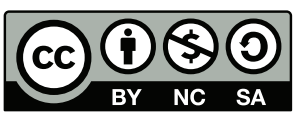
RESEARCH METHODS

For Business and Marketing

GEORGE SELF

January, 2019 – Edition 1

George Self: *Research Methods,* For Business and Marketing, January, 2019



FOREWORD

I have taught BASV 316, *Introductory Methods of Analysis*, on-line 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 in a practical way so students could use the lessons learned in their own research projects. I found an excellent book but over the years the cost of that book increased to the point that I felt like it was an unfair burden on students.

I began by looking for an acceptable "open source" book since authors make those available to students free of charge and I could modify the book to meet my own objectives. I could not find any that were focused on business research though I tried for several years—and keep looking to this day. I did, though, find a few open source books about research in the social and psychological sciences that were reasonably close to what I needed. So, I modified those books to emphasize business research and then provided my work to students free of charge.

Bhattacherjee[2], Blackstone[3], and Price[23] all released books about research that formed the major sources for this book. Those books are all open source and published under a Creative Commons license that permitted me to copy and modify them.

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.

While the book is useful in its current form, I will continually update it based on emerging trends in research.

This book is published under a Creative Commons Attribution-NonCommercial-ShareAlike license, just like the books that provided its foundation. The source is available at my GitHub account: `http://bit.ly/2xIjzXL`. It is my hope that students can use this book to learn about business research and other instructors can modify and use it for their own classes.

— George Self

BRIEF CONTENTS

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LISTINGS

Part I

BACKGROUND

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

1

RESEARCH DESIGN

1.1 INTRODUCTION

A research project always starts with someone who is curious about something observed. A grocery store clerk notices that cereal in red boxes seems to come through the checkout line more frequently than cereal in blue boxes and wonders why. An economist notices that during certain times of the year the motels seem to be full when 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.

Researchers typically "start where they are," an idea eloquently described by Kristin Esterberg[10], 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." Whether it was thinking about a question they had pondered for some time, identifying a question about their own interests and hobbies, or taking a look at patterns in their everyday life, every researcher identifies a research question that was interesting and then collect and analyze data that helped answer that question. This chapter concerns creating a worthwhile question and planning a research project. Later chapters are devoted to 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. An honest introspection is in order as they ask themselves what they may already believe about the topic and whether they believe that their perspective is the only valid one. If they determine that they have a preconceived notion that they think is the wisest perspective then that could be a problem. Researchers must also consider how they would react if their research proves them wrong about some believe. If they would be comfortable examining, and perhaps changing, what may be cherished notions based on their research then that is one thing, but if they would deny the research, hide the outcomes, or even change the data, then 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 intentionally study a topic that evokes strong feelings. Sociology professor Kathleen Blee[4] has taken this route in her research. She studies hate movement par-

ticipants and the people whose racist ideologies she does not share. Blee's research is successful because she was willing to report her findings and observations honestly, even those with which she may have personally taken strong 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, some are more prone to creating bias than others. For example, researchers may know of a topic from family history, a television program, or through casual conversations with friends. These could all introduce bias in the researcher's mind and it is important that researchers think about how they know what they know to help identify and correct biases that they may bring to the research project.

The purpose of this chapter is 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, in fact, it may not even be the best way for a given research project, but it would work as a starting point for many research investigations.

1.2 FIRST CONSIDERATIONS

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

1.2.1 *Exploration, Description, Explanation*

Three general types of research are exploratory research, descriptive research, and explanatory research. Researchers conducting exploratory research are typically at the early stages of examining a topic. Exploratory research is often designed to determine the feasibility a more extensive study. Descriptive research describes or defines a particular phenomenon. As an example, an economist publishing the gasoline prices in various parts of a city is conducting descriptive research. Finally, research that answers "why" questions is 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.

1.2.2 *Is The Topic Empirical?*

An empirical topic is one that can be investigated by observation or experience rather than one that concerns only opinions or theories. As an example, if a researcher investigated the cost of health care in order to answer the question, "What is the best way to fund health care?" that would not make an appropriate empirical study since the definition of "best way" is nebulous. The question, though, could be answered if it were re-framed a bit to simply how health care was funded then that 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[17]. Could a researcher determine if Spongebob is immoral? Of course not; this is an ethical question, not empirical. A researcher could gather facts about people's opinions concerning Spongebob and even interview the creators of the program to see what they intended, but answering the question of morality belongs to the world of ethicists or theologians, not business researchers.

1.2.3 *The Research Question*

Once a researcher finds a topic that is empirical then 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. To say that the research question is "child-free adults" or "movies" would not be a question.

2. Focused. A research question must be focused on one topic of interest and not something that is 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. Rather, 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 really nothing to research.

1.2.4 *Hypotheses*

The purpose of positivist research[1] is to test a theory and in order to do that a researcher must create a hypothesis that is derived from the theory. A hypothesis is a statement, sometimes causal, describing a researcher's expectation regarding the anticipated 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 increases trust and commitment[20]. A researcher may hypothesize that brand loyalty increases due to positive outcomes from social exchanges and then design some sort of investigation to test that hypothesis.

Sometimes, researchers hypothesize that a relationship will take a specific direction so an increase in one variable might lead to an increase in 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 hypotheses. A statement that bold implies that a relationship has been shown to exist with absolute certainty and that there is no chance that there are conditions under which the hypothesis fail. 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 will 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 to one another.

1.2.5 *Feasibility*

In Chapter **??**, **??**, ethical considerations were discussed that may make some research projects unfeasible. Certainly, no researcher is going to design an experiment where a business enterprise would intentionally injure children in order to test some theory. There are,

1 Researchers engaged in interpretive projects may not start with a hypothesis, but one would likely be developed as the research project proceeded.

though, a few practical matters 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 common as children's behavior concerning snacks can raise interesting research issues. For example, Marshall, O'Donohoe, and Kline[21] conducted a study where they interviewed $8-11$ year-old children to explore their exposure to food advertising and subsequent snack preference. While it is, generally, no trouble to find 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 intentionally lie, their responses to interview questions are almost certainly influenced by the fact than 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 the world of a child 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[27] conducted surveys to determine the demographic characteristics of shoppers who were purchasing organic food in gentrified neighborhood. Bridge and Dowling[6] 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 and record the demographics of the families who are shopping, 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 becomes unfeasible.

