Sociologists doing Research

Learning Outcomes
At the end of this chapter you will be able to do the following.
- Explain the steps in the research process.
- Define and identify dependent and independent variables.
- Explain sampling.
- Calculate the mean, median, and mode of data.
- Identify levels of measurement of variables.
- Analyze ethical concerns in research.

One of the most remarkable traits that August Comte mandated for Sociology was a core of scientific rigor. He proposed the concept of **positivism** which is the *scientifically-based sociological research that uses scientific tools such as survey, sampling, objective measurement, and cultural and historical analysis to study and understand society*. Although the current definition of positivism expands far beyond Comte's original vision, sociological scientific methodology is used by government and industry researchers and across higher education and the private sector. Comte was originally interested in **social statistics**, *why societies remain the same,* and **social dynamics**, *why societies change*. Most sociological research today falls within these broad categories. Sociologists strive for **objectivity**, *which is the ability to study and observe without distortion or bias, especially personal bias*. Bias-free research is an ideal that, which if not present, could open the door to extreme misinterpretation of research findings.

Sociology is both different from and similar to other scientific principles. It differs from chemistry, biology, and physics in that sociology does not manipulate the physical environment using established natural science theories and principles. It is similar to chemistry, biology, and physics in that statistical principles guide the discovery and confirmation of data findings. Yet, sociology has no universally social laws that resemble gravity or the speed of light, as other scientific methods do. This is because chemistry, biology, and physics have the luxury of studying phenomena which are acted upon by laws of nature. Sociologists study people, groups, communities, and societies which are comprised of **agents**, *people who use their agency to make choices based on their varied motivations.*[1]

THE RESEARCH PROCESS[2]

*Problem Recognition & Definition*Researchers start with a question such as "What do I want to know?"; "What is important for society to know?"; or "Why does this occur?" Unfortunately some questions cannot be answered, such as "How many angels can dance on the head of a pin?" Even though many would like to know the answer to this question, it cannot be empirically observed; that is it cannot be perceived through one of the five senses—sight, taste, touch, hearing or smell. After a researcher decides on what question they want to answer they must state their goals and objectives. Do they want to determine if religious service attendance causes couples to have happier marriages? Or do they want

to describe the characteristics of happy marriages. The first one is a causal study (what causes what), and the second is a descriptive study.

The next step is to conduct a literature review to establish what is already known about the topic. Why reinvent the wheel? If social scientists have already done research on the characteristics of happy marriages, why do you need to do similar research as well? Well, perhaps the researcher before you only studied certain characteristics and you have thought of more that might be important. Much research in sociology builds on existing research.

The research question is usually stated as a hypothesis. A **hypothesis** is *the researcher's educated guess or prediction about what he/she will find*, such as "Those marriages that possess the most characteristics of happy marriages will be the happiest."

Creating the Research Design

There are many different types of studies that can be conducted. The most common type in sociology is survey research, but there are also interviews, observation, action research, polls, and experiments, as well as others.

One determinant of the research design is whether the researcher wants to describe some social phenomenon or determine if one phenomenon causes another phenomenon. **Descriptive studies** *answer the questions of who, what, where, and when*. **Causal studies** are undertaken to *determine how one variable affects another, and why*. Back to our marital happiness study, do we want to describe the characteristics of a happy marriage? Or do we want to determine if the presence of many of the characteristics causes a happier marriage? In other words, how does presence of such characteristics influence marital happiness?

Sampling

Sometimes the entire **population***, the group you are interested in researching,* can be studied; however, often it is too large to study everyone. Think of distributing a survey to all the students at College of the Canyons, that is over 20,000 people. Do we really need to survey all of them? Can we realistically survey all of them? If we carefully choose a **sample***, subset of the population*, it should reflect the characteristics of the population. Also, the way the sample answers the questions will be representative of everyone in the population.

Sampling methods are classified as either probability or nonprobability. In **probability samples***, each member of the population has a known chance of being selected*. Probability methods include random sampling, systematic sampling, and stratified sampling. In **nonprobability sampling** *members are selected from the population in some nonrandom manner*. These include convenience sampling, judgment sampling, quota sampling, and snowball sampling. You will learn about all of these sampling methods in your research methods class. We will discuss only random and convenience sampling in this chapter.

In **random sampling**, *each member of the population has an equal chance of being selected*. You need a list of everyone in your population to obtain a random sample. The easiest way to draw a random sample is to assign a number to each person in the population and then use a table of random numbers to select the subset or sample. (You will learn about this in your research methods class.)

Convenience sampling *is used when you do not have a list of everyone in your population so you choose participants because they are convenient to you*.

Data Collection and Analysis

The next step is to collect your data by administering your survey, interviewing your subjects, or making observations. If you collect **quantitative data**—*data that is, or can be converted to, numbers*—you can enter it into a computer program, typically SPSS (you will use this in your statistics class). If you collect **qualitative data**—*data that can not be converted to numbers; data that is about the quality of something*-you look for themes in the results. This type of data is usually collected by conducting interviews or making observations.

