Chapter 4

Causation: Can We Say What Caused the Effect?

Chapter Overview

This chapter is about causation, the last of the four pillars of inference: strength, size, breadth, and cause. What can we say about the cause of the pattern in the data? We saw in the study of organ donation in the Preliminaries that the proportion of “Yes” responses was larger when the question asked for a forced choice rather than just offering a chance to opt in. Can we conclude that it was the form of the question that caused the difference in proportion of “Yes” responses? The answer for that study is yes, but the conclusion is valid only because of careful planning by the investigators. (For contrast: Cities with more McDonald’s restaurants record larger numbers of divorces, but it would be bad science to conclude that eating Big Macs increases the likelihood of divorce. Or: Year by year, the larger the number of licensed amateur radio operators, the larger the number of people who are mentally ill. But that doesn’t mean … You get the idea.)

Whereas Chapters 1 and 3 were about the logic of statistic inference (e.g., how strong is the evidence that the difference in the proportion of “Yes” responses was not due to random chance alone and what do we estimate to be the size of the difference), cause, along with breadth (Chapter 2), are about the scope of inference. If we establish that the difference is “real,” the breadth part of Step 5: Formulate conclusions tells us how broadly the conclusion applies. Finally, in this chapter, you will learn whether and when you can conclude that an observed difference is caused by a particular treatment.

Why do cities that sell more Big Macs have more divorces? A related variable is city size. The bigger the city, the more Big Macs are bought, and the more divorces there are. Section 4.1 explores such issues more formally. Section 4.2 tells how to choose a study design to avoid such issues. The key idea is to use a chance process to assign the conditions to be compared, as in the study of organ donation. This introduction of chance into the study design will play a different role than it has in previous chapters.

The main goal of this chapter is to explain when and why you can infer cause. When you can’t, caution is the order of the day.
Section 4.1: Association and Confounding

Introduction

To oversimplify, but not by much, you can think of science as a search for cause-and-effect relationships and for theories that unite them. In this section we illustrate two things to look for as part of using data to find causal relationships. The first is association between two variables. Consider smoking and cancer. Initially, scientists found that smokers had higher rates of lung cancer than did non-smokers. Smoking and cancer were associated. Did that prove that smoking caused cancer? No. For example, some scientists thought there might be a gene that made people both likely to smoke and likely to get lung cancer. Presence of the gene (yes/no) could be a confounding variable. The confounder could explain the association. To conclude that smoking causes lung cancer, you must be able to rule out the effect of possible confounders.

Is association evidence of possible causation? Yes. If smoking causes cancer, there has to be an association. Cancer rates will be higher for smokers. If the two are associated, one might cause the other. Association is necessary, but association alone is not enough to prove cause-and-effect.

Example 4.1: Night Lights and Near-Sightedness

Myopia, or near-sightedness, typically develops during childhood years. Recent studies have explored whether there is an association between development of myopia and the use of nightlights with infants. One study (Quinn, et al., 1999) interviewed parents of 479 children who were seen as outpatients in a university pediatric ophthalmology clinic. One of the questions asked whether the child typically slept with the room light on, with a night light on, or in darkness before age 2. Based on the child’s most recent eye examination, the children were also separated into two groups: near-sighted or not near-sighted. They found a higher percentage of near-sighted children among those using a room light (54.7%) or with a night light (33.6%) compared to children who slept in darkness (10.5%). Pediatricians and parents became concerned that the use of light in the infants’ rooms was leading to near-sightedness.

For this study, the observational units are the 479 children. We have recorded two different variables about each child: whether the child slept with the room light, a night light, or darkness and whether or not the child was near-sighted. These are both categorical variables. The lighting variable has three categories and the eye condition variable has two. The percentages of children with near-sightedness presented earlier are calculated conditional on the amount of light present in the room at night. Table 4.1 displays the data used to calculate the conditional proportions which multiplied by 100 give to percentages.
Table 4.1: Data used to calculate conditional proportions of nearsightedness based on amount of light in bedroom

<table>
<thead>
<tr>
<th></th>
<th>Darkness</th>
<th>Night light</th>
<th>Room light</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near-sighted</td>
<td>18</td>
<td>78</td>
<td>41</td>
<td>137</td>
</tr>
<tr>
<td>Not near-sighted</td>
<td>154</td>
<td>154</td>
<td>34</td>
<td>342</td>
</tr>
<tr>
<td>Total</td>
<td>172</td>
<td>232</td>
<td>75</td>
<td>479</td>
</tr>
</tbody>
</table>

Conditional proportions: 18/172 \approx 0.105 (10.5\%), 78/232 \approx 0.336 (33.6\%), and 41/75 \approx 0.547 (54.7\%). (These calculations will be discussed in more detail in Chapter 5.)

Notice that as the amount of light present in the room at night increases, so does the percentage of children that are near-sighted. Thus we can say that there exists an association between whether or not a child is nearsighted and amount of light used in the child’s room before the age of 2.

**Definition:** Two variables are associated or related if the value of one variable gives you information about the value of the other variable. When comparing groups this means that the proportions or means take different values in the groups.

Here the two variables (amount of light and eye condition) are associated: Knowing the amount of light gives you information about (helps you predict) a child’s later eye condition. For example, children who slept with a night light are much more likely to be nearsighted than are children who slept in the dark.

**Explanatory and Response Variables**

Association often raises questions about possible causal relationships. After seeing the data above we might wonder, “Does sleeping with a night light cause a child’s chance of being nearsighted to increase?” As you’ll soon see, this study does not allow us to answer this question. As the saying goes, “Association is not causation.” All the same, it is useful to have labels for the two different roles of the variables: The light condition is called the explanatory variable, because in this context we think of it as “explaining” the values of the other, the response variable, in this case, eye condition.

**Definitions:** The explanatory variable is the variable we think is “explaining” the change in the response variable and the response variable is the variable we think is being impacted or changed by the explanatory variable. The explanatory variable is sometimes called the independent variable and the response variable is sometimes called the dependent variable.
Chapter 4: Causation: Can we say what caused the effect?

Think about it: Imagine that you are planning a study to look for association between overall health (good, fair, not good) and having a pet. Which variable would be explanatory, and which would be the response?

If you think having a pet increases people’s happiness and likelihood to exercise, then you are considering pet ownership to be the explanatory variable and health to be the response variable. On the other hand, you may want to consider the health of the individual as an explanation for why some people are more likely to have pets than others. This would reverse the role of the explanatory and the response variable.

In another study, diet could be the explanatory variable and whether or not the person has heart disease could be the response. However, in some cases, there is not a clear distinction between the roles of the variables, such as looking at hair color and eye color – it’s not as though one of these variables precedes and perhaps influences the other.

Cause-and-effect

In this study, there is an association because the percentage of near-sighted children increases as the amount of light used increases. But does this convince us that the room light and night light use is causing the children to develop near-sightedness? Drawing such a “cause-and-effect” conclusion like this is much different from establishing an association between the variables.

Think about it: Can you suggest another explanation for the higher percentage of near-sightedness in children using the room light and night light compared to those with no light?

Keep in mind that the association between amount of light in the room and eyesight condition appears to be real. If we wanted to predict whether or not a child was near-sighted, it would be worthwhile to know what type of lighting they used before the age of 2. But that is still a very different statement from saying that it was in fact the lighting condition itself that directly impacted the children’s eye conditions.

One alternative explanation is that children who are already near-sighted are more likely to prefer to sleep with light in their rooms. This reverses the roles of the explanatory and response variables, though keep in mind that the researchers asked about light conditions before the age of two (when children don’t always have as much say about their environment!) and later eye condition.

Another alternative explanation is that near-sightedness is often inherited from parents (genetics). But does this explain the observed association? It does if parents who are near-sighted are more likely to use a room light or a night light with their infants, perhaps needing the extra light to more easily navigate the
room when they enter at night to check on their child. This appears to be a plausible explanation. The problem is we have no way of deciding whether this explanation is better or worse than the cause-and-effect explanation. Both explanations are consistent with the data we have and the association we observed. If the parents’ eye conditions do tend to differ among the children with well-lit rooms, night lights, and darkness, then we have no way of separating out the “effects” of this variable from those of the lighting condition. In such a case, parents’ eyesight is considered a **confounding variable**. The diagram in Figure 4.1 illustrates the confounding. The top panel shows the design of the study: Children (units) are sorted into groups according to the explanatory variable (light condition), and the response (nearsightedness) is measured. The bottom panel shows the potential confounding effect of the parents’ own vision.

**Figure 4.1:** Potential confounding in the study of light and nearsightedness

---

**Definition:** A **confounding variable** is a variable that is related both to the explanatory and to the response variable in such a way that its effects on the response variable cannot be separated from the effects of the explanatory variable.
Keep in mind that there will always be “other variables” floating around in any study. What we are most concerned about is these potential confounding variables that prevent us from isolating the explanatory variable as the only influence on the response variable. For example, eye color is another variable, but it is not likely related to eye condition or lighting use. Or maybe parents use a room light or nightlight more with girls than with boys. But since near-sightedness affects males and females equally, this wouldn’t explain the higher percentages of near-sightedness for the children using light.

The key to discounting a cause-and-effect explanation is to identify a potential confounding variable and to explain how it is linked to both the explanatory variable and the response variable in a way that also explains the observed association.

Of course, this doesn’t mean our explanatory variable is not influencing our response variable. Sometimes the explanation really is causation, as with the association between smoking and lung cancer. But we just can’t feel comfortable jumping to a cause-and-effect conclusion based solely on these kinds of studies (we’ll discuss this more in the next section!). In short, association does not imply causation, but sometimes it can be a pretty big hint.