The funding available for a study is also potentially limiting. Medical research often requires the use of very expensive 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 spend time interviewing shoppers. If the research project involves a team of survey-takers fanning out over a wide geographic ares over several weeks then 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 small, but quantifiable, cost that must be met.

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

1.2.6 *Idiographic or Nomothetic?*

In general terms, research can be described as idiographic or nomothetic, as described by Joseph Ponterotto[22]. These terms derive from Kantian philosophy and are frequently found in research reports, especially in 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 be applicable to a wider application. Idiographic research sacrifices breadth of application for deeper, richer understanding of a single case. Many case studies are idiographic in the sense that only a single individual or location is studied and applicability beyond that case is not reasonable. Much of the small business research being done is idiographic in nature.
- Nomothetic. This term comes from the Greek *nomos*, which refers to the traditional social norm. The goal of nomothetic research is 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 begin to improve but that does not guarantee that a specific business will benefit.

1.2.7 *Applied or Basic?*

The contribution that researchers hope to make to the body of knowledge depends on whether they are conducting applied research or basic research. Applied research can be immediately applied to a specific case. Applied research would help a small business owner made changes in advertising that would improve the number of customers entering the store. Basic research, on the other hand, is designed to create, or validate, theories and would be useful to a legislator considering some change in the business laws of a state.

1.2.8 *Units of Analysis*

Another point to consider when designing a research project, and 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 in the course of data collection and what can later be said about those observations. A unit of observation is the item (or items) that are actually observed, measured, or collected in the course of of the research study. A unit of analysis is the entity that is reported at the end of the study, or the "main focus" of the study. In a given study, the unit of observation might be the same as the unit of analysis, but that is not always the case. What is required, however, is for researchers to be clear about how they define their units of observation and analysis, both to themselves and to their audiences.

As an example, one common unit of analysis is an individual. A research project designed to look at the shopping habits of people 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 particular 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[7] and Philip Kotler[18].

A second common 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. As examples, 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[29][14].

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 includes entities like corporations, colleges and universities, and even night clubs.

As examples, researchers might study the economic impact of globalization or how unions influence the behavior of industry leadership, as researched by Diana Hechavarria[12] and Randall Schuler[25].

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, then, the

unit of observation would be a document while the unit of analysis is the business. This is also a good 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:

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

There are also many topics that could be studied from more than one level of analysis, though that would become a more complex study. As an example, Kuruvilla and Ranganathan researched the way micro and macro human resource policies influenced economic development strategy in India[19].

1.3 THE RESEARCH PROCESS

Broadly speaking, research methods can be grouped into two broad categories: positivism and interpretivism.

1.3.1 *Positivism*

- Goal: theory testing
- Methods: laboratory experiments and surveys
- Approach: deductive, starts from theory and generates empirical data to test the theory
- Data: quantitative in nature: numeric
- Analysis: statistical

1.3.2 *Interpretivism*

- Goal: theory building
- Methods: action research and ethnography
- Approach: inductive, starts from observations and generates theory
- Data: qualitative in nature: textural
- Analysis: coding

1.3.3 *Iterative Design*

At its core, all scientific research is an iterative process of observation, rationalization, and validation. In the observation phase, researchers observe a natural or social phenomenon, event, or behavior of interest. 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 a process of data collection and analysis and that often leads to a modification of the initial theory. However, research designs vary based on whether the researcher starts at observation and attempts to generate a theory (interpretive research) or starts at a theory and attempts to validate it with observations (positivist research).

Most traditional research tends to be positivist in nature. Figure 1 provides a schematic view of such a research project. This figure depicts a series of activities to be performed, categorized into three phases: exploration, research design, and research execution.

This generalized design does not fit all research and it should be modified to fit the needs of a specific project.

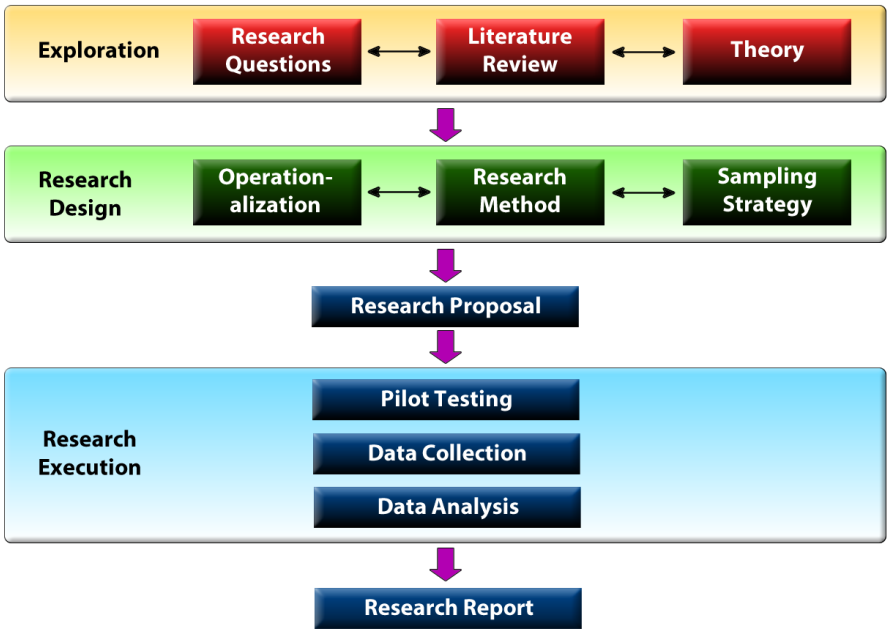


Figure 1: Positivist Research Process

1. **Exploration**. The first phase is exploration, which includes exploring and selecting research questions for further investigation, examining the published literature in the area of inquiry to understand the current state of knowledge in that area, and identifying theories that may help answer the research questions of interest. The diagram makes it clear that these three steps often run concurrently and researchers typically shift back-

and-forth between them as needed. For example, a literature review is designed to uncover pertinent theories but finding those theories may lead to further literature review.