Reporting the Results

The reason we do research is to expand the knowledge base; in order to do that we need to report our results. This is typically done by submitting articles to journals, and/or presenting them at conferences. Journal articles typically contain several sections: abstract, statement of the problem, literature review providing the theoretical framework, methods used, results, discussion of the results, and references.

The analysis is the process through which large and complicated collections of scientific data are organized so that comparisons can be made and conclusions drawn. For example, if you want to prove that marijuana use leads to heroine use, you have to prove that there are no other contributing factors such as peer pressure or emotional or mental dysfunctions. The study must show **validity**—*the study must actually test what you intended to test*. If you want to say one event is the cause of another, you will need to rule out other possibilities or explanations to show that your research is valid. The study must also demonstrate **reliability**—*the ability to repeat findings of a research study*. Our study passes the test of reliability when we have demonstrated that our research process can be replicated with similar results.

Sociologists Perform Survey Research

Since by far the most common form of research in sociology is survey research, we are going to discuss how and when it is most useful. Many sociologists may opt to study people who choose, decide, succeed, fail, harm or help others, harm or help themselves, and behave in rational and irrational ways. If you took an ounce of gasoline and dropped a burning match into it, the gas will burn. The gas has no choice just as the flame has no

choice. But, if someone placed a burning match on your arm, or the arm of your classmate, you or they might respond in any number of ways. Most would find the experience to be painful. Some might enjoy it, others might retaliate with violence, and yet others might feel an emotional bond to the one who burned them. Sociologists must focus on the subjective definitions and perceptions that people place on their choices and motivations. In general, **surveys** are *research instruments designed to obtain information from individuals who belong to a larger group, organization, or society*. The information gathered is used to describe, explain, and at times predict attitudes, behaviors, aspirations, and intended behaviors. Surveys are easily used to collect information about political views, social and religious opinions, demographic information, past or expected future behavior, and even marital happiness and characteristics such as communication style, level of commitment, and fidelity.

Polls are typically *surveys which collect opinions*, such as who one might vote for in an election, how one feels about the outcome of a controversial issue, or how one evaluates a public official or organization. Surveys can be administered **cross-sectionally**, *once*, or administered **longitudinally**, *two or more times*.

If you administer your survey and get a good **response rate**—*the percentage of people who complete your survey*—you can generalize your results to the entire population. **Generalizability** means that the *results from the sample can be assumed to apply to the population as though the population itself had been studied*.

Also important is the quality of the survey itself as a scientific instrument. **Valid survey questions** *are questions that are accurate and measure what they claim they will measure.* For example, if we wanted to know how students feel about the Lacrosse team at College of the Canyons. Which statement should we propose to find out their level of agreement? 1. "Every campus needs a Lacrosse team" or 2. "College of the Canyons would benefit from a Lacrosse team." The first asks about all campuses, not specifically this one. It is seeking an opinion about campuses and Lacrosse teams in general. The second asks specifically about this campus, and is a valid measure of what we want to know.

Reliable questions are *questions that are relatively free from bias errors which might taint the findings*. In other words, reliable survey questions are consistent. If I ask a similar group of people the same question I will get similar results.

Surveys

There are two types of survey questions. First, **open-ended questions** *are questions designed to get respondents to answer in their own words* (e.g., "What might be the benefits of having a Lacrosse team?"_____ .) Second, **closed-ended questions** *are questions designed to get respondents to choose from a list of responses you provide to them* (e.g., "Are you married?" Yes or No.) **Likert scale questions** *are statements which respondents are asked the degree to which they agree or disagree with a question or statement.* They are the most common types of questions used in surveys (e.g., "How much do you agree that the

president is doing a good job of running the country?" Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree.)

Demographic questions *are questions which provide the basic categorical information about respondents such as age, sex, race, educational level, marital status, and other applicable information..*

Levels of Measurement

Nominal level data *is data with no standard numerical values.* **This is often referred to as categorical data (**e.g., What is your favorite type of pet? __Reptile __Canine __Feline __Bird __Other). There is no numerical value associated with reptile that makes it more or less valuable than a canine or other type of pet. Other examples include sex, favorite color, or town you grew up in.

Ordinal level data *is data with categories that have an order to them.* One category is more of something than another category. For example, height measured as short, medium, and tall is ordinal because medium is more height than short and tall is more height than both short and medium. Likert scale responses are usually at the ordinal level.

Interval level data *is data with categories that have an order, but we add standard numerical values with regular intervals.* If we measure height in feet and inches we have interval data. A height of 5 feet, 3 inches is 8 inches away from 5 feet, 11 inches. Each of those 8 inches has the same value, the intervals are identical. Five feet, 3 inches is one of the categories, but in this case the categories are numbers. The Fahrenheit temperature scale is an example of an interval scale. The difference between 68 degrees and 72 degrees is the exact same four degrees as the difference between 101 degrees and 105 degrees.

