and on-demand paperback versions at a nominal cost. Many people have asked why I'm giving away something for free when I can make money selling it? Well, not just to stop my students from constantly complaining about the high price of text books, but also because I believe that scientific knowledge should not be constrained by access barriers such as price and availability. Scientific progress can occur only if students and academics around the world have affordable access to the best that science can offer, and this free book is my humble effort to that cause. However, free should not imply "lower quality." Some of the best things in life such as air, water, and sunlight are free. Many of Google's resources are free too, and one can well imagine where we would be in today's Internet age without Google. Some of the most sophisticated software programs available today, like Linux and Apache, are also free, and so is this book.

Fourth, I plan to make local-language versions of this book available in due course of time, and those translated versions will also be free. So far, I have had commitments to translate thus book into Chinese, French, Indonesian, Korean, Portuguese, Spanish versions (which will hopefully be available in 2012), and I'm looking for qualified researchers or professors to translate it into Arabic, German, and other languages where there is sufficient demand for a research text. If you are a prospective translator, please note that there will be no financial gains or royalty for your translation services, because the book must remain free, but I'll gladly include you as a coauthor on the local-language version.

The book is structured into 16 chapters for a 16-week semester. However, professors or instructors can add, drop, stretch, or condense topics to customize the book to the specific needs of their curriculum. For instance, I don't cover Chapters 14 and 15 in my own class, because we have dedicated classes on statistics to cover those materials and more. Instead, I spend two weeks on theories (Chapter 3), one week to discussing and conducting reviews for academic journals (not in the book), and one week for a finals exam. Nevertheless, I felt it necessary to include Chapters 14 and 15 for academic programs that may not have a dedicated class on statistical analysis for research. A sample syllabus that I use for my own class in the business Ph.D. program is provided in the appendix.

Lastly, I plan to continually update this book based on emerging trends in scientific research. If there are any new or interesting content that you wish to see in future editions, please drop me a note, and I will try my best to accommodate them. Comments, criticisms, or corrections to any of the existing content will also be gratefully appreciated.

Anol Bhattacherjee
E-mail: abhatt@usf.edu

Table of Contents

Chapter 1

Science and Scientific Research

What is research? Depending on who you ask, you will likely get very different answers to this seemingly innocuous question. Some people will say that they routinely research different online websites to find the best place to buy goods or services they want. Television news channels supposedly conduct research in the form of viewer polls on topics of public interest such as forthcoming elections or government-funded projects. Undergraduate students research the Internet to find the information they need to complete assigned projects or term papers. Graduate students working on research projects for a professor may see research as collecting or analyzing data related to their project. Businesses and consultants research different potential solutions to remedy organizational problems such as a supply chain bottleneck or to identify customer purchase patterns. However, none of the above can be considered "scientific research" unless: (1) it contributes to a body of science, and (2) it follows the scientific method. This chapter will examine what these terms mean.

Science

What is science? To some, science refers to difficult high school or college-level courses such as physics, chemistry, and biology meant only for the brightest students. To others, science is a craft practiced by scientists in white coats using specialized equipment in their laboratories. Etymologically, the word "science" is derived from the Latin word *scientia* meaning knowledge. **Science** refers to a systematic and organized body of knowledge in any area of inquiry that is acquired using "the scientific method" (the scientific method is described further below). Science can be grouped into two broad categories: natural science and social science. **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 can be further classified into physical sciences, earth sciences, life sciences, and others. Physical sciences consist of disciplines such as physics (the science of physical objects), chemistry (the science of matter), and astronomy (the science of celestial objects). Earth sciences consist of disciplines such as geology (the science of the earth). Life sciences include disciplines such as biology (the science of human bodies) and botany (the science of plants). In contrast, **social science** is the science of people or collections of people, such as groups, firms, societies, or economies, and their individual or collective behaviors. 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 firms, markets, and economies).

The natural sciences are different from the social sciences in several respects. The natural sciences are very precise, accurate, deterministic, and independent of the person making the scientific observations. For instance, a scientific experiment in physics, such as measuring the speed of sound through a certain media or the refractive index of water, should always yield the exact same results, irrespective of the time or place of the experiment, or the person conducting the experiment. If two students conducting the same physics experiment obtain two different values of these physical properties, then it generally means that one or both of those students must be in error. However, the same cannot be said for the social sciences, which tend to be less accurate, deterministic, or unambiguous. For instance, if you measure a person's happiness using a hypothetical instrument, you may find that the same person is more happy or less happy (or sad) on different days and sometimes, at different times on the same day. One's happiness may vary depending on the news that person received that day or on the events that transpired earlier during that day. Furthermore, there is not a single instrument or metric that can accurately measure a person's happiness. Hence, one instrument may calibrate a person as being "more happy" while a second instrument may find that the same person is "less happy" at the same instant in time. In other words, there is a high degree of *measurement error* in the social sciences and there is considerable uncertainty and little agreement on social science policy decisions. For instance, you will not find many disagreements among natural scientists on the speed of light or the speed of the earth around the sun, but you will find numerous disagreements among social scientists on how to solve a social problem such as reduce global terrorism or rescue an economy from a recession. Any student studying the social sciences must be cognizant of and comfortable with handling higher levels of ambiguity, uncertainty, and error that come with such sciences, which merely reflects the high variability of social objects.