- 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 can high school students become more creative, and why do some people commit terrorist acts. More interesting research questions are those that appeal to a broader population (e.g., "how can firms innovate" is a more interesting research question than "how can Chinese firms innovate in the service-sector"), address real and complex problems (in contrast to hypothetical problems), and where the answers are not obvious. Narrowly focused research questions (often with only a *yes/no* answer) tend to be less useful and interesting, and generally lead to unpublishable research findings.
- The next 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 of inquiry, 2) to identify key authors, articles, theories, and findings in that area, and 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 then manually browse through each article, or at least its abstract, to determine the suitability of that article for a detailed review. Literature reviews should be reasonably complete and not restricted to only a few journals, a few years, or a specific methodology. Reviewed articles may be summarized in the form of tables and can be further structured using organizing frameworks such as a concept matrix. A well-conducted literature review should indicate whether the initial research questions have already been addressed in the literature (which would obviate the need to study them again), whether there are newer or more interesting research questions available, and whether the original research questions should be modified or changed in light of findings of the literature review. The review can also provide some intuitions or potential answers to the questions of interest and/or 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 eas-

Literature review is commonly done using searches in online databases.

iest to find is the abstract, which is a short paragraph at the beginning of an article that summarizes the research question, methods used to answer the question, and key 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 the abstract, reading the discussion section is usually the next most productive. Finally, the methodology section may include important clues about a productive way to approach a research project.

- Since positivist research involves theory-testing, 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 results that are meaningful. In positivist research, theories can be used as the logical basis for postulating hypotheses needed in a later step. Obviously, 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.

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

    - Operationalization is the process of designing precise measures for abstract theoretical constructs. This is a major 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 see if there are existing measures that can be modified to measure the constructs of interest. If such measures are not available or if they reflect a different conceptualiza-

tion than that intended by the researcher then new instruments may have to be designed. This can easily be a long and laborious process, with multiple rounds of pretests and modifications before the newly-designed instrument can be accepted as "scientifically valid."

- Simultaneously with operationalization, the researcher must also decide what research method to employ for collecting data that will address the research question. Informing this stage of the process are the answers to philosophical questions like whether the research be exploratory, descriptive, or explanatory; will the approach be interpretive or positivist; is the goal to have some direct application or contribute more generally to the field; and what unit of analysis and observation will be used. Research methods may include experimentation, surveys, case studies, and others, or combinations of several methods in order 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.
- 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 a later chapter.

3. **Proposal**. At this stage, it is often a good idea to write a research proposal detailing all of the decisions made in the preceding stages of the research process and the rationale behind each decision. This multi-part proposal should address the research questions being studied and why, the current state of knowledge, theories and hypotheses to be tested, how the constructs will be measured, the research method to be employed and why, and sampling strategy. Funding agencies typically require a detailed proposal in order for them to select which to fund. Even if funding is not sought for a research project, a proposal may serve as a useful 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 is collected in a research study.

4. **Research Execution**. Having decided who 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 includes pilot testing the measurement instruments, data collection, and data analysis.

- Pilot testing is an often overlooked but extremely important part of the research process. It helps detect potential problems in the research design and instrumentation (e.g., whether survey questions are intelligible to the targeted sample), and to ensure that the measurement instruments used in the study are reliable and valid measures of the constructs of interest. 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.
- Next comes the actual collection of data. At this phase of the investigation the researcher would conduct surveys, visit field sites, interview subjects, read corporate documents, or generate whatever other data is specified by the plan.
- Following data collection, the data are analyzed and interpreted for the purpose of drawing conclusions regarding the research questions. 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).

5. **Research Report**. The final phase of research involves preparing the final research report documenting the entire research process and its findings in the form of a research paper, dissertation, or monograph. The report should outline in detail all the choices made during the research process (e.g., theory used, constructs selected, measures used, research methods, sampling, etc.) and why, as well as the outcomes of each phase of the research process. The research process must be described in sufficient detail so as 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.

1.4 MIXED METHODS

Up to this point, the research design has been treated as if it is an either/or proposition. Either a research project is positivist and numeric data are gathered or it is interpretative and textural data are gathered. In truth, researchers do not necessarily have to choose one approach over another. In fact, some of the most highly regarded business and marketing investigations combine approaches in an effort to gain the most complete understanding of their topic possible. Using

a combination of 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 were 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[2]. 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 larger number of resources, time, and expertise to complete. Also, along with gaining the benefit of the strengths of each type of research is the potential of the combined weaknesses becoming a problem.

1.5 COMMON RESEARCH MISTAKES

The research process is fraught with problems and pitfalls and novice researchers often find after investing substantial amounts of time and effort into a research project that their research questions were not sufficiently answered, or that the findings were not interesting enough, or that the research was not of "acceptable" scientific quality. Such problems typically result in research papers being rejected by journals.

- Insufficiently motivated questions. Often times, 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 time and effort on the researcher's part, the researcher must be certain (and be able to convince others) that the research questions they seek to answer in fact deal with real problems (and not hypothetical problems) that affect a substantial portion of a population and has not been adequately 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, it is possible that popular interest in these fads may die down by the time the research is completed and submitted for publication.

2 For information about using a mixed method type of research design, see John Brewer[5] and Charles Teddlie[28].

A better strategy may be to study "timeless" topics that have persisted through the years.

