

COMMONWEALTH OF AUSTRALIA
Copyright Regulations 1969

WARNING

This material has been reproduced and communicated to you by or on behalf of Harvest Bible College pursuant to Part VB of the *Copyright Act 1968 (the Act)*.

The material in this communication may be subject to copyright under the Act. Any further reproduction or communication of this material by you may be the subject of copyright protection under the Act.
Do not remove this notice.

Principles of Data Management and Presentation

JOHN P. HOFFMANN

8



UNIVERSITY OF CALIFORNIA PRESS

To Curtis

University of California Press, one of the most distinguished university presses in the United States, enriches lives around the world by advancing scholarship in the humanities, social sciences, and natural sciences. Its activities are supported by the UC Press Foundation and by philanthropic contributions from individuals and institutions. For more information, visit www.ucpress.edu.

University of California Press
Oakland, California

© 2017 by The Regents of the University of California

Library of Congress Cataloging-in-Publication Data

Names: Hoffmann, John P. (John Patrick), 1962—author.

Title: Principles of data management and presentation / John P. Hoffmann.

Description: Oakland, California: University of California Press, [2017] |

Includes bibliographical references and index. | Description based on print version record and CIP data provided by publisher; resource not viewed.

Identifiers: LCCN 2017005895 (print) | LCCN 2017012555 (ebook) | ISBN 9780520964327 (epub and ePDF) |

ISBN 9780520289956 (cloth : alk. paper) | ISBN 9780520289949 (pbk. : alk. paper)

Subjects: LCSH: Research—Methodology. | Research—Data processing—Management.

Classification: LCC Q180.55 (ebook) | LCC Q180.55 .H64 2017 (print) | DDC 001.4/2—dc23

LC record available at <https://lcn.loc.gov/2017005895>

Stata® is a registered trademark of StataCorp LP, 4905 Lakeway Drive, College Station, TX 77845 USA. SAS® and all other SAS Institute Inc. products or service names are registered trademarks in the USA and other countries.® indicates USA registration. SAS Institute, Inc., 100 SAS Campus Drive, Cary, NC 27513 USA. SPSS® is a registered trademark of SPSS, Inc., 233 S. Wacker Drive, 11th Floor, Chicago, IL 60606 USA. SPSS, Inc. is an IBM company. Their use herein is for informational and instructional purposes only.

Manufactured in United States of America

25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

Developing Research Questions

The United States has the highest imprisonment rate in the western world, with more than 2 million inmates in state and federal prisons and county jails. This represents about 1 out of every 110 adults in the United States. Each year, more than 600,000 of these inmates are released back into the community. One of the problems many of them face is employment: some companies are hesitant to hire former prisoners and thus it is difficult to find a decent paying job. Though this is not a new or surprising piece of information, the sociologist Devah Pager was curious whether race also played a role in finding a job, for previous studies had suggested that African-Americans were less likely to be hired than Caucasians even when their experience and skill levels were similar. Therefore, she set up an experiment to test the question of whether a criminal record or race had a larger impact on finding a job. Pager conducted an experiment in which she sent four college students—two African-American and two Caucasian—to apply for a large number of job openings for entry level positions (no experience needed). The background of the students was virtually identical except for their race and their report of a prison record (which was randomly assigned—in other words, they had not actually been in prison, but reported that they had to a random set of employers). As anticipated, when they reported a prison record, they were less likely to be called back for an interview. There was a surprising finding, however: whereas 17% of the time the Caucasians with a “prison record” were called for an interview, only 14% of the time were the African-Americans with no “prison record” called for an interview. In other words, being an African-American seemed more detrimental to job prospects than having a criminal record (Pager 2008). Although one might have questions about this research, consider how asking a new question of an old research topic can lead to a surprising result.

Now let's see an example of what some might think of as a more dubious research question (you decide). In 2014, the Ig Nobel Prize in Psychology was given to a research team that used data mining of a huge dataset to find that depression and cat bites are associated, especially in women. In the abstract to their article, the researchers noted that 117,000 of 1.3 million patients in the database had depression (9%) and 750 patients had experienced cat bites (0.06%). Furthermore, "depression was found in 41.3% of patients with cat bites . . . [including 47% of women] compared to 24.2% of men presenting with a similar bite" (Hanauer et al. 2013). What do you think of their research question: "is being bitten by a cat associated with depression?" (Or is it the other way around: are cats more likely to bite depressed people?) By the way, Ig Nobel Prizes are given each year to "achievements that make people laugh and then think" (www.improbable.com/ig).

Recall that figure 1.1 in chapter 1 diagrammed the typical steps in the research process. The top part of the diagram involved selecting a topic and developing a research question. This is the part of the process discussed in this chapter. But what is a research question? As discussed in the last chapter, research is a set of transparent procedures that is designed to generate knowledge, such as answering a question, checking the veracity of a claim, or solving a problem. Thus, a research question is an inquiry designed to elicit information about a claim or a problem. It assumes that there are research methods available that may be used to gather this information. We'll discuss some important characteristics of research questions later. First, though, it's important to consider a broader issue: how are the topics that motivate particular research questions chosen by researchers?

SELECTING A TOPIC

How do researchers determine the specific topics they will study? This is the first step in determining which questions to ask and how to find answers to these questions. After all, if there is no way to answer a research question, then there is little point in asking it unless you simply like to imagine what the answers might be. For example, I might be interested in determining how many people died of the Black Death in Bath, England, during the year 1351, but it is unlikely I have tools to answer this question. So let's begin by considering how to find a topic for a research project. Where should one begin? The range of topics might be limited by disciplinary boundaries, so that's a good place to start. In the field I began in during my younger days—criminology—I was expected to conduct research on juvenile delinquency or criminal behavior. It would not have made much sense to study the effects of the Black Death in Bath, England, during the fourteenth century unless I could address its influence on criminal behavior or some closely related phenomena. Of course, if you have invested time in learning the range of topics examined by your discipline—perhaps by taking classes and reading books and articles in your field of study—then you've already narrowed the topics down to a certain degree. However, the social and behavioral sciences offer a substantial range of topics, from laboratory studies of rodents (psychology) to voting behaviors (political science) to national-level studies of human birth and death rates (demography/

sociology). Thus, let's consider how we might limit the scope of topics within an academic discipline (see Booth et al. 2008, for additional suggestions).

