

1

Brief Introduction to Research in the Social, Behavioral, and Health Sciences

What Is the Purpose of Research?

Research is the process that informs us about scientific knowledge. How do we know things? Research is the mechanism that allows us to acquire and refine knowledge. It is undertaken in disciplines in the natural sciences, social sciences, health sciences, and beyond. Existing research is then used as a starting place for future research. In other words, we proceed from what we know from existing research and then use our own research to extend that further. Another way to think of it is that scientific research is like a conversation between scientists who contribute their findings to the literature and then continue communicating by conducting follow-up research that may bolster or contradict the original research, or may take the conversation (research literature) in a different direction altogether.

How Is Research Done?

Research is carried out using systematic and verifiable methods. Systematic methods are important so that bias can be minimized in research. Without implementing systematic methods, if you get two different results in two different experiments, it may well be that the methodology is the cause of

the difference if the process wasn't carried out systematically. Moreover, it is extremely important to document your methodology as well as any difficulties, changes, and so on that occurred while you were conducting your research. This verifiable evidence allows you to demonstrate the efficacy of your findings, but beyond that, it will allow you or other researchers to better use data or findings from your study in the future as a starting block for subsequent research. Without those important details, it may be more difficult to understand the process that was used, which informs the process for future study.

Scientific Method and Hypothesis Testing

Almost all students learn something about the scientific method in science classes as early as elementary school. The key things that we learned then hold true in more advanced work as well. The scientific method allows us to eliminate alternative hypotheses, leaving our research hypothesis with greater credibility (rather than "proving" a particular research hypothesis).

In scientific research, you might think it would be appropriate to say that you proved something, but that is never the appropriate way to describe what happens in scientific research. It is not like a court of law where lawyers seek to prove something. As described below, you will see that in addition to your own hypothesis, you will state the opposite hypothesis, a statement of no difference where your hypothesis is the theory that is "different." At that point, the scientific research sets out to either "reject" the statement of no difference (the competing hypothesis), or it "fails to reject" that statement of no difference. If the statement of no difference is rejected, then that leaves your hypothesis standing as the remaining explanation, but we would not use the word *proved* to describe the result. If the scientific research fails to reject that statement of no difference, then this is an indicator that the statement of no difference could be right. You have not proved that either, but this situation does not provide any support for your hypothesis.

The scientific method we use is rooted in positivist philosophy, holding the assumption that there is such a thing as objective reality. The idea is that while we may not be able to directly see this "reality," we can directly observe measures that provide information about objective reality. This is the idea that underpins research methodology procedures.

Specific hypothesis tests are covered later in this book (see Chapter 8). The format of a hypothesis test can look something like this:

H_1 : The research hypothesis

(the thing you think might be true)

H_0 : The null hypothesis

(the thing that is a statement of no difference)

This will be discussed later in this book, but for now, here's what the hypothesis test might look like, using symbols and explained below:

H_1 : The mean (of something) for Group 1 is greater than the mean for Group 2.

H_0 : The mean (of something) for Group 1 is the same as the mean for Group 2.

In symbols:

$H_1: \mu_1 > \mu_2$

$H_0: \mu_1 = \mu_2$

where

μ_1 is the population mean of Group 1

μ_2 is the population mean of Group 2

(μ is the Greek letter *mu*)

To put this into real-world terms, the mean might represent something like average salary; Groups 1 and 2 might represent men and women. In that case, the hypothesis test above would suggest that the research predicts that men's salaries will be higher, on average, than those of women. The null hypothesis (a statement of no difference) would indicate that the salaries of men and women, on average, are equal.

You can test this hypothesis right away, since the data to do so are readily available. The **General Social Survey** (GSS) contains the variables needed to address this research question.

Of course, this format works for any variables that you choose. Presumably, if you are proposing a hypothesis, you should have an idea of how you will obtain observations (data) so that you can test that hypothesis. Otherwise, the research can go little further than speculation.

Inductive Research

This approach departs from what we think of as the traditional scientific method. With **inductive research**, we begin with the observations (data that have already been collected, or perhaps are collected first for the purpose of the research). Inductive research tends to be more descriptive and qualitative, though not exclusively so. Inductive research has a focus on studying observations for trends, patterns, or irregularities. With the discovery of such, theory and prior research are then consulted to help formulate your hypothesis.