**Exploration 4.1: Home Court Disadvantage?**

Sports teams prefer to play in front of their own fans rather than at the opposing team’s site. Having a sell-out crowd should provide even more excitement and lead to an even better performance, right? Well, consider the Oklahoma City Thunder, a National Basketball Association team, in its second season (2008-09) after moving from Seattle. This team had a win-loss record that was actually worse for home games with a sell-out crowd (3 wins and 15 losses) than for home games without a sell-out crowd (12 wins and 11 losses). (These data were noted in the April 20, 2009 issue of *Sports Illustrated* in the Go Figure column.)

1. Identify the observational units and variables in this study. Also classify each variable as categorical (also binary?) or quantitative.

2. When did the Thunder have a higher winning percentage: in front of a sell-out crowd or a smaller crowd? Support your answer by calculating the proportion of sell-out games that they won and also the proportion of non-sell-out games that they won. (Write both these proportions as decimals.)

Sell-out crowd:
Smaller crowd:

**Definition:** Two variables are *associated* or related if the value of one variable gives you information about the value of the other variable. When comparing two groups this means that the proportions or means take different values in the two groups.

3. Do the two variables appear to be associated?

Often when a study involves two associated variables, it is natural to consider one the *explanatory variable* and the other the *response variable*.

**Definitions:** The *explanatory variable* is the variable we think is “explaining” the change in the response variable and the *response variable* is the variable we think is being impacted or changed by the explanatory variable. The explanatory variable is sometimes called the independent variable and the response variable is sometimes called the dependent variable.

4. Which would you consider the explanatory variable in this study? Which is the response? (That is, what are the *roles* of these variables in this study?)

There are two possible explanations for this odd finding that the team had a better winning percentage with a smaller crowds:

- The sell-out crowd *caused* the Thunder to play worse, perhaps because of pressure or nervousness.
- The sell-out crowd did *not* cause a worse performance, and some other issue (variable) explains why they had a worse winning percentage with a sell-out crowds. In other words, a third variable is at play, which is related to both the crowd size and the game outcome.

(Of course, another explanation is random chance. Using methods you will learn later, we’ve determined that you can essentially rule out random chance in this case.)

5. Consider the second explanation. Suggest a plausible alternative variable that would explain why the team would be less likely to win in front of a sell-out crowd than in front of a smaller crowd.

(Make sure it’s clear not just that your explanation would affect the team’s likelihood of winning, but that the team would be less likely to win in front of a sell-out crowd compared to a smaller crowd.)
Definition: A confounding variable is a variable that is related both to the explanatory and to the response variable in such a way that its effects on the response variable cannot be separated from the effects of the explanatory variable.

6. Identify the confounding variable based on your suggested explanation in #5. Explain how it is confounding—what is the link between this third variable and the response variable, and what is the link between this third variable and the explanatory variable? (Hint: Remember that this variable has to be recorded on the observational units: home games for the Thunder.)

Another variable recorded for these data was whether or not the opponent had a winning record the previous season. Of the Thunder’s 41 home games, 22 were against teams that won more than half of their games. Let’s refer to those 22 teams as strong opponents. Of these 22 games, 13 were sell-outs. Of the 19 games against opponents that won less than half of their games that season (weak opponents), only 5 of those games were sell-outs.

7. Was the Thunder more likely to have a sell-out crowd against a strong opponent or a weak opponent? Calculate the relevant (conditional) proportions to support your answer.

When the Thunder played a strong opponent, they won only 4 of 22 games. When they played a weak opponent, the Thunder won 11 of 19 games.

8. Was the Thunder less likely to win against a strong opponent than a weak one? Again calculate the relevant (conditional) proportions to support your answer.

9. Explain how your answers to #7 and #8 establish that strength of opponent is a confounding variable that prevents drawing a cause-and-effect conclusion between crowd size and game outcome.
10. Summarize your conclusion about whether these data provide evidence that a sell-out crowd \textit{caused} the Thunder to play worse. Write as if to a friend who has never studied statistics. Be sure to address the fact that the Thunder had a much smaller winning percentage in front of a sell-out crowd.

Confounding explains why you cannot draw a cause-and-effect conclusion from association alone: The groups defined by the explanatory variable could differ in more ways than just the explanatory variable when confounding is present. The diagram in Figure 4.2 illustrates the confounding in the Thunder study. The top panel shows the study design: Observational units (home games) are sorted into groups according to the explanatory variable (whether or not the arena was sold out). Then the response (win/lose) was observed. The bottom panel shows the confounding: sell out crowds tended to be against stronger opponents; weaker opponents tended not to sell out the arena.

\textbf{Figure 4.2:} Confounding in the study of the home court disadvantage


**Section 4.1 Summary**

Two variables are associated (related) if the values of one variable provide information about (help you predict) the values of the other variable.

Studies that involve two variables often distinguish between the roles played by the variables.
- The **explanatory variable** is the one that is suspected of possibly affecting the other.
- The **response variable** is the one that may be affected by the other.

A **confounding variable** is one whose effects on the response cannot be separated from those of the explanatory variable.

The possible presence of confounding variables is the reason that association alone does not justify a conclusion that differences in the explanatory variable *cause* differences in the response variable.

**Section 4.2: Observational Studies Versus Experiments**

**Introduction**

Many, if not most, scientific studies involve a search for cause-and-effect relationships between variables. Studies often start by looking for an association between explanatory and response variables, but the possible presence of confounding variables prevents us from being able to draw conclusions about cause.

The main challenge in designing a good study is to avoid the effect of confounding variables. If you know in advance about a particular confounding variable, you may be able to choose a design that balances the effects of that confounder. But many confounders are “lurking” variables – either not known or known but not measured. To deal with these unseen confounders, we rely on a strategy called randomization or random assignment. We use a chance device like coin tosses or drawing names from a hat to decide which subjects go into which groups. Random assignment in this way gives approximate balance for all possible confounding variables.

The importance of random assignment to avoid confounding effects leads to one of the most important distinctions in all of science: the difference between an observational study and an experimental study. In an observational study, the values of the explanatory variable are simply observed – you don’t get to choose. In an experiment, you – the investigator – get to assign the conditions to the different subjects.
Example 4.2: Lying on the Internet

Are people more likely to tell lies with email than with pencil and paper? A study reported at the August 2008 meeting of the Academy of Management involved 48 graduate students in business who participated in a “bargaining” game. Researchers kept track of whether or not students misrepresented (lied about) the size of the pot when they were negotiating with other players.

Step 1: Ask a research question.

Does the Internet encourage lying? Investigators thought that the Internet might encourage users to feel anonymous, making them more likely to lie. To narrow their hypothesis, they decided to compare two groups of students playing a competitive game. Students in one group would play on the Internet by using e-mail as their method of communication. Students in the other group would use pencil and paper as their method of communication. The researchers hypothesized that those who played on the Internet would lie more often.

Step 2: Design a study and collect data.

Definition: In an observational study, the groups you compare are “just there,” that is, they are defined by what you see rather than by what you do. In an experiment, you actively create the groups by what you choose to do to the people or experimental units. More formally, you assign the conditions to be compared. These conditions may be one or more “treatments” or a “control” (a group you do nothing to).

Think about it: Identify the explanatory and response variables in this study. Is the study observational or experimental? How can you tell?

The explanatory variable is the communication method. The response variable is whether or not the person lied. Both variables are categorical. This was an experiment rather than an observational study because the researchers actively intervened to determine the communication method used by the students.

Think about it. Suppose there were 30 male and 18 female subjects. Imagine if you had all of the 30 men in the sample play on the Internet, and the 18 women used pencil and paper. If you were concerned that males might be more competitive and therefore more likely to lie in order to win, how would this limit your ability to draw conclusions about the initial research question from this study?
If all of the men played on the Internet and all of the women used pencil and paper, then if we saw differences in the two groups in terms of the proportion who lied, you would not be able to say whether it was gender or the format of the game (Internet vs. paper and pencil) that was causing the difference.

**Think about it.** How could you address this concern?

One strategy would be to assign half the men and half the women to each group. That way, both groups will have 15 male and 9 female subjects. The groups would be balanced with respect to gender.

**Key idea:** A comparative study is balanced with respect to a possible confounding variable if the distribution of the variable is the same for each group in the study.

Note that if you have many confounders, it will be hard to balance them all. Worse yet, if you have “lurking” variables – unmeasured or unrecognized potential confounders, the challenge might seem impossible. How can you balance what you can’t even see?

Fortunately, random assignment offers a solution. If you know about and measure a potential confounder in advance, you may be able to design a study that gives exact balance for that variable. If you use a chance device to assign units to groups, you may not get exact balance. In fact, most of the time you won’t. But – this is the key – you will get approximate balance for all potential confounders, including the ones you don’t know about. If you toss a fair coin, you tend to get half heads, half tails, or at least an outcome pretty close to that. If you use a coin toss to assign subjects to two groups, you tend to get half the men in each group, and half the women in each group. You also, at the same time, tend to get half the young people in each group, half the vegetarians in each group, half the pet owners in each group, and so on.

**Key idea:** Randomly assigning experimental units to groups tends to balance out all other variables between the groups. Any variables that could have an effect on the response should be equalized between the two groups and therefore should not be confounding.

With a randomized experiment, if you observe a difference in lying between the Internet players and the pencil and paper players, there are only two explanations: Either there is a cause-and-effect relationship, or else the difference was just due to random chance, stemming from the random assignment process. We will be willing to assume there are no potential confounding variables to explain the difference.