Try to Find a Topic That Interests You in Some Way

If you are a student taking a class that requires a research paper, this is not always possible (for instance, I found the study of corrections—which was a required part of the curriculum—terribly uninteresting when I was an undergraduate student). But let's assume you are taking a course that is at least somewhat interesting or that the course requirements are general enough to accommodate your interests. Or, are there particular courses that you've completed that were particularly fascinating? What did you find interesting about them: a specific subject, a theory, a required or recommended book or article? Sit down for a few minutes and think about the most fascinating topics you've come across in the field. Use your imagination, curiosity, and creative instincts. Consider what you've seen in the news media about this topic or issues that are meaningful to you (Foss and Waters 2007). What issues might be part of this topic? Take some notes or draw some pictures of the topic or issues that interest you. Sketch some relationships that seem intriguing:

Length of queue → *number of spontaneous social interactions per minute*

Personal income → *probability of voting*

Friendship networks → *happiness*

Perhaps the best, and most common, way to find topics is to read books and articles in the discipline (Alvesson and Sandberg 2013). The Internet makes it particularly easy to search for relevant books, articles, and reports; however, some care is needed since there is information overload for many topics. And it's not a good idea to spend too much time reading in depth before you have narrowed your topic to something manageable (White 2009). However, assume you are intrigued by studies of social deviance. Read a couple of chapters that look interesting from books with the term *social deviance* in the title (e.g., Rosoff and Pontell 2010; Weitzer 2002), or from topical journals such as *Deviant Behavior* and *Social Problems*. Pay particular attention to the conclusion section of journal articles; authors often describe what needs to be done now that they have completed their research. We'll return later to how journal articles are a good source of ideas. There might also be topics that seem well researched, but some of the key studies have never been replicated. A skeptical eye could point you toward potential flaws in the research that can be examined through replication. Or, you may think that replication is needed in order to ensure that a research finding that appears to have general support is valid. For example, Devah Pager's study discussed at the beginning of the chapter was conducted in Milwaukee, Wisconsin. Would the same results be found in other parts of the United States? Talk to colleagues, including instructors, about topics that are interesting. They can often point toward the most important material to read about a specific topic.

Consider a Problem That You Think Needs to Be Solved

It is helpful to think of research as a quest to solve a problem. Think about how common it is in the sciences to solve problems: mathematicians tackle algebra problems, physicists look for ways to produce cheap energy, and medical researchers try to find vaccines for human immunodeficiency virus (HIV). In the social sciences, problem-solving often means identifying some phenomenon that needs to be understood better in order to find solutions (e.g., homelessness in a city, methamphetamine use in a rural community). The key is that there is not sufficient knowledge or there is, at best, a poor understanding of the problem. Many of those in the research community actively seek out problems to work on; in fact, most consider the topics they work on as full of problems since they wish to gain information about them.

Similar to many types of problems, a research problem involves (1) a condition and (2) a cost or a consequence that someone doesn't like (Booth et al. 2008). The general condition of a research problem is, of course, not knowing something. I don't know what the association between experiencing a parental divorce and adolescent drug use is; I want to know more about it, though. My knowledge is incomplete. In addition, if I know more about this particular problem, I'll likely learn a little more about why some young people begin to use illegal drugs. This will not only increase my understanding of and curiosity about adolescent drug use, but it might also contribute to the research community's general understanding of this type of adolescent behavior. Furthermore, I can consider practical issues by thinking about how my research might be useful for preventing drug use or at least preventing the negative consequences of drug use by young people.

Make Sure the Topic Is Specific Enough (but Not Too Specific)

Choosing a specific topic that is neither too broad nor too narrow will allow you to complete the research in a reasonable amount of time. It is also important because, in the social and behavioral sciences, we generally tackle narrow topics. For example, why young people use illegal drugs is a huge area of inquiry that hundreds of researchers contribute to each year. I recently entered the term "adolescent drug use" in Google Scholar and limited the search to the previous two years. I stopped counting the number of articles listed after the 30th page. Thus, if you attempt to address this topic, you will likely have a lot of scattered information that is difficult to present concisely or in a report (this doesn't mean you shouldn't take on broad topics, especially if you're passionate about them, but just realize that you will need to invest a significant number of years). A more focused topic could be something like "Parents' mental health problems and adolescent drug use" or "Variations in the prevalence of adolescent drug use across different types of neighborhoods." Although both of these topics are still too broad to be examined in a single research project, the subject area is much more narrow and likely to yield interesting research questions.

Is There an Audience?

Recall that in chapter 1 we discussed the importance of considering your audience (Booth et al. 2008). As you are attempting to narrow the topic, don't forget to ask whether there is an

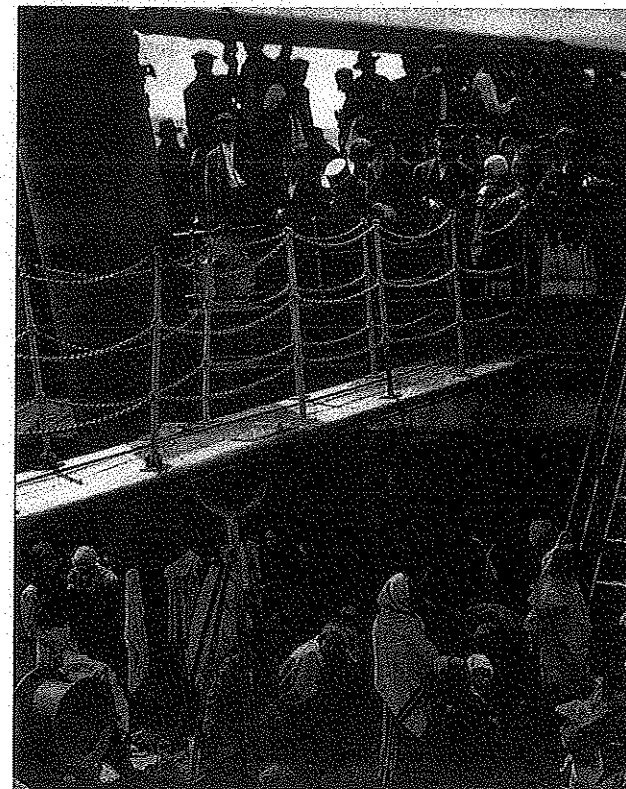


FIGURE 2.1 The steerage (photograph by Alfred Stieglitz).