Sciences can also be classified based on their purpose. **Basic sciences**, also called pure sciences, are those that explain the most basic objects and forces, relationships between them, and laws governing them. Examples include physics, mathematics, and biology. **Applied sciences**, also called practical sciences, are sciences that apply scientific knowledge from basic sciences in a physical environment. For instance, engineering is an applied science that applies the laws of physics and chemistry for practical applications such as building stronger bridges or fuel efficient combustion engines, while medicine is an applied science that applies the laws of biology for solving human ailments. Both basic and applied sciences are required for human development. However, applied sciences cannot stand on their own right, but instead relies on basic sciences for its progress. Of course, the industry and private enterprises tend to focus more on applied sciences given their practical value, while universities study both basic and applied sciences.

Scientific Knowledge

The purpose of science is to create scientific knowledge. **Scientific knowledge** refers to a generalized body of laws and theories to explain a phenomenon or behavior of interest that are acquired using the scientific method. **Laws** are observed patterns of phenomena or behaviors, while **theories** are systematic explanations of the underlying phenomenon or behavior. For instance, in physics, the Newtonian Laws of Motion describe what happens when an object is in a state of rest or motion (Newton's First Law), what force is needed to move a stationary object or stop a moving object (Newton's Second Law), and what happens when two objects collide (Newton's Third Law). Collectively, the three laws constitute the basis of classical mechanics – a theory of moving objects. Likewise, the theory of optics explains the properties of light and how it behaves in different media, electromagnetic theory explains the properties of electricity and how to generate it, quantum mechanics explains the properties of subatomic particles, and thermodynamics explains the properties of energy and mechanical work. An introductory college level text book in physics will likely contain separate chapters devoted to each of these theories. Similar theories are also available in social sciences. For instance, cognitive dissonance theory in psychology explains how people react when their observations of an event is different from what they expected of that event, general deterrence theory explains why some people engage in improper or criminal behaviors, such as illegally download music or commit software piracy, and the theory of planned behavior explains how people make conscious reasoned choices in their everyday lives.

The goal of scientific research is to discover laws and postulate theories that can explain natural or social phenomena, or in other words, build scientific knowledge. It is important to understand that this knowledge may be imperfect or even quite far from the truth. Sometimes, there may not be a single universal truth, but rather an equilibrium of "multiple truths." We must understand that the theories, upon which scientific knowledge is based, are only explanations of a particular phenomenon, as suggested by a scientist. As such, there may be good or poor explanations, depending on the extent to which those explanations fit well with reality, and consequently, there may be good or poor theories. The progress of science is marked by our progression over time from poorer theories to better theories, through better observations using more accurate instruments and more informed logical reasoning.

We arrive at scientific laws or theories through a process of logic and evidence. Logic (theory) and evidence (observations) are the two, and only two, pillars upon which scientific knowledge is based. In science, theories and observations are interrelated and cannot exist without each other. Theories provide meaning and significance to what we observe, and observations help validate or refine existing theory or construct new theory. Any other means of knowledge acquisition, such as faith or authority cannot be considered science.

Scientific Research

Given that theories and observations are the two pillars of science, scientific research operates at two levels: a theoretical level and an empirical level. The theoretical level is concerned with developing abstract concepts about a natural or social phenomenon and relationships between those concepts (i.e., build "theories"), while the empirical level is concerned with testing the theoretical concepts and relationships to see how well they reflect our observations of reality, with the goal of ultimately building better theories. Over time, a theory becomes more and more refined (i.e., fits the observed reality better), and the science gains maturity. Scientific research involves continually moving back and forth between theory and observations. Both theory and observations are essential components of scientific research. For instance, relying solely on observations for making inferences and ignoring theory is not considered valid scientific research.

Depending on a researcher's training and interest, scientific inquiry may take one of two possible forms: inductive or deductive. In **inductive research**, the goal of a researcher is to infer theoretical concepts and patterns from observed data. In **deductive research**, the goal of the researcher is to test concepts and patterns known from theory using new empirical data. Hence, inductive research is also called *theory-building* research, and deductive research is *theory-testing* research. Note here that the goal of theory-testing is not just to test a theory, but possibly to refine, improve, and extend it. Figure 1.1 depicts the complementary nature of inductive and deductive research. Note that inductive and deductive research are two halves of the research cycle that constantly iterates between theory and observations. You cannot do inductive or deductive research if you are not familiar with both the theory and data components of research. Naturally, a complete researcher is one who can traverse the entire research cycle and can handle both inductive and deductive research.

It is important to understand that theory-building (inductive research) and theory-testing (deductive research) are both critical for the advancement of science. Elegant theories are not valuable if they do not match with reality. Likewise, mountains of data are also useless until they can contribute to the construction to meaningful theories. Rather than viewing these two processes in a circular relationship, as shown in Figure 1.1, perhaps they can be better viewed as a helix, with each iteration between theory and data contributing to better explanations of the phenomenon of interest and better theories. Though both inductive and deductive research are important for the advancement of science, it appears that inductive (theory-building) research is more valuable when there are few prior theories or explanations, while deductive (theory-testing) research is more productive when there are many competing theories of the same phenomenon and researchers are interested in knowing which theory works best and under what circumstances.