The diagram in Figure 4.3 illustrates this. The top panel shows the possible confounding effect if the study design had not been randomized. We want to know whether differences in the explanatory variable cause differences in the response (dotted arrow). However, a confounding variable would affect the
response variable and be associated with the explanatory variable, making it impossible to separate the effects of explanatory and confounding variable on the response. The bottom panel shows how randomization removes the effect of the potential confounders.

**Figure 4.3:** Randomization creates approximate balance and tends to eliminate confounding effects.

![Diagram of randomization effects](image)

**Definitions:** In a *randomized experiment*, you use a chance device to make the assignments. The units in an experiment are often called *experimental units*. The role of the random assignment is to balance out potentially confounding variables among the explanatory variable groups, giving us the potential to draw cause-and-effect conclusions.

Note our change in terminology: we now will sometimes call the individuals in the study *experimental units* instead of *observational units*, because they are in an experiment instead of an observational study.

For this study each student was randomly assigned to one of two groups. Students in Group 1 used the Internet. Those in Group 2 used pencil and paper. You can think of the random assignment like drawing names out of a hat: Put all the names on slips of paper, put them in a hat, mix thoroughly, take them out one at a time, and put each of them into one of the two groups. (Internet, paper, Internet, paper, etc.)

Compare this study with the two studies in the last section, Night Lights and Home Court Disadvantage. Notice that in both of those studies, it was not possible for researchers to *assign* the conditions to be compared. By the time the children were seen at the eye clinic, it was too late to tell their parents what kind of light to use. For the study of sleep condition and vision, each child was an observational unit.
Values of the explanatory variable (level of light) were observed. They could not be chosen by the investigators. For the Thunder study, each home game was an observational unit. The strength of the opponent was observed, not assigned. If the home game is against the LA Lakers, there’s no way you can make them be a weak team just for the sake of your research.

**Step 3: Explore the data.**

The researchers found that 24 of 26 (92%) who used email were guilty of lying about the potsize, compared to 14 of 22 (64%) who used paper and pencil. (We will discuss such *conditional* percentages in more detail in Chapter 5.)

**Step 4: Draw inferences beyond the data.**

When you are comparing two groups like this, if you find a difference between the groups, you again need to consider several potential explanations. One possible explanation is that the people feel more comfortable lying by email, perhaps because it seems anonymous. In other words, the format of the game influenced how people responded. But of course another important possible explanation to consider is random chance: Maybe it was just luck of the draw that led to the big difference in the proportions of liars between the two groups. We claimed that random assignment should balance out the two groups, but there is still a chance that the random assignment process itself created slightly different groups. Notice the randomness we are considering here stems from the random assignment, not random choices or random sampling. As it turns out, for this study we can rule out this explanation. Using methods you will learn in the next chapter, it’s highly unlikely that the emailing players would lie this much more frequently, relative to the pencil and paper players, just by chance alone (i.e., the p-value is small, 0.015).

**Step 5: Formulate conclusions.**

Now it is time to formulate our final conclusions. We found a statistically significant difference between the two groups. Does this allow us to conclude that the format of the game influenced how people responded? Maybe those who use email tend to be younger and younger individuals are more likely to lie. Or there could be other personality traits that differ between the individuals in the email group and the paper and pencil group. Are these potential confounding variables?

But if you believe what we told you in the previous section, because the study used random assignment to form the email group and the paper and pencil group, we don’t expect any systematic differences between the Internet group and the pencil-and-paper group. This is because the random assignment process should create groups that are similar to each other on all respects (e.g., gender, age, propensity for lying, familiarity with technology) apart from the explanatory variable of interest.
This means that when you observed a statistically significant difference in how the groups respond or behave (too large to be expected to happen by the random chance inherent in the random assignment process), then a plausible explanation is because of the group the experimental units were assigned to.

In formulating our conclusions, we still want to consider the breadth of our study. Because the subjects in this study business school graduate students, and volunteers, we should be very cautious in generalizing these results to a larger population. See FAQ 4.2.1 and Figure 4.5 at the end of this section for more details and a helpful table.

**FAQ 4.2.1- Randomness and its implications**

Q: You make it nice and clear. If a research study doesn’t randomize the assignment, I can’t conclude anything about cause. If it doesn’t use random sampling, I can’t generalize. No chance, no inference. Right?

A: Sorry, that’s much too black and white.

Q: But wait. The theory is simple. List your units, then use chance to assign treatments. Isn’t that what you said?

A: Yes, that’s the ideal way to learn about cause-and-effect. But you can’t always do that in practice. Think about smoking and lung cancer.

Q: Those were observational studies, not randomized experiments. How did scientists conclude that smoking causes cancer?

A: Remember that a main purpose of randomization is to protect against confounding. It’s possible to design observational studies that also protect against confounding. It’s a lot harder, but it’s possible. It was over many, many of these more complicated observational studies, with more advanced statistical methods, and with support from other types of studies linking carcinogens to cigarettes, that it became an accepted scientific truth that smoking causes lung cancer.

Q: And random sampling?

A: Here, too, it’s not black and white. Randomizing the selection protects against bias in your sampling method, so if you don’t randomize, you have to be very careful about what larger group your sample represents, but that’s where judgment comes in. As in all of statistics, the mathematical theory is an ideal,
but reality is almost always more complicated. You don’t want to throw out the baby (the data) just because the bath water (design) is not ideal.

**Types of Experiments**

We just saw that randomized experiments (experiments where the explanatory variable is manipulated by the researchers and assignment of observational units to the explanatory variable’s groups is determined randomly) can potentially lead to cause-and-effect conclusions between the explanatory and response variables. What about experiments which manipulate the explanatory variable, but not randomly? These are often called *quasi-experiments*. For example, suppose we wanted to compare student learning gains when using a new curriculum to that of an old curriculum. It would be difficult to assign students randomly to which class they take. So the instructors might take the results from pre- and post-tests from students that used the old curriculum one year and compare these to the results from pre- and post-tests from students that used the new curriculum another year. Although this type of experiment may be the best option when random assignment is not feasible, there may still be reasons other than the explanatory variable (confounding variables) for the difference in outcomes between the two groups (e.g., time of day class is offered, percentage of upper-division students in the course). Therefore, even though this is considered an experiment, it is not a randomized experiment, and we should be cautious before drawing any cause-and-effect conclusions.

Also keep in mind that some explanatory variables of interest don’t lend themselves to randomized experiments. For example, gender can’t be randomly imposed on individuals, and other variables, such as smoking behavior, would not be ethical to manipulate!

**Other Considerations**

**Think about it:** In this study, the subjects were randomly assigned to the Internet group and the paper and pencil group. Do you think the researchers told the subjects in advance that they would be assigned to one of these two groups?

In most research studies, subjects are “blind” to the other treatment conditions. In fact, they often don’t know that there are multiple conditions or exactly which condition they are in. For example, suppose I want to test a new diet pill. If I give you a pill and tell you it’s going to give you more energy, you may feel a positive effect even if there is no active ingredient in the pill. If one group had this psychological suggestion and not the other, then that would be a confounding variable. So the second group, the “control group” would often be given a “placebo pill” (an empty treatment) so none of the subjects knew which group they were in. In fact, for something like measuring energy which may require subjective
judgment (as opposed to heart rate), often the person making the measurements wouldn’t know which group the subjects were in, to prevent any bias creeping into their judgments as well, making the study ‘double-blind.’ A “randomized, double-blind, placebo controlled experiment” have been considered by some to be the gold standard in scientific studies.

**Definition:** In a *double-blind* study, neither the subjects nor those evaluating the response variable know which treatment group the subject is in.

**Exploration 4.2: Have a Nice Trip**

An area of research in biomechanics and gerontology concerns falls and fall-related injuries, especially for elderly people. Recent studies have focused on how individuals respond to large postural disturbances (e.g., tripping, induced slips). One question is whether subjects can be instructed to improve their recovery from such disturbances.

Suppose researchers want to compare two such recovery strategies, lowering (making the next step shorter, but in normal step time) and elevating (using a longer or normal step length with normal step time). Subjects will have first been trained on one of these two recovery strategies, and they will be asked to apply it after they feel themselves tripping. The researchers will then induce the subject to trip while walking (but harnessed for safety), using a concealed mechanical obstacle.

Suppose the following 24 subjects have agreed to participate in such a study:

- **Females:** Alisha, Alice, Betty, Martha, Audrey, Mary, Barbie, Anna
- **Males:** Matt, Peter, Shawn, Brad, Michael, Kyle, Russ, Patrick, Bob, Kevin, Mitch, Marvin, Paul, Pedro, Roger, Sam

1. One way to design this study would be to assign the eight females to use the elevating strategy and the sixteen males to use the lowering strategy. Would this be a reasonable strategy? Why not?

2. One way to deal with this issue is to assign 4 females and 8 males to each group. Show how the proportion of males in each group is the same.
3. Now, if you saw a difference in the proportion of trips in the two groups, could it be because of gender? Why or why not? Could it be due to other variables, distinct from the recovery strategy? Why or why not?

4. Because there will always be more potential confounding variables which are distributed unevenly between the groups being compared, identify a better method for deciding who uses which strategy.

**Definitions:**

In an **observational study**, the groups you compare are “just there,” that is, they are defined by what you see rather than by what you do. In an **experiment**, you actively create the groups by what you choose to do to the people or experimental units. More formally, you assign the conditions to be compared. These conditions may be one or more “treatments” or a “control” (a group you do nothing to).