audience interested in it. Suppose, for instance, that you are a psychologist who is intrigued by the way people use grammar when they write informally. The research in this area might be of interest to some of your fellow psychologists. However, you decide that you'd like to study whether religious people are more likely than others to confuse singular possessive cases and plural possessive cases when using surnames (e.g., James' dog is vicious vs. James's dog is vicious). Perhaps many religious people who read the Old Testament and derivative literature are used to seeing terms such as "Moses' writings" or "Moses' tablets." Before embarking on such an ambitious and time-consuming project, though, it's best to consider whether other psychologists are interested in this topic or whether they can be shown its importance.

Here's another example. Sociology has seen some growth in the area of visual studies: using photographs and other visual media as a research tool or as a way to understand how

people and groups comprehend visual images, such as vacation photos and films. But will sociologists care if I wish to explore the social implications of a single photograph taken by Alfred Stieglitz, a famous early twentieth-century photographer? One of his photographs I find intriguing is called "The Steerage," taken in 1907 (figure 2.1). I read a story that Stieglitz's wife, Emmeline, loved to travel by ship. The story mentioned that, during a trip on the Kaiser Wilhelm II, Stieglitz grew tired of being around the wealthy people in first class, so he took a walk, camera in hand, to see the decks below where the lowest paying passengers had to stay, an area known as *the steerage*. The photograph he took during his walk is, I think, full of sociological implications concerning social class, early twentieth-century immigration, and even a wealthy person's discomfort with traveling in style while others were packed into much smaller spaces. However, it is unlikely—although not impossible—that I would generate much interest among sociologists if I wrote an assessment of only this photograph. Rather, a topic that would likely garner more interest is to consider how social class was represented by photographers in the early twentieth century, so a comparison of several photographs or photographers, perhaps accompanied by an analysis of commentary by photographers and art critics of the period, might make a good research project.

FROM TOPIC TO RESEARCH QUESTION

Once the topic is sufficiently narrow—although this issue should continue to be on your mind—it is next important to think about how to use it to derive a research question. First, though, it is essential to recognize that research questions must be answerable (Andrews 2003). This means that they should be limited to those inquiries that they may be answered within the confines of a research project and using the methods available. A question such as "why does religion exist?" is far too broad an inquiry for present purposes, although it might be appropriate for one's life work (if done properly). A study of the dietary practices of Matsigenka people of Brazil might seem like a good idea, but if there is no way to get information on them, then there is little to gain by spending time developing a pertinent research question.

Moreover, research questions should be relevant, or you should be able to come up with a persuasive argument why they are relevant, to your discipline (Alvesson and Sandberg 2013). Some of the things that might seem interesting, though important, may not be relevant to your fellow sociologists, social workers, psychologists, or political scientists. With these qualifications in mind, what are some steps in deriving a research question?

Once You Have Identified a Topic, Ask Questions That Seem Interesting and Applicable

This will further help narrow the topic down to a manageable level. At this point, ask various questions about the topic; this may lead to some specific issues that are especially intriguing. Thinking about the topic, what are you curious about? Suppose, for example, that a researcher is interested in the topic of adolescent depression. She might then ask about

differences among those who are depressed and those who are not. Is it just a genetic abnormality that leads to depression? Then, what effect might parents' mental health status or their parenting practices have on their children's risk of depression? Do early childhood experiences—such as neglect, punishment experiences, or being withdrawn or shy—affect this risk? Do the symptoms change over time? Does behavior affect the risk? What about relationships with siblings and friends?

Consider the previous example of early twentieth-century photographers. What was the background of Stieglitz and others who photographed people from different classes? What was the context of the photographs? Did the photographers intend them to be art, photojournalism, both, or neither? Did their photographs tend to involve urban or rural areas, or different ethnic or immigrant groups? Were these photographs commissioned by some organization or did the photographers work on their own? What were the reactions of art critics or others when the photographs were exhibited? How did these photographs compare to other media of the period?

Another set of questions directly involves imagining the audience: What is the value of these photographs for understanding social class in early twentieth-century America? Will an analysis of them provide any persuasive evidence to a social scientist interested in the general issue of social class or immigration? Can the researchers convince an audience that the questions have wider significance? Perhaps a skeptical critic would argue that photography is a marginal field for understanding social class or class differences since most photographers in the early twentieth century came from the middle or upper classes and they were opportunistic when taking photographs. Furthermore, they were idiosyncratic when it came to creating artistic representations, so there is little useful information to be obtained from their work for a social scientific or historical understanding of class or immigration. If this is a legitimate argument, then perhaps the researcher needs to find another angle or, unfortunately, another research topic.

Figuring out which questions to ask can be difficult. Fortunately, there are some useful guides. For example, psychologist and creativity expert Keith Sawyer (2013, 34–36) provides a set of questions derived from a CIA checklist that is useful for arriving at good research ideas. These include the following:

- Why does this problem need to be solved? What can we gain by solving it?
- What don't I understand yet?
- What information is available? Is it sufficient? Are there sources of information available?
- What is the scope of the problem? What is *not* the problem?
- What are some parts of the issue or problem (break it down into components)?
- What cannot be changed regarding this problem? What factors must remain constant?
- Can you think of a similar problem or a different angle? Identifying analogous problems can be helpful.