- Unresearchable problems. Some research problems may not be answered adequately based on observed evidence alone or currently accepted methods and procedures. Such problems are best avoided. However, some unresearchable, ambiguously defined problems may be modified or fine tuned into well-defined and useful researchable problems.

- Favored research methods. Many researchers have a tendency to recast a research problem so that it is amenable to their favorite research method (e. g., survey research). This is an unfortunate trend. Research methods should be chosen to best fit a research problem, and not the other way around.

- Blind data mining. Some researchers have the tendency to collect data first (using instruments that are already available), and then figure out what to do with it. In reality, data collection is only one step in the long process of planning, designing, and executing research. In fact, a series of other activities are needed in a research process prior to 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, and, particularly, for the lack of interesting research questions.

- Ecological fallacy. This occurs when claims about one lower-level unit of analysis are made based on data from some higher-level unit of analysis. In many cases, this occurs when claims are made about individuals, but only group-level data have been gathered.

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

1.6 RESEARCH DESIGNS

As noted on page 10, research designs can be classified into two categories, 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. Popular examples of positivist designs include experimental (both laboratory and field), 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 at the same time. 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.

1.6.1 *Experimental*

Experimental studies are those that are intended to test cause-effect relationships (hypotheses) in a tightly controlled setting by separating the cause from the effect in time, administering the cause to one group of subjects (the "treatment group") but not to another group ("control group"), and 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 certain ailment then a random sample of people afflicted with that ailment is found and 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 over a period of time 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 but the scientists actually administering the treatment protocol are not sure if a specific subject is receiving 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 are sure if they are 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 setting such as at a university (laboratory experiments) or in a field settings such as in an organization where the phenomenon of interest is actually occurring (field experiments). 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[3], but those from field experiments tend to be stronger in external validity.

3 Validity is more thoroughly defined in Chapter 2, page 25

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

1.6.2 *Surveys*

Field surveys are non-experimental designs that do not control for 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), while in longitudinal field surveys, dependent variables are measured at a later point in time than the independent variables. The strengths of field surveys are their external validity (since data is collected in field settings), their ability to capture and control for a large number of 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 true response) which further decreases internal validity.

1.6.3 *Secondary Data Analysis*

Secondary data analysis is analysis of data that has 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 is in contrast to 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 is available at a level of analysis suitable for answering the research questions. The limitations of this design are that the data may not have been collected in a systematic or scientific manner and hence unsuitable for scientific research. Also, since the data were collected for a presumably differ-

ent purpose, they may not adequately address the research questions of interest to the researcher. Finally, interval validity is problematic if the temporal precedence between cause and effect is unclear.

1.6.4 *Case Research*

Case research[4] is an in-depth investigation of a problem in one or more real-life settings (case sites) over an extended period of time. 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 in nature, 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.

1.6.5 *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 a period of 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. Internal validity cannot be established due to lack of controls and the findings may not be generalized to other settings because of small sample size. Hence, focus groups are not generally used for explanatory or descriptive research but are suited for exploratory research projects.

1.6.6 *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 an organizational member

4 It is important to keep in mind that case research is not the same as a business class discussing a classic Harvard Case Study. Case research is the process of actually going to a site, gathering data, and analyzing that data.

embedded within a social context, such as an organization, who initiates an action, such as new organizational procedures or new technologies, in response to a real problem, such as declining profitability or operational bottlenecks. The researcher's choice of actions must be based on theory, which should explain 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, and 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, but it is also subject to researcher bias and subjectivity, and the generalizability of findings is often restricted to the context where the study was conducted.

1.6.7 *Ethnography*

Ethnography is an interpretive research design inspired by anthropology that emphasizes the concept that a phenomenon must be studied within the context of its culture. The researcher is deeply immersed in a certain culture over an extended period of time (a few months to several years) and during that period engages, observes, and records the daily life of the studied culture. The ultimate goal is a theory about the evolution and 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.

1.7 SELECTING THE RESEARCH DESIGN

Researchers tend to select designs that they are most comfortable with and feel most competent to handle; but, ideally, the choice should depend on the nature of the research phenomenon being studied. 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 certain research phenomenon, a focus group (for individual unit of analysis) or a case study (for 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 useful. If a literature review finds competing theories then 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 both quantitative and qualitative data using a combination of techniques such as questionnaires, interviews, observations, documents, or secondary data. For example, even in a highly structured survey questionnaire intended to collect quantitative data, the researcher may leave room for a few open-ended questions to collect qualitative data that may generate unexpected insights not otherwise available from the structured quantitative data alone. Likewise, while case research employ mostly face-to-face interviews to collect most qualitative data, the potential and value of collecting quantitative data using a concurrent survey should not be ignored. As an example, in a study of organizational decision-making processes, the case interviewer could record numeric quantities such as how many months it took to make certain organizational decisions, how many people were involved in that decision process, and how many alternatives were considered, and those data can provide valuable insights not otherwise available from interviewees' narrative responses. Irrespective of the specific research design employed, the goal of the researcher 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.

1.8 SUMMARY

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Aenean commodo ligula eget dolor. Aenean massa. Cum sociis natoque penatibus et

Part II

QUANTITATIVE METHODS

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

2 DEFINING AND MEASURING CONCEPTS

2.1 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 important 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 in a precise and measurable way. 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, but there are some basic tenants on which most researchers agree when it comes to measurement.

2.1.1 *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[8], the price elasticity of gasoline[15], employee turnover[13], and automobile "lemons"[1]. Each of these topics required measurements of various types and researchers had to determine the best way to do that. As you might have guessed, researchers will measure just about anything that they have an interest in investigating.

In 1964, philosopher Abraham Kaplan wrote what has since become a classic work in research methodology, *The Conduct of Inquiry*[16]. 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. They are terms that "lend themselves to easy and confident verification." 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.