5. Let’s explore the process of random assignment to determine whether it does “work.” First, let’s focus on the gender variable. Suppose we put each person’s name on a slip, put those slips in a hat and mix them up thoroughly, and then randomly draw out 12 slips for names of people to assign to the elevating strategy. What proportion of this group do you expect will be male? What proportion of the lowering strategy do you expect will be male? Do you think we will always get an 8/8 split (8 males in each treatment group)?
6. To repeat this random assignment a large number of times to observe the long-run behavior, we will use the Randomizing Subjects applet. Open the applet, and press the Randomize button. What proportion of subjects assigned to Group 1 are men? Of Group 2? What is the difference in these two proportions?

You will notice that the difference in proportions of males is shown in the dotplot in the bottom graph. In this graph, each dot represents one repetition of the random assignment process where we are recording the difference in proportions of men between the two groups.

7. Press the Randomize button again. Was the difference in proportions of men the same this time?

8. Change the number of replications from 1 to 198 (for 200 total), uncheck the Animate option, and press the Randomize button. The dotplot will display the difference between the two proportions of men for each of the 200 repetitions of the random assignment process. Where are these values centered?

9. Does random assignment always equally distribute/balance the men and women between the two groups? Is there a tendency for there to be a similar proportion of men in the two groups? Explain.

Definition: A comparative study is balanced with respect to a possible confounding variable if the distribution of the variable is the same for each group in the study.

10. Prior research has also shown that the likelihood of falling is related to variables such as walking speed, stride rate, and height, so we would like the random assignment to distribute these variables equally between the groups as well. In the applet, use the pull-down menu to switch from the gender variable to the height variable. The dotplot now displays the differences in average height between Group 1 and Group 2 for these 200 repetitions. In the long-run, does random assignment tend to equally distribute the height variable between the two groups? Explain.
11. Suppose there is a “balance gene” that is related to people’s ability to recover from a trip. We didn’t know about this gene ahead of time, but if you select the “Reveal gene?” button and then select “gene” from the pull-down menu, the applet shows you this gene information for each subject and also how the proportions with the gene differ in the two groups. Does this variable tend to equalize between the two groups in the long run? Explain.

12. Suppose there were other “x-variables” that we could not measure such as stride rate or walking speed. Select the “Reveal both?” button and use pull-down menu to display the results for the x-variable (x-var). Does random assignment generally succeed in equalizing this variable between the two groups or is there a tendency for one group to always have higher results for the x-variable? Explain.

**Key idea:** Randomly assigning experimental units to groups tends to balance out all other variables between the groups. Any variables that could have an effect on the response should be equalized between the two groups and therefore should not be confounding.

The diagram in Figure 4.4 illustrates this. The top panel shows the possible confounding effect if the study design had not been randomized. We want to know whether differences in the explanatory variable cause differences in the response (dotted arrow). However, a confounding variable would affect both the explanatory and response variables, making it impossible to separate the effects of explanatory and confounding variables on the response. The bottom panel shows how randomization removes the effect of the potential confounders.
13. Suppose this study finds a statistically significant difference between the two groups. What conclusion would you draw? For what population? What additional information would you need to know?

14. As in #13, if you obtain a statistically significant result, what does that suggest about the potential for a cause-and-effect relationship? Why?

Another study: Cursive Handwriting

An article about handwriting appeared in the October 11, 2006 issue of the Washington Post. The article mentioned that among students who took the essay portion of the SAT exam in 2005-6, those who wrote in cursive style scored significantly higher on the essay, on average, than students who used printed block letters.
15. Identify the observational units in this study, as well as the explanatory and response variables. Also classify each variable as categorical or quantitative.

Observational units:

Explanatory variable: Type:

Response variable: Type:

16. Explain how you know that this was an observational study.

17. Is it reasonable to conclude that using a cursive writing style caused higher scores on the essay, or can you think of an alternative explanation for why students who wrote in cursive style scored higher on average than students who write with block letters? In other words, can you think of other ways in which the cursive and block letter groups might have systematically differed and identify a potential confounding variable?

The same Washington Post article also mentioned a different study in which the identical essay was shown to many graders, but some graders were randomly chosen to see a cursive version of the essay and the other graders were shown a version with printed block letters. The average score assigned to the essay with the cursive style was significantly higher than the average score assigned to the essay with the printed block letters.
18. How does this study differ from the original one? Explain.

19. Would you be willing to draw a cause-and-effect conclusion from this study, as opposed to the original study? Explain why.

See the Section summary for a helpful table (Figure 4.5) which summarizes the types of conclusions you can draw with and without random sampling and with and without random assignment.

The “placebo effect” has been found in numerous studies: when subjects are told something good is going to happen, they often have a positive response even if nothing is actually done to them. For this reason, subjects are often kept “blind” as to which treatment group they are placed in, for example by giving one group the actual treatment and the other group a “fake” treatment, like a placebo (an empty pill). Placebo treatments have even been used in studies about knee surgery!

20. Discuss why it was important in this second study for the scorers of the essays to not know that there was no difference in the essays given to different individuals other than the writing style.

21. Discuss how this study also reveals that those evaluating the response variable in a study, if it requires any subjective judgment, should be “blind” to which treatment group individuals are in.

**Definition:** In **double-blind** studies, neither the subjects nor those evaluating the response variable know which treatment group the subject is in.
Section 4.2 Summary

A well-designed experiment uses random assignment to determine which observational (experimental) units go into which explanatory variable groups.

- The goal of random assignment is to produce groups that are as similar as possible in all respects except for the explanatory variable.
- Then if the groups can be shown to differ significantly on the response variable, the difference can be attributed to a cause-and-effect relationship between the explanatory and response variables.

Random assignment is a very different use of randomness from random sampling, with different implications for scope of conclusions.

- Random sampling aims to select a representative sample from a population, so that results about the sample can be generalized to the larger population.
- Random assignment aims to produce similar treatment groups, so that a significant difference in the response variable between groups can be attributed (causally) to the explanatory variable.

Figure 4.5 summarizes how these two types of randomness lead to different impact on the scope of conclusions.

In a double-blind experiment, neither the subjects nor those evaluating the response variable know which group the subject is in. This is to guard against biases such as the placebo effect, where the mere power of suggestion can influence subjects’ responses.
**Figure 4.5:** Study design factors and their impact on scope of conclusions (adapted from Ramsey and Schafer, *The Statistical Sleuth*)

**Chapter 4 Summary**

In the first section, we saw that two variables are associated when the value of one variable provides information about (helps us predict) the value of the other variable. In a search for causal relationships, we call the possible cause variable *explanatory* and the outcome variable the *response*. However, association alone does not prove causation because of the potential for confounding variables (variables related to both the explanatory and response variables in such a way that their effects on the response variable are intertwined with those of the explanatory variable).
In Section 2, we distinguished between observational and experimental studies. In an observational study, you simply record what’s there. You don’t get to choose which group each observational unit belongs to. In an experiment, you choose and assign the conditions that define the groups to subjects or other units. Ideally, this is done using a chance device like coin tosses or drawing names out of a hat. Random assignment has the advantage that it tends to ensure balance for all possible confounding variables, both known and unknown. For a properly randomized experiment, there are only two explanations for any observed difference between groups: either the assigned explanatory variable conditions caused the difference, or else the difference is a coincidence due to the random assignment. (In the next chapter, you will see when and how we can find evidence against random chance as an explanation.) Not all experiments where the explanatory variable is manipulated are randomized. Without randomized assignment, conclusions about cause are not justified.

You have now seen that there are two main “Scope of inference” questions that should be asked as part of Step 5 of the Statistical Investigation Method.

- First, is the result from the sample generalizable to a larger population? The answer is yes, if the sample is representative of the population, and we feel more comfortable believing such if the sample is a random sample.
- Second, can we conclude that there is a cause-and-effect relationship between the explanatory and response variable? The answer is yes, if the study is a randomized experiment, meaning that the value of the explanatory variable is randomly assigned, guarding against potential confounding variables.
Chapter 4 Glossary

**association**
Two variables are associated or related if the distribution of the response variable differs across the values of the explanatory variable ................................................................. 4, 7

**cause-and-effect**
In well-designed studies (randomized experiments), can conclude the explanatory variable is causing the effect seen in the response variable ........................................................................................................... 4

**confounding variable**
A confounding variable is a variable that is related to both the explanatory and response variable in such a way that its effects on the response variable cannot be separated from the explanatory variable. .... 5, 8

**control treatment**
A *do nothing* treatment in an experiment ........................................................................................................... 12, 19

**double-blind**
A study design where neither the subjects nor those evaluating the response know which treatment group the subject is in ................................................................. 18, 25

**experiment**
A study in which researcher actively assign subjects to treatment groups ............................................. 12, 19

**experimental units**
What observational units are called in an experiment study ........................................................................ 14, 19

**explanatory variable**
The variable that, if the alternative hypothesis is true, is explaining changes in the response variable; sometimes known as the independent or predictor variable ................................................................. 3, 7

**observational study**
Studies in which researchers observe individuals and measure variables of interest, but do not intervene in order to attempt to influence responses. ................................................................. 12, 19

**quasi-experiments**
Experiments that manipulate the explanatory variable, but not randomly ........................................ 17

**randomized experiment**
An experiment where experimental units are randomly assigned to two or more *treatment* conditions and the explanatory variable is actively imposed of the subjects. ................................................................. 14, 19
**response variable**

The variable that, if the alternative hypothesis is true, is impacted by the explanatory variable; sometimes known as the dependent variable.