You may be able to come up with other questions (see Foss and Waters 2007, Sak 2011). The point of this exercise is to narrow the topic down in such a way that a research question can emerge.

Identify a Gap

Earlier, we discussed a key way to find topics: reading books and articles in your discipline. As you narrow down the issues that interest you and read material about these issues, you will find that there are often well-settled questions that should normally be avoided (Campbell et al. 1982; unless replication is needed), but there are also almost always gaps in the literature (White 2009). Gaps exist for various reasons. Most often, they are present because a large majority of studies can offer only limited information about any research topic. Since research questions are, by their nature, rather narrow, there will almost always be more to know about a subject area.

It is not always easy to identify gaps, however. One efficient technique is to pay particular attention to the discussion and limitations sections of journal articles. These sections often provide a good way to find out what needs to be done regarding a particular topic. For example, suppose a researcher is interested in the association between stressful life conditions and illegal drug use. She reads an article in a recently published academic journal and notes the following sentence in the discussion section: "The next step is to determine whether adolescent males and females are affected differently by stress and whether this has any implications for initiating drug use." This suggests—although it is important to locate and read other articles to make sure—that the study of stress and illegal drug use has not evaluated differences between males and females. Perhaps this is worth considering.

Here is an example from an actual study. In the discussion section of a journal article on voting behavior and climate change skepticism, the authors wrote:

Voting for a politically conservative party leads to higher levels of climate change skepticism, whereas voting for more liberal parties leads to lower levels of skepticism. . . . [But] in countries where there is more bipartisan political support for responses to climate change, we might expect to find lower and more stable levels of climate change skepticism. (McCrea et al. 2016, 1327, 1329).

In other words, the researchers studied this issue with data from one nation, yet if we examined this issue across nations, we might find that the answers are more nuanced than their results showed. This suggests a research project that uses data from different countries and asks whether the presumed connection between voting for a particular political party and climate change skepticism depends on how closely the parties agree about how governments should respond to climate change.

I've noticed that something that many students are worried about is disagreeing with or criticizing experts in their field about the choice or utility of a research question. ("He's a PhD level researcher so he must know so much more than me about research. Who am I to

criticize?") However, this should not cause any anxiety because considering potential gaps, omissions, or even flaws in research is a common way that new ideas are generated or new research gets done. As mentioned earlier, most research questions are by design narrow and researchers can only tackle a thin slice of any topic in a single study. Thus, there is nothing wrong with thinking that there is another way to approach a research problem.

In addition, there may be healthy disagreement about the results of research projects because, from one's perspective, there are (1) flaws in the arguments, (2) flaws in the research methods and analysis, or (3) flaws in the interpretations. I once read an article in a prestigious criminology journal that claimed that school-level influences (e.g., student-teacher ratio; the proportion of poor children in a school) had no association with whether students misbehaved in school. However, individual characteristics, such as gender and school performance, did predict poor behavior. I looked at the way they did their study and realized that their data came from only seven schools. Knowing a bit about statistics, I suspected that this was not a large enough sample of schools to determine if there were school-level influences. I then used a dataset that had students from 1,000 schools and showed that when more schools were considered, school-level influences emerged as important. For example, schools where students, in general, were more involved in activities such as clubs and sports tended to have less misbehavior than in those schools with less student involvement.

As you read articles and look for gaps, consider some ways that research may be underdeveloped or even flawed. Moreover, think about alternative ways to tackle a particular research problem. Some issues to consider include the following:

- The researchers thought they were measuring some phenomenon, but you think there are other, perhaps even better ways, of measuring it. For instance, what did the researchers mean by "climate change skepticism" and how did this influence the way it was measured? Are there alternative questions that might be used to measure this form of skepticism?
- The researchers used a particular method, such as a quantitative approach, to examine a research question. It might also be useful to try a different method, such as a qualitative approach. Perhaps a study relied on survey data to examine the association between stress and alcohol use among young people, but you think qualitative interviewing might yield additional insight.
- There are alternative statistical methods one might use to study a problem. For example, a researcher may have used a linear technique when there was—in all likelihood—a nonlinear association. Suppose you come across a study of stressful life events and drug use that used a statistical model showing a linear relationship. You may think that these events only affect drug use up to a certain level, but the probability of use then decreases because highly stressed kids become depressed or withdrawn. Thus, understanding this issue might be advanced by studying nonlinear associations.

- From where were the data collected? Suppose that the sample used to examine a research question consisted of middle-class people from a Midwestern US city. Perhaps a sample from a Southern or Western US city would result in a different conclusion. Here, it is important to think about potential differences between people who live in different geographic areas. Similar questions about the source of the data include those about gender, race/ethnicity, socioeconomic status, nationality, and so forth.
- The researchers claimed that X influences Y, but actually you think that Y influences X. For example, the researchers examined whether voting for a particular political party affected subsequent views of climate change, but perhaps those views were already in place and then affected which party was more attractive to the voter.
- They claimed that X influences Y, but you think that some other factor affects both X and Y. For example, stress may not affect drug use; rather, poor parenting causes kids to experience stress and leads to drug use. In statistical terms, we say that poor parenting *confounds* the presumed association between stress and drug use.
- X influences Y only under certain circumstances. Suppose that some researchers studying welfare policies examined one nation that has a parliamentary democratic form of government. You might think it interesting to study these policies in nations with different forms of government (e.g., constitutional monarchy).