Photo by lindsay Cotter on Unsplash

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 was only observed indirectly. Birthplace might be another indirect observable. Researchers can ask study participants where they were born, but chances are good that they will not directly observe any of those people being born in the locations they report.

Sometimes the measures that we are interested in are more complex and more abstract than observational terms or indirect observables. 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 very difficult to measure. In both cases, ethnocentrism and bureaucracy, the theoretical notions represent ideas whose meaning is known but the measurement of the concept may be nearly impossible. Kaplan referred to these more abstract things as constructs. Constructs are "not observational either directly or indirectly" but they can be defined based on observables.

2.1.2 *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.

As an 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. Perhaps one method of coping is determined to be more effective than others so the report may contain a recommendation that future research be conducted that measures just that one method of coping. The point is that measurement considerations are important throughout the research project.

The measurement process could also involve multiple stages. Starting with identifying and defining key terms to determining how to observe and measure them to assessing the quality of the measure-

ments, 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 particular types of data elements.

Data types are discussed on page ??.

2.2 CONCEPTUALIZATION

One of the first steps in the measurement process is conceptualization, which is defining 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 the notion or image that is conjured up when the researcher thinks of some cluster of 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" and that definition may well change over time, from culture to culture, and even from individual to individual, as explained by George Mosse[11]. This is why 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 42^{d} President of the United States, famously stated "It depends upon what the meaning of the word 'is' is."[1] 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 the process of measurement called conceptualization, which involves writing out clear, concise definitions for key concepts. Brainstorming may help to 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 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 merely applying a definition from a dictionary, it requires careful consideration and evaluating alternative concepts.

One important decision while conceptualizing constructs is specifying whether they are unidimensional or multidimensional. Unidimensional constructs are those that are expected to have a single un-

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

derlying dimension and can be measured using a single measure or test. Examples include simple constructs such as a person's weight, wind speed, and even complex constructs like self-esteem (if self-esteem is conceptualized as consisting of a single dimension, which of course, may be unrealistic). Multidimensional constructs consist of two or more underlying dimensions. For instance, if a person's academic aptitude is conceptualized as consisting of two dimension, mathematical and verbal ability, then academic aptitude is a multidimensional construct. Each of the underlying dimensions in this case must be measured separately by using different tests for mathematical and verbal ability, and then combine the two scores, possibly in a weighted manner, to create an overall value for the academic aptitude construct.

Before moving on to the next steps in the measurement process, it would be wise to consider one of the dangers associated with conceptualization. While it is important to consult prior scholarly definitions of key concepts, it would be wrong to assume that those definitions are any more real than whatever current definitions are generated by the researcher. It would also be wrong to assume that just because definitions exist for some concept that the concept itself exists beyond some abstract idea. This idea, assuming that abstract concepts exist in some concrete way is 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 the term has been conceptualized by others. But they should also question past conceptualizations. Today's conceptualization of "family" would be very different from one that was used a hundred 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 difficult to navigate through everyday living. But at the same time, it is important to remember that a society has assigned those definitions and that they are no more real than any other, alternative definition a researcher might choose to assign.

2.3 OPERATIONALIZATION

Once a theoretical construct is defined, indicators for measuring the construct are defined in a process called operationalization. For instance, if an unobservable theoretical construct such as socioeconomic status is defined as the level of family income then it can be operationalized using an indicator that asks respondents the question: what is your annual family income? Given the high level of subjec-

tivity and imprecision inherent in social science constructs, most (except a few demographic constructs such as age, gender, education, and income) are measured using multiple indicators.

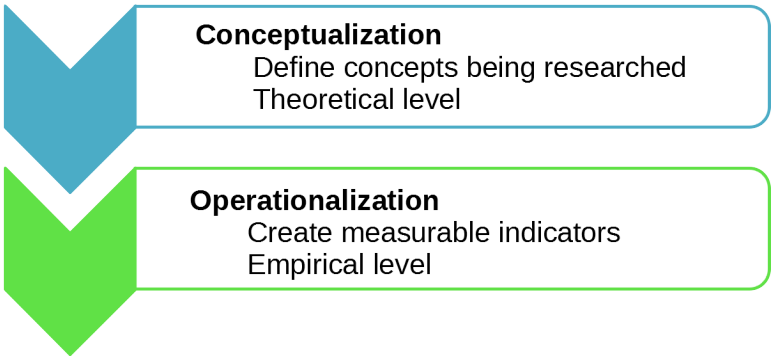


Figure 2: Conceptualization Vs. Operationalization

Indicators operate at the empirical level in contrast to constructs, which are conceptualized at the theoretical level. The combination of indicators at the empirical level representing a given construct is called a variable, and those may be independent, dependent, mediating, or moderating, depending on how they are employed in a research study. Also each indicator may have several attributes (or levels) and each attribute represent a value. For instance, a "gender" variable may have two attributes: male or female. 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 can be analyzed using techniques like regression or structural equation modeling while qualitative data uses techniques like coding. Note that many variables in business research are qualitative, even when represented in a quantitative manner. For instance, a customer satisfaction indicator with five attributes: strongly dissatisfied, somewhat dissatisfied, neutral, somewhat satisfied, and strongly satisfied, can assign the numbers $1-5$ respectively for these five attributes, so that we can use sophisticated statistical tools for quantitative data analysis. However, note that the numbers are only labels associated with respondents' personal evaluation of their own satisfaction, and the underlying variable (satisfaction) is still qualitative even though it is represented numerically.

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, then attending religious services may be a reflective indicator of religiosity. 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 re-

ligiosity is defined as composed of a belief dimension, a devotional dimension, and a ritual dimension, then indicators chosen to measure each of these different dimensions will be considered formative indicators. Unidimensional constructs are measured using reflective indicators (even though multiple reflective indicators may be used for measuring abstruse constructs such as self-esteem), while multidimensional constructs are measured as a formative combination of the multiple dimensions, even though each of the underlying dimensions may be measured using one or more reflective indicators.