**treatments**

The assigned conditions in an experiment.
Section 4.1

GPAs and pulling all-nighters*

4.1.1 Exercises 4.1.1-3 refer to the following study. A study published in the January 2008 issue of the journal *Behavioral Sleep Medicine* involved a survey of 120 students at St. Lawrence University, a small liberal arts college in upstate New York. Researchers found that students who claimed to have never pulled an all-nighter had average GPAs of 3.1, compared to 2.9 for those students who do claim to have pulled all-nighters.

a) In this context, what is the observational unit?
   A. Researcher
   B. Student
   C. GPA
   D. Whether or not have pulled an all-nighter

b) Which of the following is the explanatory variable?
   A. Whether or not have pulled an all-nighter
   B. Student
   C. GPA
   D. Those students who have pulled all-nighters

c) Which of the following is the response variable?
   A. GPA
   B. Average GPA
   C. Whether or not have pulled an all-nighter
   D. Students who have pulled all-nighters tend to have lower GPAs, on average

d) The explanatory variable is _________ (pick one: CATEGORICAL  QUANTITATIVE)
e) The response variable is _________ (pick one: CATEGORICAL  QUANTITATIVE)

4.1.2 Reconsider the previous study about GPAs and pulling all-nighters. Suppose that the difference between these two averages is shown to be statistically significant. Which of the following is a potential confounding variable, and provides an alternative explanation for why the all-nighter group would have a significantly lower average GPA?
   A. Students
   B. Type of lifestyle choices made by student
C. Difficulty level of classes offered at St. Lawrence
D. Class sizes

4.1.3 Reconsider the previous exercise about the potential confounding variable that provides an alternative explanation for why the all-nighter group would have a significantly lower average GPA. Sketch a well-labeled diagram showing how this explanation works.

Using candy cigarettes

4.1.4 In a study published in a 2007 issue of the journal *Preventive Medicine*, researchers found that smokers were more likely to have used candy cigarettes as children than non-smokers were. When hearing about this study, John responded: “But isn’t the smoking status of the person’s parents a confounding variable here?” When Karen asked what he meant, John said: “Children whose parents smoke are more likely to become smokers themselves when they become adults.” What else does John need to say in order to explain how the parents’ smoking status could be a confounding variable in this study?

Marriage views*

4.1.5 Do different generations view marriage differently? A 2010 survey conducted by the Pew Research Center asked the following question of each participant, “Is marriage becoming obsolete?” Here is how the data from Pew Research Center’s study worked out:

<table>
<thead>
<tr>
<th>Generation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Millennial (ages 18-29)</td>
<td>536</td>
</tr>
<tr>
<td>Gen X (ages 30-45)</td>
<td>729</td>
</tr>
<tr>
<td>Boomers (ages 46-65)</td>
<td>1146</td>
</tr>
<tr>
<td>Age 65+</td>
<td>211</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Marriage obsolete?</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>1018</td>
</tr>
<tr>
<td>No</td>
<td>1604</td>
</tr>
</tbody>
</table>

a) Identify the observational units.
b) Identify the explanatory and response variables. Also identify the type (categorical or quantitative) of each.
c) The conditional proportions of people answer Yes, marriage is become obsolete in each generation of the sample are: 0.44 (Millenial), 0.429 (Gen X), 0.350 (Boomers) and 0.322 (Age 65+). Based on these conditional proportions calculated in does it appear that different generations tend to view marriage differently? Explain.
Heart attacks

4.1.6 Studies conducted in New York City and Boston have noticed that more heart attacks occur in December and January than in all other months. Some people have tried to conclude that holiday stress and overindulgence causes the increased risk of heart attack.
Identify a confounding variable whose effect on heart attack rate might be confounded with that of the month, providing an alternative explanation for the increased risk of heart attack in December and January. Also, include a diagram showing how this explanation works.

Number of TVs and life expectancy*

4.1.7 Exercises 4.1.7-8 contain questions about the following study. Analysis of data on a sample of 22 countries shows a strong positive association between the average life expectancy in the country and the number of TVs per 1000 people in the country. Meaning, countries that have more TVs per 1000 people also have higher average lifetimes. Based on this, is it reasonable to conclude that by sending TVs to countries with lower life expectancies, we can increase their inhabitants’ lifetimes? If yes, explain how. If no, give an alternative explanation for the association between number of TVs and life expectancy.

4.1.8 Regarding the previous question about number of TVs and life expectancy.
a) Identify the observational units.
   A. People
   B. TVs
   C. Countries
   D. Life expectancy

b) Identify the explanatory variable.
   A. People
   B. TVs
   C. Countries
   D. Life expectancy

c) The explanatory variable is ________ (pick one: CATEGORICAL QUANTITATIVE)

d) Identify the response variable.
   A. People
   B. TVs
   C. Countries
   D. Life expectancy
e) The response variable is _______ (pick one: CATEGORICAL QUANTITATIVE)

**Mediterranean diet**

4.1.9 Based on a four-year (2003 – 2007) study of over 30,000 people who were 45+ years, where all individuals were followed up regularly for diet and health changes, researchers Tsivgoulis et al. (*Neurology*, 2013) reported that the Mediterranean diet was linked to better memory and cognitive skills.

a) Identify the explanatory and response variables.
b) Identify a confounding variable whose effect on memory and cognitive skills might be confounded with that of the explanatory variable identified in (a), providing an alternative explanation for the improved memory and cognitive skills among those on the Mediterranean diet. Also, include a diagram showing how this explanation works.

**VCR Ownership**

4.1.10 A Gallup poll conducted in December 2013 found that 74% of respondents who were 65 years or older owned a VCR compared to 41% of 18 to 29 years old.

a) Identify the explanatory variable.
   A. Respondents
   B. Age group
   C. Whether own a VCR
   D. Are older folks more likely to own a VCR?
b) Identify the response variable.
   A. Respondents
   B. Age group
   C. Whether own a VCR
   D. Are older folks more likely to own a VCR?
c) Do the results from Gallup’s poll indicate that there is an association between owning a VCR and age? Explain.
**Hormone Therapy**

4.1.11 In a study published in *Preventive Medicine* (1991), researchers Stampfer and Colditz observed that women who underwent hormone replacement therapy (HRT) showed a lower risk of coronary heart diseases.

a) Identify the explanatory and response variables.

b) Identify a confounding variable that provides an alternative explanation for the reduction in CHD risk among women who underwent HRT, compared to those who didn’t undergo HRT. Also, include a diagram showing how this explanation works.

**Eating Breakfast***

4.1.12 Based on a survey of almost 3000 adults, researchers Wyatt et al. (*Obesity Research*, 2002) reported that those who ate breakfast regularly tended to be more successful at maintaining their weight loss.

a) Identify the explanatory and response variables.

b) Identify a confounding variable that provides an alternative explanation for the improved success rate of weight loss among regular breakfast eaters, compared to those who don’t eat breakfast regularly. Also, include a diagram showing how this explanation works.

**Dinner and drugs**

4.1.13 In June 2012, *Time* magazine ran an article titled “Do Family Dinners Really Reduce Teen Drug Use?” Answer the following questions in this context.

a) Identify the observational units.

b) Identify the explanatory and response variables.

c) Identify a confounding variable that provides an alternative explanation for the lower drug use among teens whose families ate dinner together, compared to those whose didn’t. Also, include a diagram showing how this explanation works.

**Smoking while pregnant***

4.1.14 Many studies have shown that women who smoke while pregnant tend to have babies that weigh significantly less, on average, than women who do not smoke while pregnant.

a) Identify the population(s) of interest in these studies.

b) Identify the explanatory variable in these studies. Also classify this variable as categorical or quantitative.

c) Identify the response variable in these studies. Also classify this variable as categorical or quantitative.
d) Socio-economic status is potentially a confounding variable in these studies. Explain what it means for this to be a confounding variable in this context, and describe how this could provide an alternative explanation to concluding that smoking while pregnant causes lower birth weight in babies.

**Spanking and IQ**

4.1.15 Studies have shown that children in the U.S. who have been spanked (yes/no) have a significantly lower IQ score on average than children who have not been spanked.

a) Identify the explanatory variable in these studies. Also classify this variable as categorical or quantitative.

b) Identify the response variable in these studies. Also classify this variable as categorical or quantitative.

c) Identify a confounding variable that provides an alternative explanation for the lower average IQ score for children who have been spanked, compared to those who haven’t.

**Overweight friends***

4.1.16 In July 2013, Gallup surveyed 2027 randomly selected U.S. adults. They found that of the 921 people who described themselves as overweight, 424 reported having some or many friends/family member who were overweight (424/921 = 0.46). Also, of the 1106 people who described their weight as being “about right,” 332 reported having some or many friends/family member who were overweight (0.30). Use the conditional proportions to explain whether the data indicate that there is an association between whether one perceives oneself as being overweight and how many friends/family members one perceives as being overweight.

**Children and lifespan**

4.1.17 Do men with children tend to live longer than men without children? To investigate, a group of students at Cal Poly randomly sampled men from the obituaries page on the San Luis Obispo Tribune’s webpage between June – November 2012. For each man selected, they noted the age at which the person died, and number of children under “survived by.”

a) Identify the explanatory variable in this study. Also classify this variable as categorical or quantitative.

b) Identify the response variable in this study. Also classify this variable as categorical or quantitative.

Given below is a set of dotplots displaying the data:
c) Based on the above dotplots, is there any indication of an association between the explanatory variable and the response variable? Explain how you are deciding.

d) Identify a confounding variable that provides an alternative explanation for the higher lifetime values among men who have children, compared to those who don’t.