Use Theory

There are many definitions of the term *theory* and many ways that theories are used across academic disciplines. For example, Jonathan Turner (2013, 843) defines theory as “a mental activity revolving around the process of developing ideas that explain how and why events occur.” As suggested by this definition, it is common to view theories as explanations of some phenomena that may be tested in some way (although this often differs across disciplines also). Inquiring into *how* and *why* things occur also provides the basis for many theories. For example, why did the US Civil War occur? How did the earth’s surface temperature rise during the last 50 years? If a theory is a good explanation for a phenomenon, it should be a clear and thorough account and it should offer a better explanation than some other theory (Hoffmann 2003). Research questions are often designed as a test of theory through a deductive process or as a way to build a theory via induction. Figure 2.2 provides an illustration of the role that theories and research questions play in an ideal view of the way science progresses. Although some ideas have been labeled as laws in the social sciences—the law of demand, the iron law of oligarchy—it is rare to actually confirm that these laws are invariable across different conditions. In any event, this section focuses on using a theory to generate a research question.

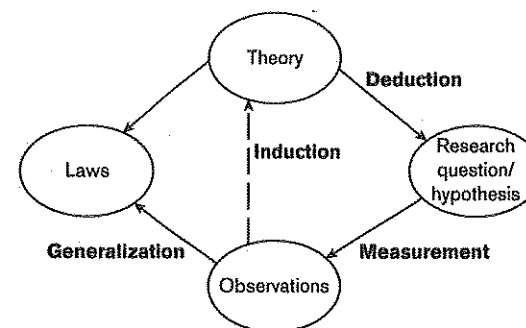


FIGURE 2.2 How science progresses.

A type of theory that is useful for generating research questions is identified by sociologist Robert K. Merton (1968) as “theories of the middle range.” These are less abstract than broad explanations of human and societal actions and behaviors, such as rational choice theory, but more general than specific hypotheses or research questions. Middle range theories tend to focus on a particular set of variables that may be used to explain some outcome. For instance, Robert Agnew’s (2006) general strain theory (GST) describes various types of subjective and objective stressors that are presumed to affect whether youth become involved in juvenile delinquency. Agnew also elaborates a set of intervening pathways that affect whether these stressors lead to illegal behaviors or some other outcome, such as depression. For example, if youth tend to respond to stressful situations with anger, they are likely to lash out violently and perhaps physically assault others when facing this type of condition. The theory may be used to develop specific research questions, such as one of those intimated earlier: do stressful life events, such as family fights or a death in the family, lead to a higher probability of initiation of marijuana use among adolescents?

Develop a Concept Map

A concept map is a way to depict the researcher’s thought process as she attempts to move from a topic to a research question (Alvarez and Gowin 2010). It shows explicitly a set of terms that begin generally and then narrows the focus down until she reaches a research area that is appropriate for a single study. Concept maps are similar to *idea maps* that are designed to help people find solutions to problems (Sawyer 2013). Both types are intended to show a path from a broader topic to a narrow issue. A key part of this strategy involves using *concepts*. Although concepts are addressed in more detail later in the chapter, they may be defined for now as “generalized statements about [entire] classes of phenomena rather than specific statements of fact” (Becker 2008, 109). These statements are designed to say something about people, organizations, or other social units in general, without addressing only those that apply at some particular time or place. Some common concepts used in the social

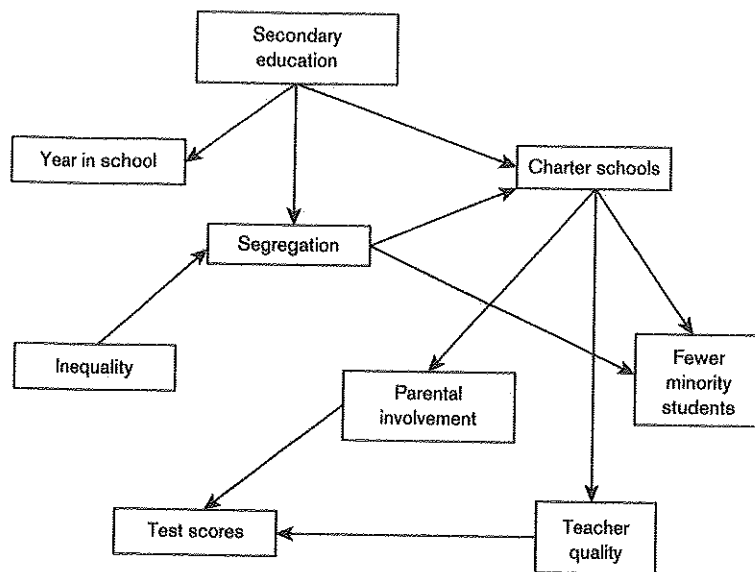


FIGURE 2.3 Concept map example. Research question: Do charter school students tend to have higher test scores because of greater parental involvement or because of higher quality teachers?

sciences are racial inequality, bureaucracy, juvenile delinquency, nationalism, prejudice, stereotype, and gender.

Although there are different ways to construct concept maps, the most useful is hierarchical in which the researcher

1. selects a topic and develops some concepts related to the topic;
2. ranks the concepts hierarchically from the most general to the most specific; and
3. arranges the concepts according to their hierarchical position and relationship to one another (Alvarez and Gown 2010, 5).

There are sometimes branches in the map that do not end up contributing to the research question, but are useful because they provide additional context for the topic. For instance, suppose a researcher is interested in the general topic of secondary education, such as what happens in high schools in the United States. What are some concepts related to the topic of education? From the literature in sociology and educational research, here are a few: test

scores, grades, year in school, school choice, charter schools, curriculum, teacher quality, parental involvement, segregation, school attachment and commitment, inequality, and extracurricular activities. In reading the literature, the researcher also notices that several descriptive studies indicate that students in charter schools have higher scores on standardized tests than students in noncharter schools. How might this information be used in a concept map? Assuming the researcher is familiar with the literature, has some interest in particular aspects of education, and has carefully thought about the process that affects student performance, figure 2.3 provides an example.

This is a simple and unrefined example of a concept map, but it illustrates some key points. However, one area that likely needs improvement involves the enhancement of concepts, which can be ambiguous (what is meant by parental involvement? teacher quality?). So some revisions may be needed (O'Leary 2014). Nonetheless, carefully constructing and refining concept maps is a useful approach to developing research questions.