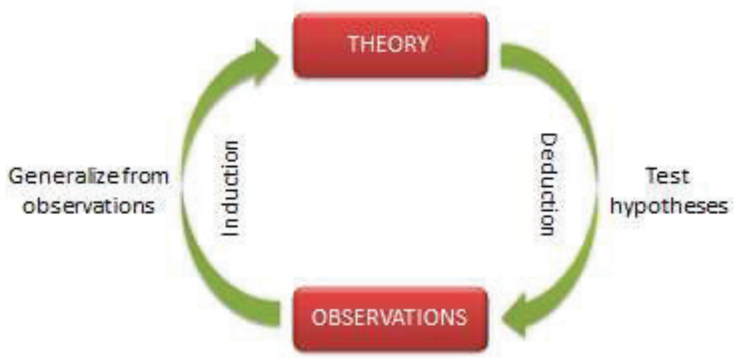

Figure 1.1. The Cycle of Research

Theory building and theory testing are particularly difficult in the social sciences, given the imprecise nature of the theoretical concepts, inadequate tools to measure them, and the presence of many unaccounted factors that can also influence the phenomenon of interest. It is also very difficult to refute theories that do not work. For instance, Karl Marx's theory of communism as an effective means of economic production withstood for decades, before it was finally discredited as being inferior to capitalism in promoting economic growth and social welfare. Erstwhile communist economies like the Soviet Union and China eventually moved toward more capitalistic economies characterized by profit-maximizing private enterprises. However, the recent collapse of the mortgage and financial industries in the United States demonstrates that capitalism also has its flaws and is not as effective in fostering economic growth and social welfare as previously presumed. Unlike theories in the natural sciences, social science theories are rarely perfect, which provides numerous opportunities for researchers to improve those theories or build their own alternative theories.

Conducting scientific research, therefore, requires two sets of skills – theoretical and methodological – needed to operate in the theoretical and empirical levels respectively. Methodological skills ("know-how") are relatively standard, invariant across disciplines, and easily acquired through doctoral programs. However, theoretical skills ("know-what") is considerably harder to master, requires years of observation and reflection, and are tacit skills that cannot be "taught" but rather learned though experience. All of the greatest scientists in the history of mankind, such as Galileo, Newton, Einstein, Neils Bohr, Adam Smith, Charles

Darwin, and Herbert Simon, were master theoreticians, and they are remembered for the theories they postulated that transformed the course of science. Methodological skills are needed to be an ordinary researcher, but theoretical skills are needed to be an extraordinary researcher!

Scientific Method

In the preceding sections, we described science as knowledge acquired through a scientific method. So what exactly is the "scientific method"? **Scientific method** refers to a standardized set of techniques for building scientific knowledge, such as how to make valid observations, how to interpret results, and how to generalize those results. The scientific method allows researchers to independently and impartially test preexisting theories and prior findings, and subject them to open debate, modifications, or enhancements. The scientific method must satisfy four key characteristics:

- *Logical:* Scientific inferences must be based on logical principles of reasoning.
- *Confirmable:* Inferences derived must match with observed evidence.
- *Repeatable:* Other scientists should be able to independently replicate or repeat a scientific study and obtain similar, if not identical, results.
- *Scrutinizable:* The procedures used and the inferences derived must withstand critical scrutiny (peer review) by other scientists.

Any branch of inquiry that does not allow the scientific method to test its basic laws or theories cannot be called "science." For instance, theology (the study of religion) is not science because theological ideas (such as the presence of God) cannot be tested by independent observers using a logical, confirmable, repeatable, and scrutinizable. Similarly, arts, music, literature, humanities, and law are also not considered science, even though they are creative and worthwhile endeavors in their own right.

The scientific method, as applied to social sciences, includes a variety of research approaches, tools, and techniques, for collecting and analyzing qualitative or quantitative data. These methods include laboratory experiments, field surveys, case research, ethnographic research, action research, and so forth. Much of this book is devoted to learning about these different methods. However, recognize that the scientific method operates primarily at the empirical level of research, i.e., how to make observations and analyze these observations. Very little of this method is directly pertinent to the theoretical level, which is really the more challenging part of scientific research.

Types of Scientific Research

Depending on the purpose of research, scientific research projects can be grouped into three types: exploratory, descriptive, and explanatory. **Exploratory research** is often conducted in new areas of inquiry, where the goals of the research are: (1) to scope out the magnitude or extent of a particular phenomenon, problem, or behavior, (2) to generate some initial ideas (or "hunches") about that phenomenon, or (3) 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 regarding during an economic recession, exploratory research may be directed at measuring the extent of citizens' dissatisfaction, understanding how such dissatisfaction is manifested, such as the frequency of public protests, and the presumed causes of such dissatisfaction, such as ineffective government policies in dealing with inflation, interest rates, unemployment, or higher taxes. Such research may include examination of publicly reported figures, such as estimates of economic indicators, such as gross domestic product (GDP), unemployment, and consumer price index, as archived by third-party sources, obtained through interviews of experts, eminent economists, or key government officials, and/or derived from studying historical examples of dealing with similar problems. This research may not lead to a very accurate understanding of the target problem, but may be worthwhile in scoping out the nature and extent of the problem and serve as a useful precursor to more in-depth research.

Descriptive research is directed at making careful observations and detailed documentation of a phenomenon of interest. These observations must be based on the scientific method (i.e., must be replicable, precise, etc.), and therefore, are more reliable than casual observations by untrained people. Examples of descriptive research are tabulation of demographic statistics by the United States Census Bureau or employment statistics by the Bureau of Labor, who use the same or similar instruments for estimating employment by sector or population growth by ethnicity over multiple employment surveys or censuses. If any changes are made to the measuring instruments, 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 ethnographic reports of gang activities among adolescent youth in urban populations, the persistence or evolution of religious, cultural, or ethnic practices in select communities, and the role of technologies such as Twitter and instant messaging in the spread of democracy movements in Middle Eastern countries.