Happiness and income*

4.1.18 To investigate whether there is an association between happiness and income level, we will use data from the 2002 General Social Survey (GSS), cross-classifying a person’s perceived happiness with their family income level. The GSS is a survey of randomly selected U.S. adults who are not institutionalized. Here are the data:

<table>
<thead>
<tr>
<th>Happy?</th>
<th>Income</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Above average</td>
<td>Average</td>
</tr>
<tr>
<td>Very happy</td>
<td>110</td>
<td>221</td>
</tr>
<tr>
<td>Pretty happy</td>
<td>159</td>
<td>372</td>
</tr>
<tr>
<td>Not too happy</td>
<td>21</td>
<td>53</td>
</tr>
<tr>
<td>Total</td>
<td>290</td>
<td>646</td>
</tr>
</tbody>
</table>

a) Identify the explanatory variable. Is it a categorical variable or a quantitative variable?
b) Identify the response variable. Is it a categorical variable or a quantitative variable?
c) Among above average income individuals 0.379 are Very happy (110/290), among average income individuals 0.342 are Very happy (221/646) and among below average individuals 0.195 are very happy (83/426). Do the data provide any indication of an association between happiness and income level? Why or why not?
d) Is it okay to conclude that income affects happiness? If yes, explain why. If no, then identify a confounding variable that provides an alternative explanation for the association between happiness and income level.

Gender and body image

4.1.19 A 2013 Gallup poll asked randomly selected U.S. adults whether they wanted to stay at their current body weight or change. One purpose was to investigate whether there was any difference between men and women with regard to this aspect.
Political party and evolution*

4.1.20 Is there an association between political party affiliation and beliefs about human evolution? A survey of a random sample of U.S. adults by The Pew Research Center for the People and the Press, conducted in March – April 2013 recorded each participant’s political party affiliation and belief about human evolution. Here is how the data from Pew Research Center’s study worked out:

<table>
<thead>
<tr>
<th>Belief about human evolution</th>
<th>Political party affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Republican</td>
</tr>
<tr>
<td>Humans have evolved over time</td>
<td>196</td>
</tr>
<tr>
<td>Humans have existed in present form</td>
<td>218</td>
</tr>
<tr>
<td>Don’t know</td>
<td>41</td>
</tr>
</tbody>
</table>

a) Identify the explanatory and response variables. Also identify the type (categorical or quantitative) of each.

b) Based on the table, 0.43 Republicans believe Humans have evolved over time, 0.67 Democrats and 0.65 Independents. Do the data provide any indication of an association between political party affiliation and beliefs about human evolution?

Education and the census

4.1.21 A survey of a random sample of U.S. adults by The Pew Research Center for the People and the Press, conducted in early January 2010 recorded each participant’s highest level of education completed, and whether they knew that responding to the Census was required by law. Of the 973 participants who had some college or less, 27.9% (271/973) knew that responding to the Census was required by law. Of the 526 participants who had a college degree or more, 37.1% (195/526) knew that responding to the Census was required by law.

a) Identify the explanatory variable. Is it a categorical variable or a quantitative variable?

b) Identify the response variable. Is it a categorical variable or a quantitative variable?

c) Do the data provide any indication of an association between education level and the awareness that responding to the Census is required by law? Explain.
Prayer and blood pressure*

4.1.22 In August 1998, an article titled “Prayer can lower blood pressure” appeared in the *USA Today*. The article was based on the findings of a study by the National Institutes of Health Initiatives that followed 2391 people aged 65 years or more. The article said,

“People who attended a religious service once a week and prayed or studied the Bible once a day were 40% less likely to have high blood pressure than those who don’t go to church every week and prayed and studied the Bible less.”

a) Identify the explanatory variable in this study. Also classify this variable as categorical or quantitative.
b) Identify the response variable in this study. Also classify this variable as categorical or quantitative.
c) Identify a confounding variable that provides an alternative explanation for the lower blood pressure among those who attended religious services regularly and studied the Bible regularly, compared to those who didn’t.

Colds and exercise

4.1.23 In November 2010, an article titled “Frequency of Colds Dramatically Cut with Regular Exercise” appeared in the *Medical News Today*. The article was based on the findings of a study by researchers Nieman et al. (*British Journal of Sports Medicine*, 2010) that followed 1002 people aged 18 – 85 years for 12 weeks, asking them to record their frequency of exercise (5 or more days a week? Yes or No), as well as incidences of upper respiratory tract infections (Cold during last week? Yes or No).

a) Identify the explanatory variable in this study. Also classify this variable as categorical or quantitative.
b) Identify the response variable in this study. Also classify this variable as categorical or quantitative.
c) Identify a confounding variable that provides an alternative explanation for the lower frequency of colds among those who exercised 5 or more days per week, compared to those who were largely sedentary.

Politics and gun ownership*

4.1.24 In February 2013, the Pew Research Center surveyed randomly selected U.S. adults about their opinions on gun ownership and related issues. One of the questions they asked was, “Should states be allowed to ignore federal gun laws?” Of the 366 Republicans, 57.9% (212/366) responded “yes,” compared to 18.1% (85/470) Democrats, and 38.1% (230/604) Independents.

a) Identify the explanatory variable in this study. Also classify this variable as categorical or quantitative.
b) Identify the response variable in this study. Also classify this variable as categorical or quantitative.

c) Do the data provide any indication that there is an association between political party affiliation and opinion of whether states should be allowed to ignore federal gun laws? Explain.

Red meat and heart disease

4.1.25 In March 2012, an article titled “Eating red meat regularly ‘dramatically increases the risk of death from heart disease,’” appeared in the Daily Mail (www.dailymail.co.uk). The article was based on the findings of a study that followed over 120,000 men and women for almost 30 years, the data for which were analyzed by the Harvard School of Public Health in Boston.

a) Identify the explanatory variable in this study. Also classify this variable as categorical or quantitative.

b) Identify the response variable in this study. Also classify this variable as categorical or quantitative.

c) Identify a confounding variable that provides an alternative explanation for the higher risk of death from heart disease among those who ate red meat regularly, compared to those who didn’t.

Chocolate and Nobel prizes*

4.1.26 Researcher Messerli published an article titled “Chocolate Consumption, Cognitive Function, and Nobel Laureates,” in the October 2012 issue of the New England Journal of Medicine. The article shows a positive association between a country’s chocolate consumption and the number of Nobel prizes won by the country (adjusted for population size).

a) Identify the observational units.

b) Identify the explanatory variable and the response variable in this study.

c) Identify a confounding variable that provides an alternative explanation for the higher number of Nobel prizes won by countries with higher chocolate consumption, compared to those who with lower chocolate consumption.

Million Women study

4.1.27 The Million Women study in England followed more than 1,000,000 women aged 50 to 64 years from 1996 to 2001, tracking their living habits, and maintaining records on their medical and social factors as well as cancer data. The researchers found that after around 12 years, 3.7% (217/5,877) of the South Asian women developed breast cancer, compared to 3.6% (180/4,919) of the black women, and 4.4% (45,191/1,038,144) of the white women.

a) Identify the explanatory and response variables in this context. Also, identify whether each variable is categorical or quantitative.

b) Use the conditional proportions to explain whether the data indicate that there is an association between one’s race and whether or not one has breast cancer?
c) Identify a possible confounding variable that provides an alternative explanation for the observed differences in breast cancer rates among the different races.

Section 4.2

4.2.1* Which of the following is the primary purpose of randomly assigning subjects to treatments in an experiment?

A. To produce similar (experimental) groups so any differences in the response variable can be attributed to the explanatory variable
B. To give each subject a 50-50 chance of obtaining a successful outcome
C. To produce a representative sample so results can be generalized to a larger population
D. To simulate what would happen in the long run
E. Both A and C

4.2.2 A randomized experiment allows for the possibility of drawing a cause-and-effect conclusion between ________ and ________.

A. The subjects and the treatments
B. The observational units and the variables
C. The explanatory variable and the response variable
D. Statistical significance and statistical confidence

4.2.3* Which of the following is true of experiments?

A. The researchers assign the explanatory variable to subjects.
B. The researchers assign the response variable to subjects.
C. The researchers assign both the explanatory and response variables to subjects.
D. The researchers assign neither the explanatory nor response variable to subjects.

4.2.4 Which of the following must happen in a study to allow us to determine cause and effect?

A. Taking a random sample from a population.
B. Randomly assign the observational units to different treatment groups.
C. Simulation-based inference techniques.
D. Theory-based inference techniques.
4.2.5* Is random sampling or random assignment the more important consideration if the research question is whether faculty tend to drive older cars than students drive on your campus?

4.2.6 Is random sampling or random assignment the more important consideration if the research question is whether Facebook users tend to have lower grade point averages than students who do not use Facebook?

4.2.7* Is random sampling or random assignment the more important consideration if the research question is whether students tend to receive higher scores on essays if they are encouraged to submit a draft than if they are not so encouraged?

4.2.8 Is random sampling or random assignment the more important consideration if the research question is whether members of one political party tend to donate more to charities than members of another political party?

4.2.9* Is random sampling or random assignment the more important consideration if the research question is whether a waitress generates higher tips by giving her name when she first greets customers?

4.2.10 Reconsider previous exercises (4.2.5 through 4.2.9). Notice that the phrase “tend to” appears in many of the research questions. Explain what this phrase means and why it is important in these questions.

4.2.11* Can a study have both random sampling and random assignment? If so, explain what can be determined from such a study if significance is found.

4.2.12 Researchers could design an experiment where there is a gender balance between two experimental groups by putting half the females in one group and half in the other and do the same for the males. Why don’t researchers always just force variables to be balanced out between groups, but often use random assignment?

4.2.13* Does random assignment always equally balance the variables between two groups? Is there a tendency for there to be a balance? Explain.

4.2.14 From which of the following studies can cause and effect conclusions potentially be drawn? If cause and effect can be determined, explain what may cause what.
Chapter 4: Causation: Can we say what caused the effect?

a) From a random sample of city residents, it was found that those with higher incomes utilize the recycling services significantly more than those with lower incomes.
b) Students were randomly assigned to two groups. One group listened to music while taking a math test and one group did not. The one that did not listen to music scored significantly higher on the math test than the other.
c) A teacher gave one of his classes a math test while music was playing and another without music playing. The class that did not listen to music scored significantly higher on the math test than the other.