Identify a Hypothesis

Many social and behavioral scientists think that their studies should try to mimic, as close as possible, the natural and physical sciences. Recall from chapter 1 that a hypothesis is a statement about the presumed relationship between two or more variables. It is also sometimes described as an educated guess regarding an answer to a question. Most descriptions in the natural and physical sciences outline a specific set of five steps for the scientific method, one of which is to develop a hypothesis:

1. Observe a phenomenon.
2. Ask a question about the phenomenon.
3. *Develop a hypothesis.*
4. Conduct an experiment to test the hypothesis (which includes collecting and analyzing data).
5. Evaluate the evidence from the experiment to determine the veracity of the hypothesis (McClelland 2015).

The hypothesis must be testable and worded in such a way that it could be true or false (falsifiability). For some researchers, it is easier to think about research questions by identifying a specific hypothesis or set of hypotheses during the course of their study (Andrews 2003). This can be useful, especially if one thinks there is a clear or particularly important association between two concepts or variables. For example, a recent psychological study hypothesized that "job-based psychological ownership is positively related to job satisfaction" (Peng and Pearce 2015, 157). Job-based ownership means that the person thinks of her job and duties as her own rather than simply part of an organization. Hypotheses are normally derived directly from research questions, but specifying a hypothesis can also help refine a research question.

REFINING RESEARCH QUESTIONS

As mentioned earlier, research questions in the social sciences are often derived from theories. Since theories are fundamentally about explanation, it is important to recall that an explanation is a clear and thorough account of the issue that a researcher is studying. We wish to know why something occurred or how it happened. This is different than descriptive research questions that focus on when or where something occurred or who is involved in some phenomenon. Of course, research questions can apply to both descriptive and explanatory research.

Regardless of the type of research, questions may be modeled and refined by using concepts and propositions. As noted earlier, concepts are general statements about classes of a phenomenon rather than specific instances of a phenomenon. For example, a concept such as *stereotype* is designed to elicit an image of a general belief—often mistaken—about a group of people (“all hipsters wear skinny jeans”). Concepts are useful because they can provide a summary statement or term regarding a whole class of events or outcomes (Becker 1998). In a research project, they should be defined clearly: What does the concept mean? In some projects, the concepts used are understood well because they are so common in a particular research area. This lessens the need to define them. In addition, researchers must think about how to move from defining a concept to measuring it. Practically speaking, how will the data be used to identify a concept? This is a vital part of a process known as *operationalization* and is key to any research project. Concepts are measured using variables, which in quantitative social analyses are usually based on questions asked during surveys or interviews. We return to the issue of measurement and operationalization in chapters 4 and 5. For now, the focus is on the role of concepts in the formation and refinement of research questions.

A proposition is a statement that links two concepts together. The proposition “hipsters wear skinny jeans” connects the concept of hipster with the concept of skinny jeans. Similarly, the statement “job-based psychological ownership is positively related to job satisfaction” (Peng and Pearce 2015, 157) connects one concept, psychological ownership, with another, job satisfaction. Thus, a hypothesis is an example of a proposition.

Using Arguments

Chapter 1 includes a discussion of research as argumentation. It is helpful to revisit this section since thinking about arguments can help refine research questions. For example, consider that propositions are one type of claim (“crime rates are rising,” “self-esteem is positively related to grades in high school”). When asking a research question, it is common to have a claim in mind, usually in the form of a proposition or hypothesis. For instance, the concept map shown in figure 2.3 includes the following research question: Do charter school students tend to have higher test scores because of greater parental involvement or because of higher quality teachers? This suggests two claims: (1) students in charter schools have higher test scores, than those in other public schools, because of greater parental involvement

and (2) students in charter schools have higher test scores, than those in other public schools, because of higher-quality teachers. What arguments might support one of these two propositions? Are there reasons based on previous studies or a particular theoretical perspective? Note that the conjunction “because” is used in both propositions. However, it is useful, even if grammatically messy, to extend the claims by using “because” again: students in charter schools have higher test scores, than those in other public schools, *because* of higher-quality teachers *because* these teachers *do something better than other teachers* (Booth et al. 2008). This will help point toward (a) whether answering the research question is feasible, (b) what evidence is needed to evaluate it, and (c) whether it is relevant.

Another important aspect of using arguments is to be prepared to consider alternatives. Members of high school and college debate clubs know this principle well: they are taught that considering alternative explanations and arguments is essential to being a successful debater. Moreover, chapter 1 discussed cognition traps in which a person's views are rigid, such as when only one explanation is considered when evaluating some phenomenon. Given that the researcher is—or should be—in dialogue with the audience, it is essential that she be prepared to consider alternatives. For example, suppose that neither claim about charter schools is correct. An alternative explanation is that many charter schools are selective about which students they accept. They tend to reject applicants who have shown behavioral problems, low academic achievement, or don't have much motivation to do schoolwork. Therefore, charter schools self-select high achieving students who get high test scores regardless of what type of schools they attend or the quality of their teachers (Nichols-Barrer et al. 2016). Thinking about alternative explanations and attempting to consider them explicitly during the study will help improve the dialogical argument process that leads to supporting one proposition over the other and can, eventually, be used to suggest future directions for research on this issue.

Given the basic model of an argument (see chapter 1), some questions to ask that may help refine research questions and claims include the following:

- What are some alternative explanations to the claim or alternative answers – other than what I've presumed – to the question? (Direct a skeptical eye toward the claim and try to list some alternatives; see Klass [2012, chapter 3] for some reasons to be skeptical.)
- Has the question already been answered sufficiently that more research is not needed (e.g., we know that kids who use marijuana tend to associate with peers who use marijuana)?
- What evidence do I need to answer the question or determine if one claim is supported over another?
- Should this evidence be primarily qualitative, quantitative, experimental, historical, textual, based on a case study, or some combination of these? What type of data are needed? Should they come from surveys, experiments, administrative entities, or some other source?