BOX 1.1**THE GENERAL SOCIAL SURVEY**

The National Opinion Research Center (NORC) at the University of Chicago administers the GSS. The GSS was started in 1972 and continues today. The data used for the examples in this book come from the latest available completed version of the GSS collected in 2016. According to NORC, with the exception of the U.S. Census, the GSS is the most frequently analyzed source of information in the social sciences. NORC acknowledges that there are at least 14,000 instances where the GSS has been used for articles in scholarly journals, books, and doctoral dissertations. Furthermore, it is estimated that more than 400,000 students annually use the GSS in their work.

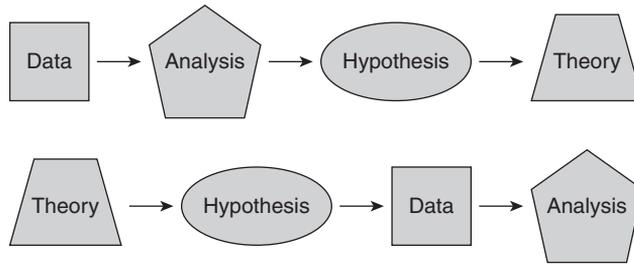
The GSS contains many demographic and attitudinal questions, as well as rotating topics of special interest. A number of core questions have remained unchanged in each survey since 1972. This allows for rich longitudinal research about attitudes, opinions, and demographics in the United States. Topical questions appear sometimes for just one year; other times, they can appear for a period of years. Therefore, the GSS is versatile as a longitudinal data resource and a relevant cross-sectional resource.

To maximize the amount of information that can be collected in this massive interviewing project, the GSS uses a **split ballot design**, in which NORC's interviewers asked some questions in a random subsample of the households and asked other questions in other households. Some questions, including demographic items, were asked of all respondents. When we begin analyzing the GSS data, you will notice that some data items have a substantial number of respondents for whom data are marked as missing. For the most part, this refers to respondents who were not asked that particular question as a result of the split ballot design.

Although many items were asked of only a subsample of respondents, you can still take the responses as representative of the U.S. adult noninstitutionalized population, subject to normal sampling error. For more information about how the GSS data were collected, see Appendix B, Field Work and Interviewer Specifications, and Appendix C, General Coding Instructions, in the *General Social Survey 1972–2016 Cumulative Codebook* (Smith, DalVERN, Freese, & Hout, 2017).

Deductive Research

This approach directly involves what we think of as the scientific method. With **deductive research**, we begin with a theory and a hypothesis. So, at this point you will propose your hypothesis and at least one competing hypothesis (usually the statement of no difference described previously). Then we collect

FIGURE 1.1 • Deductive and Inductive Research

data (observations) to conduct the hypothesis test. Deductive research lends itself well to quantitative research. (See Figure 1.1 for a comparison of the two types of research.)

Research Designs

How to collect data (observations), as well as when and where to collect data, are critical concerns in developing an appropriate research design. You will need to consider available access to the population you wish to study, time, and other resources available. If you are studying a hard-to-reach or hard-to-identify population, that will be a pervasive foundation of your research, where you will need to address the difficulties and specify how you will overcome them. In proposing any research, it is important to be realistic about resources that are needed to carry out the research. Time, money, assistance, equipment, and other resources should be calculated and planned before embarking on the project. Without the resources needed to carry out the research, the project will fall short of an effective completion. With regard to time, you will need to determine the time period for data collection: Will the data be collected at one point in time, or will they be collected at more than one point in time to assess changes over the interval(s) of time?

Cross-Sectional Research Design

With **cross-sectional research**, the data are collected at one point in time. The data (observations) should represent a cross-section, or representative slice, of the population you intend to study. To obtain this cross-section, or representative sample, a variety of methods can be used (see Chapter 3).

Extra detail: To be clear, it's usually not possible to collect all data at the same second, or even in the same day, so that interval during which you collect data (say, 2 months) is considered as one point in time.

Longitudinal Research Design

If you decide to collect data from respondents at more than one point in time for the express purpose of comparison of change over time, then you

would be conducting **longitudinal research**. There are a number of different ways to facilitate this, but they all share at least one thing in common: data collection will need to be scheduled for at least two points or intervals. If you are interested in multiple time points, such as collecting data annually for 10 years, then you would need to schedule and budget for the future data collection required in that design.