Explanatory research seeks explanations of observed phenomena, problems, or behaviors. While descriptive research examines the what, where, and when of a phenomenon, explanatory research seeks answers to why and how types of questions. It attempts to "connect the dots" in research, by identifying causal factors and outcomes of the target phenomenon. Examples include understanding the reasons behind adolescent crime or gang violence, with the goal of prescribing strategies to overcome such societal ailments. Most academic or doctoral research belongs to the explanation category, though some amount of exploratory and/or descriptive research may also be needed during initial phases of academic research. Seeking explanations for observed events requires strong theoretical and interpretation skills, along with intuition, insights, and personal experience. Those who can do it well are also the most prized scientists in their disciplines.

History of Scientific Thought

Before closing this chapter, it may be interesting to go back in history and see how science has evolved over time and identify the key scientific minds in this evolution. Although instances of scientific progress have been documented over many centuries, the terms "science," "scientists," and the "scientific method" were coined only in the 19^{th} century. Prior to this time, science was viewed as a part of philosophy, and coexisted with other branches of philosophy such as logic, metaphysics, ethics, and aesthetics, although the boundaries between some of these branches were blurred.

In the earliest days of human inquiry, knowledge was usually recognized in terms of theological precepts based on faith. This was challenged by Greek philosophers such as Plato, Aristotle, and Socrates during the 3rd century BC, who suggested that the fundamental nature of being and the world can be understood more accurately through a process of systematic logical reasoning called **rationalism**. In particular, Aristotle's classic work *Metaphysics* (literally meaning "beyond physical [existence]") separated *theology* (the study of Gods) from *ontology* (the study of being and existence) and *universal science* (the study of first principles, upon which logic is based). Rationalism (not to be confused with "rationality") views reason as the source of knowledge or justification, and suggests that the criterion of truth is not sensory but rather intellectual and deductive, often derived from a set of first principles or axioms (such as Aristotle's "law of non-contradiction").

The next major shift in scientific thought occurred during the 16th century, when British philosopher Francis Bacon (1561-1626) suggested that knowledge can only be derived from observations in the real world. Based on this premise, Bacon emphasized knowledge acquisition as an empirical activity (rather than as a reasoning activity), and developed **empiricism** as an influential branch of philosophy. Bacon's works led to the popularization of inductive methods of scientific inquiry, the development of the "scientific method" (originally called the "Baconian method"), consisting of systematic observation, measurement, and experimentation, and may have even sowed the seeds of atheism or the rejection of theological precepts as "unobservable."

Empiricism continued to clash with rationalism throughout the Middle Ages, as philosophers sought the most effective way of gaining valid knowledge. French philosopher Rene Descartes sided with the rationalists, while British philosophers John Locke and David Hume sided with the empiricists. Other scientists, such as Galileo Galilei and Sir Issac Newton, attempted to fuse the two ideas into **natural philosophy** (the philosophy of nature), to focus specifically on understanding nature and the physical universe, which is considered to be the precursor of the natural sciences. Galileo (1564-1642) was perhaps the first to state that the laws of nature are mathematical, and contributed to the field of astronomy through an innovative combination of experimentation and mathematics.

In the 18th century, German philosopher Immanuel Kant sought to resolve the dispute between empiricism and rationalism in his book *Critique of Pure Reason*, by arguing that experience is purely subjective and processing them using pure reason without first delving into the subjective nature of experiences will lead to theoretical illusions. Kant's ideas led to the development of **German idealism**, which inspired later development of interpretive techniques such as phenomenology, hermeneutics, and critical social theory.

At about the same time, French philosopher Auguste Comte (1798–1857), founder of the discipline of sociology, attempted to blend rationalism and empiricism in a new doctrine called **positivism**. He suggested that theory and observations have circular dependence on each other. While theories may be created via reasoning, they are only authentic if they can be verified through observations. The emphasis on verification started the separation of modern science from philosophy and metaphysics and further development of the "scientific method" as the primary means of validating scientific claims. Comte's ideas were expanded by Emile Durkheim in his development of sociological positivism (positivism as a foundation for social research) and Ludwig Wittgenstein in logical positivism.

In the early 20th century, strong accounts of positivism were rejected by interpretive sociologists (antipositivists) belonging to the German idealism school of thought. Positivism was typically equated with quantitative research methods such as experiments and surveys and without any explicit philosophical commitments, while **antipositivism** employed qualitative methods such as unstructured interviews and participant observation. Even practitioners of positivism, such as American sociologist Paul Lazarsfield who pioneered large-scale survey research and statistical techniques for analyzing survey data, acknowledged potential problems of observer bias and structural limitations in positivist inquiry. In response, antipositivists emphasized that social actions must be studied though interpretive means based upon an understanding the meaning and purpose that individuals attach to their personal actions, which inspired Georg Simmel's work on symbolic interactionism, Max Weber's work on ideal types, and Edmund Husserl's work on phenomenology.