It is best if you write down some notes about each of these issues. Keeping notes along the way will make it easier if you have to backtrack and change your specific question.

Think Like an Experiment

It is often helpful to think in terms of how might we answer the research question if we had extensive control over the groups we are interested in comparing. Suppose we could, for instance, randomly assign youth to charter schools or regular public schools. Given what we know from previous studies and from considering different possibilities, what might have happened if youth who attended regular public schools had attended charter schools? In other words, imagine the counterfactual situation. A variation of this is what the sociologist Howard Becker (2008, 20–21) calls the *null hypothesis trick*. This involves imagining if the phenomenon under consideration was the outcome of a purely random group of people. If interest is in test scores among charter school students, assume that charter schools are comprised of a completely random selection of students with various backgrounds. Would we expect to find the particular outcomes, such as higher test scores, if this was the case? Or would this be more likely if the students were not randomly selected? What would the selection process look like if it was not random? For example, good students and their parents—rather than students at random—may be more likely than bad or mediocre students to select charter schools.

Think Like a Historian (or a Detective)

Historians typically consider events that have already happened and then try to figure out why and how the event came about. Similarly, detectives and medical examiners usually don't begin their work until after a crime has been committed. Both types of researchers attempt to reconstruct events based on evidence that they acquire and then develop the most likely explanations given this evidence. Of course, they also need to consider the validity and strength of the evidence. Moreover, historians and detectives are interested in motives: what were the reasons that people or groups did the things they did that brought about the event (Shore 2008)? Once again, Howard Becker (2008, 39) has an analogous idea for conducting research: the *machine trick*. Imagine you are given a machine and wish to discover how it works. Perhaps the machine produces something, such as computer microchips (for fascinating examples, see the Science Channel show "How It's Made" [<http://www.sciencechannel.com/tv-shows/how-its-made/videos>]). What goes into its construction? How do the parts work together to make the machine operate? This is likened to *reverse engineering*: taking a machine apart in order to duplicate it. To think in this way, begin with the outcome that you wish to explain, such as higher test scores among students. Then, think backward, based on the background knowledge you've accumulated, to figure out what personal and social conditions would produce higher test scores. What evidence would be needed to understand these conditions better? Do the conditions include attendance at charter schools, certain types of teachers, or particular sorts of families? What else would such a machine

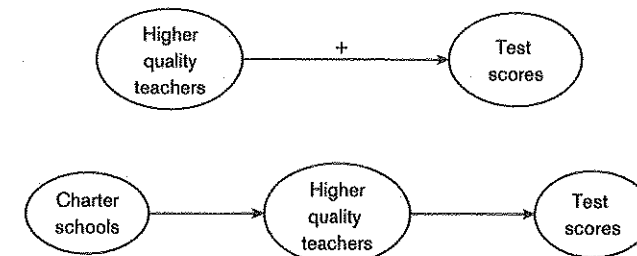


FIGURE 2.4 Example of two conceptual models.

need in order to operate? How does it produce higher test scores? What might be missing that would affect the success of the machine? This way of approaching a research question or issue is often useful for determining the quality and direction of the research.

Using Conceptual Models

In the social sciences, a conceptual model is a physical or diagrammatic representation of concepts and propositions. Some researchers prefer to call them "causal models," but this term implies more than most social science research projects can provide (see chapter 1), so it is avoided here. The diagrams are similar to flowcharts, yet the process shown relates to how one concept is thought to affect or be associated with another concept. They have also been used to depict theories, although some theories are too complex to be represented in this manner. Conceptual models are sometimes derived from concept maps, but they are more focused. For our purposes, we discuss just a couple of simple models by focusing on only two or three concepts linked by one or two propositions.

One of the values of conceptual models is that they strip research questions and claims down to their basic structure. Although there is a risk of oversimplification, it is a good idea to construct a conceptual model to determine if your research question and the claims you derive from it make sense (Whetten 2002). To construct one of these models in simple form, use oval shapes to represent concepts and arrows to represent propositions. Plus and minus signs are often placed on the arrows to indicate that the researcher thinks the association among the concepts is in a positive or a negative direction. For example, consider the earlier research question regarding test scores among students in charter schools. A couple of conceptual models are represented using the diagram in figure 2.4.

Although what these models represent is not entirely clear without additional context, they may be read as "the presence of higher quality teachers is positively related to higher test scores" and "charter schools tend to have higher quality teachers; thus, test scores tend to be higher." The second model assumes the first, which is much simpler, but the second model clarifies the claim better. Note that neither model explicitly shows the comparisons,

though. What are charter schools being compared with? What constitutes the concept outside the boundary of high-quality teachers (low-quality teachers or all other teachers?). What types of test scores are the research questions concerned with? Thus, to understand conceptual models, a narrative that explains the research question or claim and the concepts should accompany them. As with the previous section, it is also important to ask additional questions regarding these conceptual models before moving along in the research process:

- Are the concepts clear?
- Is there a way to measure them?
- Is evidence available with which to test whether the models are accurate?
- What type of evidence is needed?
- Are there alternative explanations that need to be considered?

AN EXAMPLE

We have touched upon a substantial number of ideas to guide the development of research questions. Let's now consider an example that uses a few of these ideas to narrow a research topic down to a research question. Suppose that we are given the task—perhaps by our employer or college teacher—to develop a research study on *recidivism*. The first thing to do is to figure out what this term means. According to the most common definitions, recidivism refers to a person's involvement in criminal behavior after he or she have been given sanctions or experienced some sort of formal state intervention for previous criminal activity. A literature review of research on recidivism further specifies this term as criminal behavior after serving a court-imposed sentence, such as time in jail or prison. Although there might be other ways to understand recidivism, this seems like a reasonable way to understand it.