It is important 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 you will use for data collection. A survey implies one way of measuring concepts while field research implies a very different way. The data-collection strategy employed will play a major role in shaping how concepts are operationalized.

2.4 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 don't exist in reality), and multi-dimensional (in which case, we have the added problem of identifying their constituent dimensions). Hence, it is not adequate just to measure constructs using any scale, the scales must be tested to ensure that:

1. they measure the construct consistently and precisely (i.e., the scales are "reliable") and
2. they actually 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 are evaluated in scientific research. A measure can be reliable but not valid if it is measuring consistently but it is the wrong construct. Likewise, a measure can be valid but not reliable if it is measuring the right construct but not doing so in a consistent manner. Using the analogy of a shooting target, as shown in Figure 3, a measure that is both reliable and valid is like a group that is tightly clustered 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 group that is widely scattered and off-center.

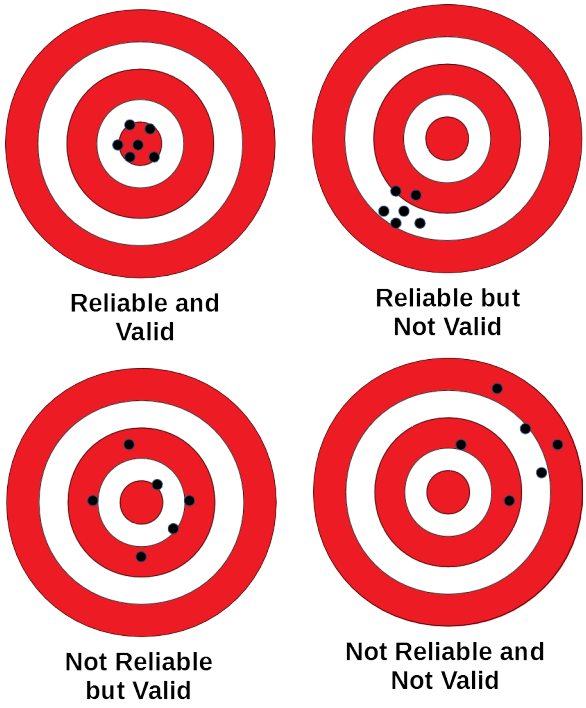


Figure 3: Reliability Analogy

2.4.1 *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."[9]

Any score obtained by a measuring instrument (the observed score) is composed of both 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 true 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, the if the bathroom scale mentioned above constantly read five pounds heavy that would be an error, but it would be one that is consistent and could be corrected in the research

data. That is an example of a systematic error. The second type of error is a random error. If the bathroom scale were accurate but the person reading it one day read 151 and the next as 149 then that would be a random error. Random errors cannot be corrected but tend to cancel out due to the random nature of the error (sometimes the reading will be a bit high and other times low), especially if there are many data points.

Unreliable measurements in business research could be for several reasons. One is the researcher's subjectivity. For example, if employee morale in a firm is being measured by watching whether the employees smile at each other, whether they make jokes, and so forth, then different observers may infer different measures of morale if they are watching the employees on a very busy 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 levels of morale on the same day, depending on what they view as a joke and what is not. "Observation" is a qualitative measurement technique.

Sometimes, reliability may be improved by using quantitative measures. Counting the number of grievances filed over one month as a measure of (the inverse of) morale. Of course, grievances may or may not be a valid measure of morale, but it is less subject to human subjectivity, and therefore more reliable.

A second source of unreliable observation is asking imprecise or ambiguous questions. For instance, if people are asked to report their salary some may state a monthly salary, some an annual salary, and some even an hourly wage. Thus, the resulting observations will be divergent and unreliable.

A third source of unreliability is asking questions about issues that respondents are not very familiar with or care about, such as asking an American college graduate about Canada's relationship with Slovenia or asking a Chief Executive Officer to rate the effectiveness of his company's technology strategy (which was likely delegated to a technology executive).

To improve reliability, start by replacing subjective data collection techniques (observation) with those that are more objective (questionnaire), ask respondents only questions that they may know or care about, avoid ambiguous items (e.g. clearly indicate annual salary), 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.

- **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, this is assessed in a pilot study and can be done in two ways, depending on the level of measurement being used. If the measure is categorical, a set of all categories is defined, raters check

off which category each observation falls in, and the percentage of agreement between the raters is used as an estimate of inter-rater reliability. For instance, if there are two raters rating 100 observations into one of three possible categories, and their ratings match for 75% of the observations, then inter-rater reliability is 0.75. If the measure is interval or ratio scaled (e.g., classroom activity is being measured once every five minutes by two raters on one to seven scale), then 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 is a measure of 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 is the time gap, the greater is the chance that the two observations may change during this time (due to random error), and the lower will be the test-retest reliability.

- **Split-half reliability**. Split-half reliability is a measure of consistency between two halves of a construct measure. For instance, if you have a ten-item measure of a given construct, randomly split those ten items into two sets of five (unequal halves are allowed if the total number of items is odd), and administer the entire instrument to a sample of respondents. Then, calculate the total score for each half for each respondent, and the correlation between the total scores in each half is a measure of split-half reliability. The longer the instrument, the more likely it is that the two halves of the measure will be similar (since random errors are minimized as more items are added), and hence, this technique tends to systematically overestimate the reliability of longer instruments.

- **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 rate those items in a similar manner is a reflection of 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*.

2.4.2 *Validity*

Validity is concerned with the meaningfulness of research results. In brief, does the research actually measure what it was purported to measure? For example, does the Scholastic Aptitude Test (SAT) actually predict the likelihood of a high school student successfully completing college?[9] There are numerous types of validity found in the literature, but 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).

2.4.2.1 *Measurement Validity*

The *theoretical* assessment of validity focuses on how well an abstract construct is translated into an operational measure, which is called translational validity, and 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[26].