In the mid-to-late 20th century, both positivist and antipositivist schools of thought were subjected to criticisms and modifications. British philosopher Sir Karl Popper suggested that human knowledge is based not on unchallengeable, rock solid foundations, but rather on a set of tentative conjectures that can never be proven conclusively, but only disproven. Empirical evidence is the basis for disproving these conjectures or "theories." This metatheoretical stance, called **postpositivism** (or postempiricism), amends positivism by suggesting that it is impossible to verify the truth although it is possible to reject false beliefs, though it retains the positivist notion of an objective truth and its emphasis on the scientific method.

Likewise, antipositivists have also been criticized for trying only to understand society but not critiquing and changing society for the better. The roots of this thought lie in *Das Capital*, written by German philosophers Karl Marx and Friedrich Engels, which critiqued capitalistic societies as being social inequitable and inefficient, and recommended resolving this inequity through class conflict and proletarian revolutions. Marxism inspired social revolutions in countries such as Germany, Italy, Russia, and China, but generally failed to accomplish the social equality that it aspired. **Critical research** (also called critical theory) propounded by Max Horkheimer and Jurgen Habermas in the 20th century, retains similar ideas of critiquing and resolving social inequality, and adds that people can and should consciously act to change their social and economic circumstances, although their ability to do so is constrained by various forms of social, cultural and political domination. Critical research attempts to uncover and critique the restrictive and alienating conditions of the status quo by analyzing the oppositions, conflicts and contradictions in contemporary society, and seeks to eliminate the causes of alienation and domination (i.e., emancipate the oppressed class). More on these different research philosophies and approaches will be covered in future chapters of this book.

Chapter 2

Thinking Like a Researcher

Conducting good research requires first retraining your brain to think like a researcher. This requires visualizing the abstract from actual observations, mentally "connecting the dots" to identify hidden concepts and patterns, and synthesizing those patterns into generalizable laws and theories that apply to other contexts beyond the domain of the initial observations. Research involves constantly moving back and forth from an empirical plane where observations are conducted to a theoretical plane where these observations are abstracted into generalizable laws and theories. This is a skill that takes many years to develop, is not something that is taught in graduate or doctoral programs or acquired in industry training, and is by far the biggest deficit amongst Ph.D. students. Some of the mental abstractions needed to think like a researcher include unit of analysis, constructs, hypotheses, operationalization, theories, models, induction, deduction, and so forth, which we will examine in this chapter.

Unit of Analysis

One of the first decisions in any social science research is the unit of analysis of a scientific study. The **unit of analysis** refers to the person, collective, or object that is the target of the investigation. Typical unit of analysis include individuals, groups, organizations, countries, technologies, objects, and such. For instance, if we are interested in studying people's shopping behavior, their learning outcomes, or their attitudes to new technologies, then the unit of analysis is the *individual*. If we want to study characteristics of street gangs or teamwork in organizations, then the unit of analysis is the *group*. If the goal of research is to understand how firms can improve profitability or make good executive decisions, then the unit of analysis is the *firm*. In this case, even though decisions are made by individuals in these firms, these individuals are presumed to represent their firm's decision rather than their personal decisions. If research is directed at understanding differences in national cultures, then the unit of analysis becomes a *country*. Even inanimate objects can serve as units of analysis. For instance, if a researcher is interested in understanding how to make web pages more attractive to its users, then the unit of analysis is a *web page* (and not users). If we wish to study how knowledge transfer occurs between two firms, then our unit of analysis becomes the *dyad* (the combination of firms that is sending and receiving knowledge).

Understanding the units of analysis can sometimes be fairly complex. For instance, if we wish to study why certain neighborhoods have high crime rates, then our unit of analysis becomes the *neighborhood*, and not crimes or criminals committing such crimes. This is because the object of our inquiry is the neighborhood and not criminals. However, if we wish to compare different types of crimes in different neighborhoods, such as homicide, robbery, assault, and so forth, our unit of analysis becomes the *crime*. If we wish to study why criminals engage in illegal activities, then the unit of analysis becomes the *individual* (i.e., the criminal). Like, if we want to study why some innovations are more successful than others, then our unit of analysis is an *innovation*. However, if we wish to study how some organizations innovate more consistently than others, then the unit of analysis is the *organization*. Hence, two related research questions within the same research study may have two entirely different units of analysis.

Understanding the unit of analysis is important because it shapes what type of data you should collect for your study and who you collect it from. If your unit of analysis is a web page, you should be collecting data about web pages from actual web pages, and not surveying people about how they use web pages. If your unit of analysis is the organization, then you should be measuring organizational-level variables such as organizational size, revenues, hierarchy, or absorptive capacity. This data may come from a variety of sources such as financial records or surveys of Chief Executive Officers (CEO), who are presumed to be representing their organization (rather than themselves). Some variables such as CEO pay may seem like individual level variables, but in fact, it can also be an organizational level variable because each organization has only one CEO pay at any time. Sometimes, it is possible to collect data from a lower level of analysis and aggregate that data to a higher level of analysis. For instance, in order to study teamwork in organizations, you can survey individual team members in different organizational teams, and average their individual scores to create a composite team-level score for team-level variables like cohesion and conflict. We will examine the notion of "variables" in greater depth in the next section.

Concepts, Constructs, and Variables

We discussed in Chapter 1 that although research can be exploratory, descriptive, or explanatory, most scientific research tend to be of the explanatory type in that they search for potential explanations of observed natural or social phenomena. Explanations require development of **concepts** or generalizable properties or characteristics associated with objects, events, or people. While objects such as a person, a firm, or a car are not concepts, their specific characteristics or behavior such as a person's attitude toward immigrants, a firm's capacity for innovation, and a car's weight can be viewed as concepts.