Since this is such a broad topic, let's try to narrow it down a bit. Again, from the literature, it appears that most research on recidivism concerns those who have served time in state or federal prisons. Most of this research has also used large datasets collected by federal government agencies to track released prisoners and determine whether they are arrested for a crime from 1 to 3 years after release. Some questions we might ask at this stage include the following:

- Who tends to recidivate (e.g., males or females, violent offenders or property offenders)?
- When do they tend to recidivate (e.g., within the first three months)?
- Why do they tend to recidivate (e.g., because they go back to hanging out with the same friends; they cannot find a job)?
- Assuming the research is done in the United States, are those released in certain areas of the country more likely to recidivate?



FIGURE 2.5 Employment and recidivism conceptual model.

Imagine we decide that the most interesting issue involves the *why* question. It seems most important since if we understand why people recidivate, perhaps we can help inform programs to reduce it.

At this point, we review specific studies of recidivism and look for gaps, such as by carefully reading articles and paying close attention to hints at additional research that might be useful. One of the first things we notice is that most of the research over the past 10–15 years has been conducted on men rather than women. So it might be appropriate to limit our attention to women. There is also an article we find that provides an excellent summary of research on interventions that help reduce recidivism among women offenders (Gobeil et al. 2016). The authors of this article find that interventions such as drug treatment, employment training, and parenting programs while in prison led to a reduced likelihood of recidivism once female inmates were released. But, at this point, since the studies that were summarized in this article relied on general involvement in these intervention programs, we have little idea why or in what way they reduced recidivism. For example, did employment training improve their chances of getting a job once they were released?

At this point, let's consider a theory and how it might inform our attempt to hone in on a research question or hypothesis. In the field of criminology, one of the most popular explanations for criminal behavior is known as *social bond theory*. Briefly, this theory claims that people who are more attached and involved in conventional activities are less likely than others to commit criminal acts. For example, social bond theory has motivated studies that have found that when adults are gainfully employed in legitimate occupations, they are less likely to commit a crime, even if they are otherwise predisposed toward this sort of behavior (Uggen 1999). Thus, we reason that if employment training in prison does lead to employment after release, those women will be less likely to recidivate compared with women who did not receive this training. This leads to the conceptual model in figure 2.5.

It is now a good idea to consider the questions listed earlier: Are the concepts clear? Can they be measured? Is there evidence available to assess this conceptual model? Furthermore, what alternative explanations might also elucidate why employment programs lead to lower recidivism risk? For instance, perhaps especially violent prisoners cannot qualify for employment training and they are more likely than others to recidivate upon release. This example provides only one pathway to a potentially testable research question, but we could also have used other approaches, such as concept maps, to arrive at a research question.

SIDEBAR 2.1 RESEARCH TOPIC/QUESTION CHECKLIST

1. Is the topic right for me?
 - a. Is it interesting?
 - b. Is the topic overly broad? Is it too narrow?
 - c. Does it lend itself to a research question or set of questions?
 - d. Can I manage my biases?
2. Is the research question right for the field?
 - a. Has the question already been answered?
 - b. Does a relevant audience, such as experts in the field, think it is relevant and important?
 - c. Will the findings (potentially) be important?
 - d. Will they make a contribution?
3. Is the research question well articulated?
 - a. What theories, arguments, experimental thinking, and concept maps can be used to develop and refine the question?
 - b. Can I articulate a hypothesis or a conceptual model based on the question?
 - c. Are the concepts and propositions well defined?
 - d. Are there unarticulated assumptions?
 - e. What are some alternative explanations?

In addition to the issues outlined in this chapter, you should also ask yourself the following:

4. Can the research question be answered?
 - a. Are data available?
 - b. Are there datasets that already exist that might be used to answer the research question?
 - c. If data sources do exist, can I gain access to them?
 - d. Do I need to collect my own data? Is this feasible?
 - e. Can I measure the concepts adequately to answer the research question?
 - f. Do I have the right skills?
 - g. Can I complete the research in time?
 - h. Will it cost too much?
 - i. Are there ethical problems?
5. Does the question meet the approval of my supervisor, advisor, or committee?
 - a. Do they think I am on the right track?

FINAL WORDS

Research questions are an essential component of the research process. The proper identification of research questions ensures that the scope of the research is sufficiently narrow, valuable, interesting, answerable, and nonredundant. The research question is also central to the research process because it is related to each component of this process (White 2009). Research questions are typically developed by narrowing down a topic that the researcher finds interesting and important, and that remains at least partly unanswered. They often emerge through engagement with the research literature and by thinking creatively about puzzles or problems, but may also be derived from theories and concept maps, translated into hypotheses, and elaborated by considering relevant dialogical arguments, approaching them from different perspectives, and constructing conceptual models. Some combination of these methods will more often than not lead to a good research question, as well as testable claims.

As a final step, sidebar 2.1 provides a checklist of questions designed to determine whether a research question is generally appropriate (O'Leary 2014, 45). It follows most of the steps recommended in this chapter, and also includes additional considerations as you think about developing a research question and making it the centerpiece of a research project.

EXERCISES FOR CHAPTER 2

1. Select one of the following topics and use the research topic/question checklist, steps 1–3, to try to narrow it down to a research question, problem, or hypothesis. Make sure this is grounded in the research literature in some way.
 - a. Binge alcohol use among young adults.
 - b. Political activism among older adults.
 - c. Access to mental health treatment in impoverished communities in the rural United States.
 - d. Divorce and its effects on parent-child relationships.
 - e. Trust in the police in urban areas of the United States and Canada.
 - f. Eating disorders among minority female teenagers.
 - g. Bullying and its consequences for elementary school students.
 - h. Co-occurring mental health disorders among the homeless.
 - i. The successes and failures of nonprofit environmental groups in Europe.
 - j. China-Taiwan political relations in the twenty-first century.
2. If you didn't do this as part of exercise 1, draw a conceptual model of the research question or hypothesis.
3. Choose one of the concepts you would need to measure if you were to pursue a project on this research question or hypothesis. Conduct a literature review to determine how this concept has been measured in previous studies.