4.2.15* From which of the following studies can cause and effect conclusions potentially be drawn? If cause and effect can be determined, explain what may cause what.

a) Subjects were randomly assigned to watch one of two videos, one that was about a sad situation and one that was about a happy situation. The group that watched the sad video scored significantly lower on a quiz about their mood.
b) Students collected data on city residents shopping in their city’s downtown stores and another sample from city residents that were shopping at the mall on the outskirts of the city. They found that those shopping downtown were significantly more likely to vote for the school millage than those shopping at the mall.
c) From a random sample of students, it was found that those who are members of Greek organizations have significantly lower grade point averages.

4.2.16 What is the difference between random sampling and random assignment and what types of conclusions can be drawn from each?

Smoking while pregnant*

4.2.17 Many studies have shown that babies born to women who smoked while pregnant tend to weigh less at birth than babies born to mothers who did not smoke while pregnant.
   a) Are these studies observational studies or experiments?
   b) Will a cause-and-effect conclusion be possible?
   c) Identify the observational units in these studies.
   d) Identify the explanatory variable in these studies.
   e) Identify the response variable in these studies.

4.2.18 Refer to the information provided in the previous question.
   a) While it’s possible in principle to conduct a randomized experiment to investigate this issue, it would be unethical to do so. Explain why, as if to someone who has never studied statistics.
   b) Identify some possible confounding variables in these studies.
Why is your baby a boy?

4.2.19 Many studies have looked at the sex of babies, and then surveyed the mothers and fathers about their behaviors (diet, health, etc.) at or immediately before the time of conception attempting to find connections between these behaviors and the sex of the baby.
   a) Are these studies observational studies or experiments?
   b) Will a cause-and-effect conclusion be possible?
   c) Identify the observational units in these studies.
   d) Identify the explanatory variables in these studies.
   e) Identify the response variable in these studies.

Classical music and intelligence*

4.2.20 Many studies have found that children who listen to classical music while in the womb or when younger tend to have higher IQs later in life. These studies are typically done by surveying parents of older children about whether or not they played classical music when the child was younger, and conducting an IQ test of the child.
   a) Are these studies observational studies or experiments?
   b) Will a cause-and-effect conclusion be possible?
   c) Identify the observational units in these studies.
   d) Identify the explanatory variable in these studies.
   e) Identify the response variable in these studies.

Consumer attitudes

4.2.21 A team of researchers (Singer, et al., 2000) used the Survey of Consumer Attitudes to investigate whether incentives would improve the response rates on telephone surveys. A national sample of 735 households was randomly selected, and all 735 of the households were sent an “advance letter” explaining that the household would be contacted shortly for a telephone survey. However, 368 households were randomly assigned to receive a monetary incentive along with the advance letter, and the other 367 households were assigned to receive only the advance letter. Here are the data on how many households responded to the telephone survey.

<table>
<thead>
<tr>
<th>Received an incentive?</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responded to the telephone survey?</td>
<td>Yes</td>
<td>286</td>
<td>245</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>82</td>
<td>122</td>
</tr>
</tbody>
</table>
4.2.22 Is the sex of a person making a friend request on Facebook to someone they do not know associated with whether or not the request is accepted? To answer this question, student researchers at Hope College made up phony Facebook profiles, one representing a Hope College female student and one representing a Hope College male student. The profiles were made to look as similar as possible except for a couple of pictures of the phony students, their names, and of course their sex. From a group of 219 students from one class at the college, 118 were randomly assigned to receive the friend request from the female “student” and 101 were randomly assigned to receive the friend request from the male “student.” The results of the acceptance of these requests are shown in the following table.

<table>
<thead>
<tr>
<th>Accepted friend request?</th>
<th>Female</th>
<th>Male</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>61</td>
<td>18</td>
<td>79</td>
</tr>
<tr>
<td>No</td>
<td>57</td>
<td>83</td>
<td>140</td>
</tr>
<tr>
<td>Total</td>
<td>118</td>
<td>101</td>
<td>219</td>
</tr>
</tbody>
</table>

Facebook friends*

a) Was this an observational study or an experiment? How are you deciding?
b) What are the observational units?
c) What are the variables recorded? For each variable, identify the type of the variable (categorical or quantitative), and the role of the variable (explanatory or response).
d) Did the study involve random **sampling**? If yes, what is the advantage of a randomly selected sample? If no, what is the disadvantage?
e) Did the study involve random assignment to either get the friend request from either the female or male? If yes, what is the advantage? If no, what is the disadvantage?

f) An appropriate analysis of the data shows there is strong evidence that people receiving the request from the female are more likely to accept. Is it appropriate to conclude that the sex of the requestor is affecting the acceptance rates? Why or why not? Explain your reasoning.

Cuteness and aggression
4.2.23 Are cuteness and aggression related? A study done at Yale University tested this by showing people pictures of cute animals (like kittens and puppies) or pictures of older more serious looking animals. They tested the aggression of the subjects by giving them bubble wrap and let them pop the bubbles. The subjects in the group shown the cute animals popped an average of 120 bubbles compared 100 for the group seeing the pictures of the older animals.

a) Identify the explanatory variable in this study and tell whether it is categorical or quantitative.

b) Identify the response variable in this study and tell whether it is categorical or quantitative.

c) What should the researchers do so they have the possibility of determining that the type of picture affects how many bubbles are popped?

Power poses*
4.2.24 A research article “Power Posing: Brief Nonverbal Displays Affect Neuroendocrine Levels and Risk Tolerance,” published in Psychological Science, September 2010, describes a study involving 42 volunteer participants (male and female), where participants were randomly assigned to hold either low-power poses (contractive positions, closed limbs), or high-power poses (expansive positions, open limbs) for 2 minutes. All participants were told that the aim of this exercise was to see if their heart rate changed. After the exercise, each participant was given $2, and told that they could keep the money, or roll a die for a double or nothing. Of the 20 participants who held low-power poses, 3 took the “double or nothing” bet, whereas of the 22 participants who held high-power poses, 18 took the bet. Suppose we want to know, Are people who hold “high-power poses” more likely to take risks (such as, the double or nothing bet) compared to those hold “low-power poses”?

a) Was this an observational study or an experiment? How are you deciding?

b) What are the observational units?

c) What are the variables recorded? For each variable, identify the type of the variable (categorical or quantititative), and the role of the variable (explanatory or response).

d) Did the study involve random sampling, random assignment, or both? How are you deciding?

e) An appropriate analysis of the data shows that there is strong evidence that people in the high-power are more likely to take the double or nothing bet. Is it appropriate to conclude that the pose affects likelihood of taking the bet? Why or why not? Explain your reasoning.
Skittles taste test

4.2.25 Students recruited at the cafeteria at Hope College were blindly given a single Skittles candy to put in their mouth. They were told the five possible flavors and then were asked which flavor they had. Of the 154 healthy students tested, 78 gave the correct answer. Of the 118 students that had stuffy or runny noses tested, 44 gave the correct answer.

a) Was this an observational study or an experiment? How are you deciding?
b) What are the observational units?
c) What are the variables recorded? For each variable, identify the type of the variable (categorical or quantitative), and the role of the variable (explanatory or response).
d) Did the study involve random sampling, random assignment, both or neither? How are you deciding?
e) An appropriate analysis of the data shows that there is strong evidence that healthy students were more likely to give the correct answer. Is it appropriate to conclude that the health of the student affects ability to taste? Why or why not? Explain your reasoning.

End of Chapter Exercises

4.CE.1* When can you legitimately draw a cause-and-effect conclusion from a randomized experiment?
   A. When the p-value is small
   B. When the p-value is large
   C. Always, regardless of the p-value
   D. Never, regardless of the p-value

4.CE.2 Answer the following questions about explanatory and response variables.

   a) Must an explanatory variable always be categorical?
   b) Must a response variable always be quantitative?

Animal therapy*

4.CE.3 In a recent study of animal-assisted therapy (Cole, et. al., 2007), researchers investigated whether patients hospitalized with heart failure could be helped by a visit from a dog. The 76 patients in the study were randomly assigned to one of three groups: one group received a 12-minute visit from a volunteer and a dog, another group received a 12-minute visit from a volunteer only, and the third group received no visit. Many variables, including heart rate, blood pressure, and anxiety levels, were measured on the patients prior to and after the visit.
a) Identify the explanatory variable(s) in this study.
b) Identify the response variable(s) in this study.
c) Is this an experiment or an observational study? Explain.
d) Describe the purpose of randomly assigning patients to treatment groups.
e) Researchers also reported summary information on a variety of background variables such as age, gender, and smoking status. Do you think the researchers were hoping to find significant differences among the three groups on these variables? Explain why or why not.

Walking in socks

4.CE.4 In a study conducted in New Zealand, researchers Parkin et al. randomly assigned volunteers to either wear socks over their shoes (intervention), or wear usual footwear (control), as they walked downhill on an inclined icy path. Researchers standing at the bottom of the inclined path would then rate how confident the participant appeared (“confident,” “cautious but did not hold onto supports,” or “held onto supports”), and how much time (seconds) the participant took to walk down the path.

a) Was this an experiment or an observational study?
b) Identify the variables recorded. For each variable, identify the type (categorical or quantitative) and role (explanatory or response).
c) Did the study involve randomness? If so, how: random sampling, or random assignment, or both?