Knowingly or unknowingly, we use different kinds of concepts in our everyday conversations. Some of these concepts have been developed over time through our shared language. Sometimes, we borrow concepts from other disciplines or languages to explain a phenomenon of interest. For instance, the idea of *gravitation* borrowed from physics can be used in business to describe why people tend to "gravitate" to their preferred shopping destinations. Likewise, the concept of *distance* can be used to explain the degree of social separation between two otherwise collocated individuals. Sometimes, we create our own concepts to describe a unique characteristic not described in prior research. For instance, *technostress* is a new concept referring to the mental stress one may face when asked to learn a new technology.

Concepts may also have progressive levels of abstraction. Some concepts such as a person's *weight* are precise and objective, while other concepts such as a person's *personality* may be more abstract and difficult to visualize. A **construct** is an abstract concept that is specifically chosen (or "created") to explain a given phenomenon. A construct may be a simple concept, such as a person's *weight*, or a combination of a set of related concepts such as a person's *communication skill*, which may consist of several underlying concepts such as the person's *vocabulary, syntax*, and *spelling*. The former instance (weight) is a **unidimensional construct**, while the latter (communication skill) is a **multi-dimensional construct** (i.e., it consists of multiple underlying concepts). The distinction between constructs and concepts are clearer in multi-dimensional constructs, where the higher order abstraction is called a construct and the lower order abstractions are called concepts. However, this distinction tends to blur in the case of unidimensional constructs.

Constructs used for scientific research must have precise and clear definitions that others can use to understand exactly what it means and what it does not mean. For instance, a seemingly simple construct such as *income* may refer to monthly or annual income, before-tax or after-tax income, and personal or family income, and is therefore neither precise nor clear. There are two types of definitions: dictionary definitions and operational definitions. In the more familiar dictionary definition, a construct is often defined in terms of a synonym. For instance, attitude may be defined as a disposition, a feeling, or an affect, and affect in turn is defined as an attitude. Such definitions of a circular nature are not particularly useful in scientific research for elaborating the meaning and content of that construct. Scientific research requires **operational definitions** that define constructs in terms of how they will be empirically measured. For instance, the operational definition of a construct such as *temperature* must specify whether we plan to measure temperature in Celsius, Fahrenheit, or Kelvin scale. A construct such as *income* should be defined in terms of whether we are interested in monthly or annual income, before-tax or after-tax income, and personal or family income. One can imagine that constructs such as *learning*, *personality*, and *intelligence* can be quite hard to define operationally.

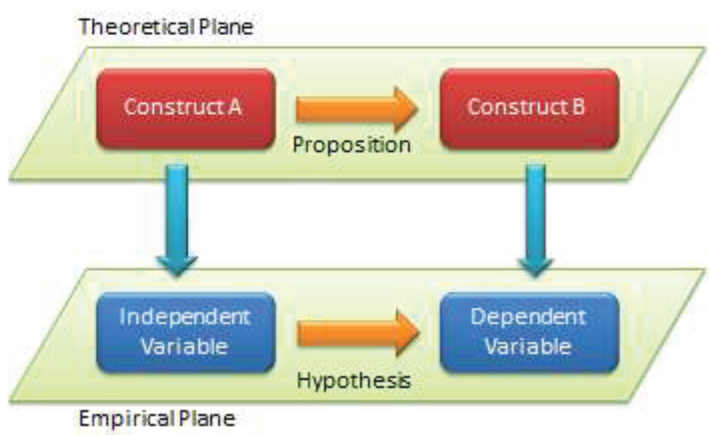

Figure 2.1. The theoretical and empirical planes of research

A term frequently associated with, and sometimes used interchangeably with, a construct is a variable. Etymologically speaking, a variable is a quantity that can vary (e.g., from low to high, negative to positive, etc.), in contrast to constants that do not vary (i.e., 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*, which is an index generated from an analytical and pattern-matching test administered to people. In this case, *intelligence* is a construct, and *IQ score* is a variable that measures the intelligence construct. Whether IQ scores truly measures one's intelligence is anyone's guess (though many believe that they do), and depending on whether how well it measures intelligence, the IQ score may be a good or a poor measure of the intelligence construct. As shown in Figure 2.1, scientific research proceeds along two planes: a theoretical plane and an empirical plane. Constructs are conceptualized at the theoretical (abstract) plane, while variables are operationalized and measured at the empirical (observational) plane. Thinking like a researcher implies the ability to move back and forth between these two planes.

Depending on their intended use, variables may be classified as independent, dependent, moderating, mediating, or control variables. Variables that explain other variables are called **independent variables**, those that are explained by other variables are **dependent variables**, those that are explained by independent variables while also explaining dependent variables are **mediating variables** (or intermediate variables), and those that influence the relationship between independent and dependent variables are called **moderating variables**. As an example, if we state that higher intelligence causes improved learning among students, then intelligence is an independent variable and learning is a dependent variable. There may be other extraneous variables that are not pertinent to explaining a given dependent variable, but may have some impact on the dependent variable. These variables must be controlled for in a scientific study, and are therefore called **control variables**.