Ratio level data *adds a real zero starting point for the numerical values.* We can create ratios with ratio level data. With ratio data we can say that someone who has two children has twice as many children as someone having only one child, and someone having four children has twice the children of someone who has just two children, and the person with four children has four times the number of children as the person with only one child. Also, it is feasible to have zero children, or a real zero pertaining to the numerical data of quantity of children. Unrelated to this particular example, it is important to note that when examining the variables of height and age, it is not possible to be completely void of either Ratio data is used to compare to other data. For example, the sex ratio is the number of males per 100 females in a society. In 2005, the sex ratio for Alaska, Rhode Island, and the US was Alaska 103.1; Rhode Island 93.1, and U.S. 96.[3] We can say that Alaska had more males than females (103.1 males per 100 females) while Rhode Island had more females than males (93.1 males per 100 females). The U.S. overall has more females than males (96 males per 100 females).

Number of males and females, opinions about a Lacrosse team, marital happiness, height, and sex are variables. **Variables** *vary by respondent* (one is male, the next is female, the next is female, etc.). Sex is the variable and male or female are the **attributes***, or the*

possible category choices. Everyone in your class is human, so humanness is not a variable—it does not vary. But almost everything else you can observe is a variable.

Two types of variables are dependent and independent variables. **Dependent variables** *change in response to the influence of independent variables; they depend upon the independent variables.* **Independent variables** *are variables that when manipulated will stimulate a change upon the dependent variables.* If I know the independent variable, then can I predict what the dependent variable will be? If I know that you possess many of the characteristics of happy marriages, then can I predict your level of happiness? Yes. That does not mean that everyone with many of the characteristics will be the happiest, but more often than not, they will be. So possession of characteristics is the independent variable and happiness is the dependent variable. How happy you are depends on how many of the characteristics you possess.

Is this a causal relationship or merely an association or correlation? A **causal relationship** *is when one variable actually causes the other to occur*, such as eating lots of Krispy Kreme donuts causes you to gain weight. That is pretty clear, but in sociology most relationships are not that clear. Do I know for certain that possession of many of the characteristics that are found in happy marriages causes a marriage to be happy? No. What if there is something else that is causing both happiness and possession of characteristics? Maybe it is religion, or optimistic personality, or something else. If this is true, then this is an association or correlation. They go hand in hand, but one does not cause the other. In social science research, it is important to know that correlation does not equal causation.

Quantitative Analysis

When basic statistics are performed on data, we call them measures of central tendency (mean, median, and mode). Consider this list of numbers which represents the number of movies that nine students have seen in the last two weeks: 0, 1, 1, 1, 3, 4, 4, 5, 8.

The **mean** *is the arithmetic score of all the numbers divided by the total number of students* (i.e., 27÷9=3). The **median** *is the exact mid-point value in the ordered list of scores* (e.g., 0, 1, 1, & 1 fall below and 4, 4, 5, & 8 fall above the number 3 thus 3 is the median). The **mode** *is the number which occurs most often* (e.g., 1 occurs the most, so the mode is 1). The **extreme values** or **outliers** are *the especially low or high numbers in the series* (e.g., 8). Notice that if you removed the 9th student's score and averaged only the remaining scores the mean would be 2.375. Extreme values can increase or decrease the mean. You will cover these basic and more interesting statistics in your statistics class.

ETHICS OF RESEARCH

Ethics *are standards of what is right and wrong.* They are a general agreement shared by researchers as to what is proper and improper in scientific research. Our culture and sociology have ethical standards that may be different from other disciplines or other cultures. Standards may arise from religious, political, or pragmatic sources. Standards

differ over time; for example, long ago we did not have formal considerations about how to treat people who participate in scientific studies.

There are four major **ethical issues that protect research subjects**. First, voluntary participation means that subjects must participate voluntarily, they must understand the risks of participating, and they must be able to withdraw from the study at any time. Second, researchers can do no harm to participants. This includes anything from killing someone to causing them undue stress. Third, every study must be confidential which means that the researcher can never divulge the participants' identities. Some studies are anonymous which means the researcher does not know the participants' identities. Fourth, deception cannot be used to get people to participate in research they would not want to participate in.

To be sure that subjects know what they are getting into when they agree to participate in a study they sign an informed consent form which tells them the general purpose of the study, explains their right to withdraw, explains the confidentiality of the study, tells whether it is anonymous, explains the potential risks, and describes how to contact the researcher.

A university or research institution will have an **Intuitional Review Board (IRB)**, which *oversees and makes sure all research meets ethical standards*. Often researchers will need to get written plans for their projects approved by the IRB before they begin official research.

Even with ethical standards, you can probably think of times when it would be necessary to deceive a subject or when you might need to cause just a little stress to investigate something. Of course there are exceptions, but we will leave that to your research methods class to clarify.

An Outline of the Process of Social Research

1. Develop a research question
2. Consult prior research (literature review)
3. Select a population and sample
4. Select a method of data collection
5. Select a method of data analysis (qualitative or quantitative)
6. Collect data
7. Analyse data
8. Report research findings

---

[1] Google Anthony Giddens-human agency, January 18, 1938 British Sociologist.

[2] http://www.statpac.com/research-papers/research-process.htm

[3] http://factfinder.census.gov/servlet/GRTTable?_bm=y&-_box_head_nbr=R0102&-ds_name=ACS_2005_EST_G00_&-_lang=en&-format=US-30) 7 January, 2012.