The *empirical* assessment of validity examines how well a given measure relates to one or more external criterion, based on empirical observations. This type of validity is called criterion validity, which is 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-related validity examines whether a given measure behaves the way it should, given the theory of that construct. The distinction between theoretical and empirical assessment of validity is illustrated in Figure 4. However, both approaches are needed to adequately ensure the validity of measures in business research.

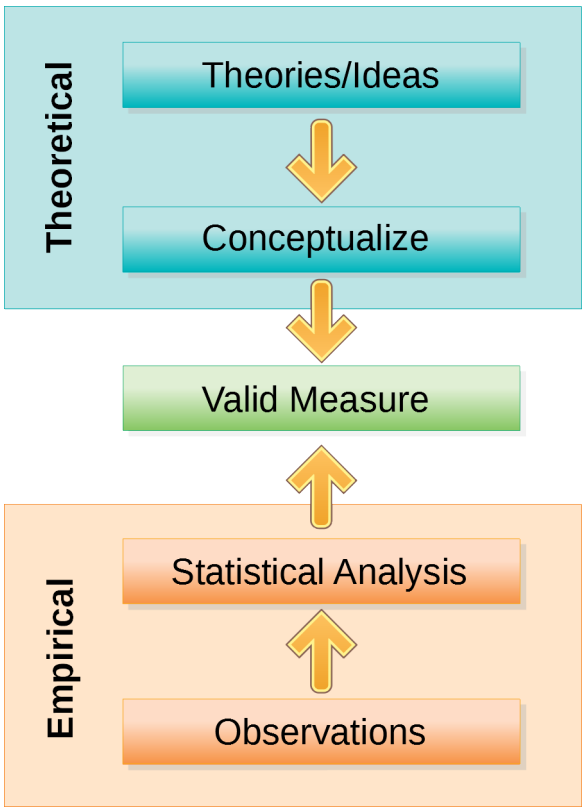


Figure 4: Assessing Theoretical and Empirical Validity

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 a lot of explanation. Hence this indicator has face validity. However, if we were to suggest how many books were checked out of an office library as a measure of employee morale, then such a measure would probably lack face validity because it does not seem to make much sense. Interestingly, some of the popular measures used in organizational research appear to lack face validity. For instance, 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 constructs that are highly abstract or are hard to conceptually separate from each other (e. g., compassion and empathy), it may be worthwhile to use a panel of experts to 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 in-

stance, to measure the construct "satisfaction with restaurant service," then the content domain should include the quality of food, courtesy of wait staff, duration of wait, and the overall ambiance of the restaurant (i.e., whether it is noisy, smoky, etc.). The project should measure the extent to which a restaurant patron is satisfied with the quality of food, courtesy of wait staff, the length of wait, and the restaurant's ambiance. 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 content validity of constructs.

CRITERION-RELATED VALIDITY

CONVERGENT VALIDITY refers to the closeness with which a measure, or group of measures, relates to (or converges on) the construct that it is purported to measure. Convergent validity can be established by comparing the observed values of one indicator with those of other indicators to attempt to find high correlation between those indicators. Compare this to discriminant validity.

DISCRIMINANT VALIDITY refers to the degree to which 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 then a measure of organizational knowledge must actually measure organizational knowledge (convergent validity) and 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 that is presumed to occur simultaneously. 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.

2.4.2.2 *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 hypothesized independent variable and not by variables extraneous to the research context. This is sometimes called "causality" and it requires three conditions:

1. covariation of cause and effect (if cause happens then effect also happens and if cause does not happen effect does not happen)
2. temporal precedence (cause must precede effect in time)
3. lack of plausible alternative explanation (or spurious correlation).

Certain research designs, such as laboratory experiments, are strong in internal validity since researchers can manipulate the independent variable (cause) via a treatment and observe the effect (dependent variable) of that treatment after a certain point in time while controlling for the effects of extraneous variables. Other designs, such as field surveys, are poor in internal validity because researchers cannot manipulate the independent variable (cause) and because cause and effect are measured at the same point in time which defeats temporal precedence making it equally likely that the effect actually brought about the presumed cause rather than the reverse.

EXTERNAL VALIDITY refers to whether the observed associations can be generalized from the sample to the population (population validity), or to 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 the population of all financial firms it would have strong population validity and if to other types of firms it would have strong ecological validity. Survey research, where data are sourced from a wide variety of individuals, firms, or other units of analysis, tends to have broader generalizability than laboratory experiments where artificially contrived treatments and strong control over extraneous variables render the findings less generalizable to real-life settings where treatments and extraneous variables cannot be controlled.

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.

CONSTRUCT VALIDITY examines how well a given measurement scale is measuring the theoretical construct that it is designed to measure. One frequent problem with construct validity is simply

defining the construct in such a way that it is measurable. As one example, "property ownership" is a construct of a market economy explained by Robert Reich[24]. That is, the fact that people can own property drives a local economy. But this construct relies on a number of external forces that cannot be controlled, such as local politics (a city's eminent domain can take a person's property) and the value of the property on the open market. Measuring the influence of property ownership on a local economy (the construct) would be very difficult since there are so many confounding variables.

STATISTICAL CONCLUSION VALIDITY examines the extent to which conclusions derived from a statistical procedure are valid. For example, it examines whether the right statistical method was used and whether the variables meet the assumptions of that statistical test (such as sample size or distributional requirements).

2.4.3 *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, 2) elimination, 3) inclusion, and 4) randomization.

1. Manipulation involves the researcher manipulating the independent variables in one or more ways (called "treatments"), and compares the effects of the treatments against a control group where subjects do not receive the treatment. Treatments may include a new drug or different dosage of 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, such as by restricting the study to a single gender or a single socioeconomic status.

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 also improves external validity, allowing inferences drawn from the sample to be generalized to the population from which the sample is drawn; however, generalizability across populations is harder to ascertain since populations may differ on multiple dimensions and only a few of those dimensions can be controlled.