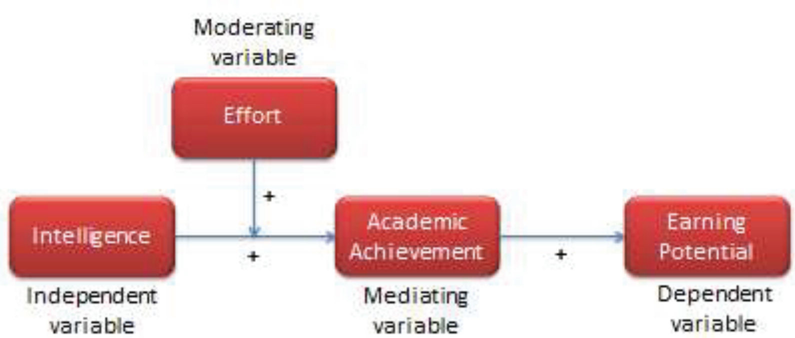

Figure 2.2. A nomological network of constructs

To understand the differences between these different variable types, consider the example shown in Figure 2.2. If we believe that intelligence influences (or explains) students' academic achievement, then a measure of intelligence such as an *IQ score* is an independent variable, while a measure of academic success such as *grade point average* is a dependent variable. If we believe that the effect of intelligence on academic achievement also depends on the effort invested by the student in the learning process (i.e., between two equally intelligent students, the student who puts is more effort achieves higher academic achievement than one who puts in less effort), then *effort* becomes a moderating variable. Incidentally, one may also view effort as an independent variable and intelligence as a moderating variable. If academic achievement is viewed as an intermediate step to higher earning potential, then *earning potential* becomes the dependent variable for the independent variable *academic achievement*, and academic achievement becomes the mediating variable in the relationship between intelligence and earning potential. Hence, variable are defined as an independent, dependent, moderating, or mediating variable based on their nature of association with each other. The overall network of relationships between a set of related constructs is called a **nomological network** (see Figure 2.2). Thinking like a researcher requires not only being able to abstract constructs from observations, but also being able to mentally visualize a nomological network linking these abstract constructs.

Propositions and Hypotheses

Figure 2.2 shows how theoretical constructs such as intelligence, effort, academic achievement, and earning potential are related to each other in a nomological network. Each of these relationships is called a proposition. In seeking explanations to a given phenomenon or behavior, it is not adequate just to identify key concepts and constructs underlying the target phenomenon or behavior. We must also identify and state patterns of relationships between these constructs. Such patterns of relationships are called propositions. A **proposition** is a tentative and conjectural relationship between constructs that is stated in a declarative form. An example of a proposition is: "An increase in student intelligence causes an increase in their academic achievement." This declarative statement does not have to be true, but must be empirically testable using data, so that we can judge whether it is true or false. Propositions are generally derived based on logic (deduction) or empirical observations (induction).

Because propositions are associations between abstract constructs, they cannot be tested directly. Instead, they are tested indirectly by examining the relationship between corresponding measures (variables) of those constructs. The empirical formulation of propositions, stated as relationships between variables, is called **hypotheses** (see Figure 2.1). Since IQ scores and grade point average are operational measures of intelligence and academic achievement respectively, the above proposition can be specified in form of the hypothesis: "An increase in students' IQ score causes an increase in their grade point average." Propositions are specified in the theoretical plane, while hypotheses are specified in the empirical plane. Hence, hypotheses are empirically testable using observed data, and may be rejected if not supported by empirical observations. Of course, the goal of hypothesis testing is to infer whether the corresponding proposition is valid.

Hypotheses can be strong or weak. "Students' IQ scores are related to their academic achievement" is an example of a weak hypothesis, since it indicates neither the directionality of the hypothesis (i.e., whether the relationship is positive or negative), nor its causality (i.e., whether intelligence causes academic achievement or academic achievement causes intelligence). A stronger hypothesis is "students' IQ scores are *positively* related to their academic achievement", which indicates the directionality 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). The signs in Figure 2.2 indicate the directionality of the respective hypotheses.

Also note that scientific hypotheses should clearly 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). Later on in this book, we will examine how to empirically test such cause-effect relationships. Statements such as "students are generally intelligent" or "all students can achieve academic success" are not scientific hypotheses because they do not specify independent and dependent variables, nor do they specify a directional relationship that can be evaluated as true or false.

Theories and Models

A **theory** is a set of systematically interrelated constructs and propositions intended to explain and predict a phenomenon or behavior of interest, within certain boundary conditions and assumptions. Essentially, a theory is a systemic collection of related theoretical propositions. While propositions generally connect two or three constructs, theories represent a *system* of multiple constructs and propositions. Hence, theories can be substantially more complex and abstract and of a larger scope than propositions or hypotheses.

I must note here that people not familiar with scientific research often view a theory as a *speculation* or the opposite of *fact*. For instance, people often say that teachers need to be less theoretical and more practical or factual in their classroom teaching. However, practice or fact are not opposites of theory, but in a scientific sense, are essential components needed to test the validity of a theory. A good scientific theory should be well supported using observed facts and should also have practical value, while a poorly defined theory tends to be lacking in these dimensions. Famous organizational research Kurt Lewin once said, "Theory without practice is sterile; practice without theory is blind." Hence, both theory and facts (or practice) are essential for scientific research.

Theories provide explanations of social or natural phenomenon. As emphasized in Chapter 1, these explanations may be good or poor. Hence, there may be good or poor theories. Chapter 3 describes some criteria that can be used to evaluate how good a theory really is. Nevertheless, it is important for researchers to understand that theory is not "truth," there is nothing sacrosanct about any theory, and theories should not be accepted just because they were proposed by someone. In the course of scientific progress, poorer theories are eventually replaced by better theories with higher explanatory power. The essential challenge for researchers is to build better and more comprehensive theories that can explain a target phenomenon better than prior theories.