Repeated Cross-Sectional (Longitudinal) Research Design

This longitudinal design is exactly as it sounds: a repeated cross-section. In the **repeated cross-sectional research design**, a cross-section is taken at Time 1, then another independent cross-section of the same population is taken at Time 2. If the design calls for it, additional independent cross-sections can be taken at Time 3, Time 4, and so on, as many times as prescribed. Sampling methods used in each of the time periods should be the same or similar, unless you have documented evidence of improved sampling over time and you can demonstrate that the sample taken at each point in time is representative of the target population.

Fixed Sample Panel (Longitudinal) Research Design

The **fixed sample panel research design** starts just like a repeated cross-sectional design, with a cross-section at Time 1. However, any similarities with the former longitudinal design end there. At Time 2, the researcher must follow up with the same group selected for this original sample at Time 1. Also, it is important to point out that this is generally not something that can be effectively carried out spontaneously. In other words, it is not typical to begin a longitudinal research study not knowing whether you will use a repeated cross-sectional design or a fixed sample panel design. Part of the reason for this is that since the fixed sample panel design requires following up with the same respondents at a future point in time, the selection process is usually attuned to this so that you will have a better chance of having respondents follow up. Also, it is important to disclose to subjects that the nature of the study necessitates continuing data collection in the future and they will be contacted again. Moreover, there are strategies to increase retention of respondents over time, such as reminder e-mails, calls, postcards, thank-you notes, rewards for continued participation, and so on. It will be important to prepare a retention plan and budget up front at the very beginning of the research project.

This method has a tremendous advantage in that the change over time can be seen directly among the same respondents over time. The toll that must be paid for this advantage, however, can be high. Following up with the same respondents can be difficult, time-consuming, and costly. In some cases, the respondents may need to be located if they have changed addresses, stopped responding, gone away for vacation or other purposes, and so on. In other cases, respondents may not be able to be located at all. The two most obvious

situations like this would be (a) a respondent's illness or death or (b) willful refusal of a respondent to cooperate with further data collection. Some of the retention methods described previously can be useful to improve continued response rates, but it is important to understand and to estimate what the attrition rate will be and how that will be handled. Will the sample diminish in size over time (in which case it will have to be larger than normal at the outset), or will there be a method of resampling for those cases that are lost? The latter can be useful for overall comparative analysis, but still does not address the loss of continuous subject data over time. (See Figure 1.2 for a comparison of the two types of research design.)

Cohort (Longitudinal) Research Design

The **cohort (longitudinal) research design** resembles the repeated cross-sectional design, but adds a requirement for the members who are to

FIGURE 1.2 • Fixed Sample Panel Versus Repeated Cross-Section

Fixed Sample	Repeated Cross-Section
<p><i>Pros</i></p> <p>Follows same respondents Observe changes in individuals</p>	<p><i>Pros</i></p> <p>Less expensive Easier to sample</p>
<p><i>Cons</i></p> <p>Difficult to keep same respondents More expensive</p>	<p><i>Cons</i></p> <p>Different respondents over time</p>

BOX 1.2

MINI-CASE: FRAMINGHAM HEART STUDY

The Framingham Heart Study is a longitudinal study of the residents of Framingham, Massachusetts. The study began in 1948 and began a fixed sample panel design. Additional generations were later added to the study, but the design follows up with the same respondents over time to observe changes in outcome variables related to experiences of diet, exercise, and other factors. A great deal of knowledge about heart disease and the effects of cigarette smoking was first uncovered through this study. The research design enables the observation of particular behaviors to determine their role in health. For more information, go to the official website of the study: <https://www.framinghamheartstudy.org>.

be selected into the sample: They must have some common starting point. Typical cohorts include birth cohorts (born in the same year) and graduation cohorts (graduated from high school or college in the same year).

Terms

cohort (longitudinal) research
design 7
cross-sectional research 5
deductive research 4
fixed sample panel research
design 6

General Social Survey 3
inductive research 3
longitudinal research 6
repeated cross-sectional research
design 6
split ballot design 4

Reference

Smith, T. W., Dalvern, M., Freese, J., & Hout, M. (2017). *General social surveys, 1972–2016: Cumulative codebook*. Chicago, IL: National Opinion Research Center.