2.5 SUMMARY

- In social science, our variables can be one of four different levels of measurement: nominal, ordinal, interval, or ratio.
- Indexes and typologies allow us to account for and simplify some of the complexities in our measures.

Part III

QUALITATIVE METHODS

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

Part IV

MIXED METHODS

All quantitative and qualitative research methods have certain strengths and weaknesses. Mixed methods are an attempt to use more than one research method on a given project to utilize the strengths of each method while mitigating their weaknesses.

Part V

REPORTING

After a research project is completed, the investigator must report the results of the project, often in both written and oral forms. This chapter concerns the reporting process.

Part VI

APPENDIX

3

APPENDIX

No appendix content yet.

GLOSSARY

descriptive research	Research that is designed to describe observed phenomena. The goal is to improve understanding rather than explore new ideas. *see* exploratory research
explanatory research	Research that is designed to explain an observed phenomena or process. *see* exploratory research
IRB	Institutional Review Board. 3, 10, 11, 13
paradigm	A pattern or model of how things work in the world. *see* theory
qualitative research	Qualitative research typically intends to explore observed phenomena with a goal of developing hypotheses and dive deep into a problem. Qualitative data collection involves semi-structured activities like focus groups and ethnographies. *see* quantitative research
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
theory	A system of ideas that is intended to explain phenomena. Theories that are accepted by scientists 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

BIBLIOGRAPHY

[1] George A Akerlof. "The market for "lemons": Quality uncertainty and the market mechanism." In: *Uncertainty in Economics*. Elsevier, 1978, pp. 235–251.

[2] Anol Bhattacherjee. "Social science research: Principles, methods, and practices." In: (2012).

[3] Amy Blackstone. *Principles of sociological inquiry–qualitative and quantitative methods*. The Saylor Foundation, 2012.

[4] Kathleen M Blee. "Racial violence in the United States." In: *Ethnic and Racial Studies* 28.4 (2005), pp. 599–619.

[5] John Brewer and Albert Hunter. *Multimethod research: a synthesis of styles.* Sage Publications, Inc, 1989.

[6] Gary Bridge and Robyn Dowling. "Microgeographies of retailing and gentrification." In: *Australian Geographer* 32.1 (2001), pp. 93–107.

[7] Alan D Bright. "The role of social marketing in leisure and recreation management." In: *Journal of leisure research* 32.1 (2000), pp. 12–17.

[8] Daniel R Denison. *Corporate culture and organizational effectiveness.* John Wiley & Sons, 1990.

[9] Ellen A Drost et al. "Validity and reliability in social science research." In: *Education Research and perspectives* 38.1 (2011), p. 105.

[10] Kristin G Esterberg. "Qualitative methods in social research." In: (2002).

[11] Moose George. *The image of man*. 1996.

[12] Diana M Hechavarria and Paul D Reynolds. "Cultural norms & business start-ups: the impact of national values on opportunity and necessity entrepreneurs." In: *International Entrepreneurship and Management Journal* 5.4 (2009), p. 417.

[13] Peter W Hom and Rodger W Griffeth. *Employee turnover*. South-Western Pub, 1995.

[14] John F Huck and Thaddeus McEwen. "Competencies needed for small business success: perceptions." In: *Journal of Small Business Management* 29.4 (1991), p. 90.

[15] Jonathan E Hughes, Christopher R Knittel, and Daniel Sperling. *Evidence of a shift in the short-run price elasticity of gasoline demand*. Tech. rep. National Bureau of Economic Research, 2006.

[16] Abraham Kaplan. *The conduct of inquiry: Methodology for behavioural science*. Routledge, 2017.

[17] David D Kirkpatrick. "Conservatives pick soft target: A cartoon sponge." In: *The New York Times* 20 (2005).

[18] Philip Kotler and Eduardo L Roberto. "Social marketing. Strategies for changing public behavior." In: (1989).

[19] Sarosh Kuruvilla and Aruna Ranganathan. "Economic development strategies and macro-and micro-level human resource policies: The case of India's "outsourcing" industry." In: *ILR Review* 62.1 (2008), pp. 39–72.

[20] C Jay Lambe, C Michael Wittmann, and Robert E Spekman. "Social exchange theory and research on business-to-business relational exchange." In: *Journal of Business-to-Business Marketing* 8.3 (2001), pp. 1–36.

[21] David Marshall, Stephanie O'Donohoe, and Stephen Kline. "Families, food, and pester power: beyond the blame game?" In: *Journal of Consumer Behaviour: An International Research Review* 6.4 (2007), pp. 164–181.

[22] Joseph G Ponterotto. "Qualitative research in counseling psychology: A primer on research paradigms and philosophy of science." In: *Journal of counseling psychology* 52.2 (2005), p. 126.

[23] Paul Christopher Price, I-Chant A Chiang, and Rajiv Jhangiani. *Research methods in psychology*. BCcampus, BC Open Textbook Project, 2015.

[24] Robert B Reich. *Saving capitalism: For the many, not the few*. Vintage, 2016.

[25] Randall S Schuler and Nikolai Rogovsky. "Understanding compensation practice variations across firms: The impact of national culture." In: *Journal of International Business Studies* 29.1 (1998), pp. 159–177.

[26] Pnina Shinebourne. "Using Q method in qualitative research." In: *International Journal of Qualitative Methods* 8.1 (2009), pp. 93–97.

[27] Daniel Monroe Sullivan. "From food desert to food mirage: Race, social class, and food shopping in a gentrifying neighborhood." In: *Advances in Applied Sociology* (2014).

[28] Charles Teddlie and Abbas Tashakkori. "A general typology of research designs featuring mixed methods." In: *Research in the Schools* 13.1 (2006), pp. 12–28.

[29] Attahir Yusuf. "Critical success factors for small business: Perceptions of South Pacific entrepreneurs." In: *Journal of small business management* 33.2 (1995), p. 68.

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured here:

http://postcards.miede.de/