A term often used in conjunction with theory is a model. A **model** is a representation of all or part of a system that is constructed to study that system (e.g., how the system works or what triggers the system). While a theory tries to explain a phenomenon, a model tries to represent a phenomenon. Models are often used by decision makers to make important decisions 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 parameters such as prior year's advertising expenses, sales, market growth, and competing products. Likewise, weather forecasters can use models to predict future weather patterns based on parameters such as wind speeds, wind direction, temperature, and humidity. While these models are useful, they may not necessarily explain advertising expenditure or weather forecasts. 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. An advertising expenditure model may be a descriptive model. Predictive models (e.g., a regression model) allow forecast of future events. Weather forecasting models are predictive models. Normative models are used to guide our 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. Recall from Chapter 1 that **deduction** is the process of drawing conclusions about a phenomenon or behavior based on theoretical or logical reasons and an initial set of premises. As an example, if a certain bank enforces a strict code of ethics for its employees (Premise 1) and Jamie is an employee at that bank (Premise 2), then Jamie can be trusted to follow ethical practices (Conclusion). In deduction, the conclusions must be true if the initial premises and reasons are correct.

In contrast, **induction** is the process of drawing conclusions based on facts or observed evidence. For instance, if a firm spent a lot of money on a promotional campaign (Observation 1), but the sales did not increase (Observation 2), then possibly the promotion campaign was poorly executed (Conclusion). However, there may be rival explanations for poor sales, such as economic recession or the emergence of a competing product or brand or perhaps a supply chain problem. Inductive conclusions are therefore only a hypothesis, and may be disproven. Deductive conclusions generally tend to be stronger than inductive conclusions, but a deductive conclusion based on an incorrect premise is also incorrect.

As shown in Figure 2.3, inductive and deductive reasoning go hand in hand in theory and model building. Induction occurs when we observe a fact and ask, “Why is this happening?” In answering this question, we advance one or more tentative explanations (hypotheses). We then use deduction to narrow down the tentative explanations to the most plausible explanation based on logic and reasonable premises (based on our understanding of the phenomenon under study). Researchers must be able to move back and forth between inductive and deductive reasoning if they are to post extensions or modifications to a given model or theory, or built better ones, which are the essence of scientific research.

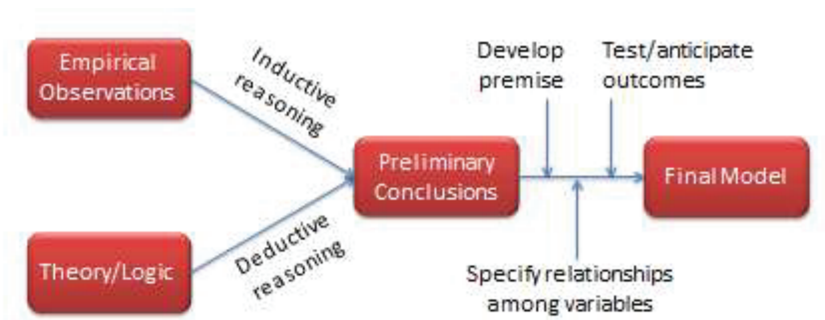

Figure 2.3. The model-building process

Chapter 3

The Research Process

In Chapter 1, we saw that scientific research is the process of acquiring scientific knowledge using the scientific method. But how is such research conducted? This chapter delves into the process of scientific research, and the assumptions and outcomes of the research process.

Paradigms of Social Research

Our design and conduct of research is shaped by our mental models or frames of references that we use to organize our reasoning and observations. These mental models or frames (belief systems) are called **paradigms**. The word "paradigm" was popularized by Thomas Kuhn (1962) in his book *The Structure of Scientific Revolutions*, where he examined the history of the natural sciences to identify patterns of activities that shape the progress of science. Similar ideas are applicable to social sciences as well, where a social reality can be viewed by different people in different ways, which may constrain their thinking and reasoning about the observed phenomenon. For instance, conservatives and liberals tend to have very different perceptions of the role of government in people's lives, and hence, have different opinions on how to solve social problems. Conservatives may believe that lowering taxes is the best way to stimulate a stagnant economy because it increases people's disposable income and spending, which in turn expands business output and employment. In contrast, liberals may believe that governments should invest more directly in job creation programs such as public works and infrastructure projects, which will increase employment and people's ability to consume and drive the economy. Likewise, Western societies place greater emphasis on individual rights, such as one's right to privacy, right of free speech, and right to bear arms. In contrast, Asian societies tend to balance the rights of individuals against the rights of families, organizations, and the government, and therefore tend to be more communal and less individualistic in their policies. Such differences in perspective often lead Westerners to criticize Asian governments for being autocratic, while Asians criticize Western societies for being greedy, having high crime rates, and creating a "cult of the individual." Our personal paradigms are like "colored glasses" that govern how we view the world and how we structure our thoughts about what we see in the world.

Paradigms are often hard to recognize, because they are implicit, assumed, and taken for granted. However, recognizing these paradigms is key to making sense of and reconciling differences in people' perceptions of the same social phenomenon. For instance, why do liberals believe that the best way to improve secondary education is to hire more teachers, but conservatives believe that privatizing education (using such means as school vouchers) are