Cost of drugs*

4.CE.5 In a randomized, double-blind study reported in the Journal of American Medical Association, researchers Waber et al. (2008) administered a pill to each of the 82 healthy paid volunteers from Boston, Massachusetts, but told half of them that the drug had a regular price of $2.50 per pill, whereas the remaining participants were told that the drug had a discounted price of $0.10 per pill, mentioning no reason for the discount. In the recruitment advertisement, the pill was described as an FDA approved opioid analgesic, but in reality both groups were administered placebo pills. To simulate pain, the participants were administered electric shocks to the wrist, and the researchers recorded the proportion of subjects who reported a reduction in pain after taking the pill.

a) Identify the explanatory variable and response variable. For each variable, identify the type (categorical or quantitative).
b) Explain what “randomized” means in this study.
c) Explain what “double-blind” means in this study.
d) The researchers report “pain reduction was greater for the regular-price pill (p-value < 0.001).” Explain what the researchers mean by this. Be sure to comment, with correct justification, on
   • Whether the results are statistically significant.
   • Whether a cause-and-effect conclusion can be drawn.
   • To whom the results of the study can be generalized.
Rebates and bonuses

4.CE.6 An article in a 2006 issue of *Journal of Behavioral Decision Making* reports on a study involving 47 undergraduate students in a class at Harvard. All of the participants were given $50, but some (chosen at random) were told that this was a “tuition rebate,” while the others were told that this was “bonus income.” After one week, the students were contacted again and asked how much of the $50 they had spent and how much they had saved. Those in the “rebate” group had spent an average of $22.04, while those in the “bonus” group had spent an average of $9.55.

a) Identify the explanatory and response variables in this study. Also classify each variable as categorical or quantitative.
b) Is this an observational study or an experiment? Explain.
c) Did this study make use of random sampling, random assignment, both, or neither?
d) If the difference in average spending amounts between the two groups is determined to be statistically significant, would it be legitimate to draw a cause-and-effect conclusion between what the money was called and how much was spent? Justify your answer.
e) If the difference in average spending amounts between the two groups is determined to be statistically significant, would it be legitimate to generalize the result of this study to all adult Americans? Justify your answer.

Satisfaction with attractiveness*

4.CE.7 A poll conducted by the Gallup organization asked American adults whether or not they are generally satisfied with their physical attractiveness. One goal of the study was to investigate whether men and women differ with regard to responses on this issue.

a) Identify the explanatory and response variables in this study.
b) Is this an observational study or an experiment? Explain briefly.
c) If this study were conducted well, would it have involved random sampling, random assignment, or both? Explain.

Drinking at college

4.CE.8 A study found that college students who live off-campus are significantly more likely to drink alcohol than those who live on-campus.

a) Do you suspect that this was an observational study or an experiment? Explain.
b) Is it appropriate to conclude that living off-campus causes a student to be more likely to drink alcohol? Explain why or why not.
c) Is it appropriate to conclude that being a drinker causes a student to be more likely to live off-campus? Explain why or why not.
How to name your book*

4.CE.9 Ian Ayers, the author of a popular book titled Super Crunchers, conducted a study to help him decide what to name his book. He placed an ad on google.com, with the ad sometimes giving the title as Super Crunchers and other times giving the title as The End of Intuition. Google provided Ayres with data on how often a person clicked on the ad to obtain more information, depending on the title given in the ad. It turned out that viewers were significantly more likely to click through on the Super Crunchers ad than on The End of Intuition ad, so that’s the title that Ayres chose.

a) Was this an observational study or an experiment? Explain how you know.

b) Identify the explanatory and response variables.

c) Was Ayres justified in concluding that the name Super Crunchers caused a higher click-through rate than the other name? Explain why or why not.

Spanking and IQ

4.CE.10 Studies have shown that children in the U.S. who have been spanked have a significantly lower IQ score on average than children who have not been spanked.

a) Is it legitimate to conclude from this study that spanking a child causes a lower IQ score? Explain why or why not.

b) Explain why conducting a randomized experiment to investigate this issue (of whether spanking causes lower IQs) would be possible in principle but ethically objectionable.

Reading Harry Potter*

4.CE.11 You want to investigate whether teenagers in the United Kingdom (UK) tend to have read more Harry Potter books, on average, than teenagers in the United States (US).

a) Identify and classify (as categorical or quantitative) the explanatory and response variable.

b) Would you ideally use random sampling for this study, or random assignment, or both? Explain.

Restaurant customer behavior

4.CE.12 Do restaurant customers tend to order more expensive meals when classical music is playing in the background than when other kinds of music are playing in the background? Describe how you could design a randomized experiment to investigate this question.
Pen color and grades*

4.CE.13 Do college professors who use a red pen to grade student essays tend to assign lower scores, on average, than professors who use a blue pen to grade student essays? Describe how you could design a randomized experiment to investigate this question.

Video games and kids

4.CE.14 Are American adults who have children significantly more likely to play video games than American adults who do not have children? Is it possible to design a randomized experiment to investigate this question? If so, describe how you would design such an experiment. If not, explain why not.

Smoking and church attendance*

4.CE.15 Are people who attend church regularly less likely to smoke than people who do not attend church regularly? Is it possible to design a randomized experiment to investigate this question? If so, describe how you would design such an experiment. If not, explain why not.

Investigation: High Anxiety and Sexual Attraction

Social Psychologists throughout the years have shown that an aggression-sexuality link exists in not only various animal species, but also in humans. Dutton and Aron (Journal of Personality and Social Psychology, 1974) set out to show that a more general link exists in humans, namely emotional arousal of all kinds and sexual attraction. They set up their study to compare men in a high emotional arousal situation to men in a low emotional arousal situation.

Researchers wanted to test the notion that an attractive female is seen as more attractive by males who encounter her in a fear-arousing situation than by males who encounter her in a non fear-arousing situation. In the high emotional arousal group, men deemed to be between the ages of 18 and 35 who crossed a suspension bridge 230 feet above rocks and shallow rapids in Capilano Canyon, North Vancouver, British Columbia, Canada and were not accompanied by a female were approached on the bridge by an attractive female interviewer. The same interviewer also approached men who fit the same criteria but crossed a solid wood bridge 10 feet above a rivulet that ran into the main river. Both groups of men were interviewed on the bridge and were told that the interview was for a Psychology class project on the effects of exposure to scenic attractions on creative expression. The men filled out a short questionnaire after which the interviewer wrote her phone number down on a slip of paper and said that if...
they were interested in the results of the experiment they could call her. The researcher talked to 18 men on the suspension bridge and 16 men on the wooden bridge.

**Step 1: Ask a research question**

The researcher is now interested in seeing whether a higher proportion of men in the group that crossed the Capilano Canyon bridge will call her versus the proportion of calls she will receive from the group that crossed the solid wooden bridge.

**Step 2: Design a study and collect data**

1. Is this an observational study or an experiment? How are you deciding?

2. What are the observational units?

3. We have two variables: bridge type and whether the subject called or not. Which of these variables is explanatory and which is response?

You will learn in Chapter 5 how to state null and alternative hypotheses for studies like this. Here they are for this study:

Null: The long-run probability of calling the researcher from men in the study is the same in both groups

Alt: The long-run probability calling the researcher from men in the study is larger in the group of men crossing the Capilano suspension bridge

Of the 18 men on the suspension bridge that accepted her phone number, 9 called her. Of the 16 men on the solid wooden bridge that accepted her phone number, 2 called her.

**Step 3: Explore the data**

4. Fill in the two-way table with the data from the study.

<table>
<thead>
<tr>
<th></th>
<th>Suspension bridge</th>
<th>Wooden bridge</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject called interviewer</td>
<td></td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>Subject did not call interviewer</td>
<td></td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>Totals</td>
<td>18</td>
<td>16</td>
<td>34</td>
</tr>
</tbody>
</table>
5. What proportion of men on the suspension bridge called the interviewer?

6. What proportion of men on the wooden bridge called the interviewer?

Step 4: Draw inferences

In Chapter 5, we will learn how to use the 3S strategy to help us investigate how much evidence the sample data provide to support our conjecture that the long-run probability of men who call is greater for those on the suspension bridge than for those on the wooden bridge. As a result of the simulation analysis method used in Chapter 5 we received a p-value of 0.023.

7. Use the null and alternative hypotheses stated on the previous page and the p-value of 0.023 to write a conclusion about strength of evidence in this study. *Hint: Even though the study design is different, the way you interpret p-values in terms of the null and alternative hypothesis is still the same.*

8. How would the p-value have changed if the alternative hypothesis was stated as: *Alt: The long-run probability of calling the researcher from men in the study is different between the two groups*

9. Write a conclusion about the strength of evidence in the study based on the p-value from #8.

10. In studies comparing two groups, the parameter of interest is often stated as the true difference in the long run probabilities in the two groups. In this study, this is the long-run difference in the probabilities of men calling the researcher comparing the men crossing the Capilano suspension bridge to the men crossing the wooden bridge. A 95% CI for this parameter is $0.375 \pm 0.35$. Explain how the fact that the confidence interval does not include 0 corresponds to the size of the p-value obtained in #9.
Step 5: Formulate conclusions

11. Generalization. Now, let’s step back a bit and think about the scope of our inference. Was this a random sample? Does this study represent all people? How about all men? Is there any population to which we can generalize our conclusion?

12. Causation. What about cause and effect conclusions? Did the researchers use random assignment to determine which of the 34 men crossed the suspension bridge and which crossed the wooden bridge? Are we able to conclude that the cause of the difference between the two groups of men was the type of bridge they crossed?

Step 6: Look back and ahead

The researchers acknowledged the fact that there were confounding variables involved in this study. They tried to minimize these by doing other studies to try to answer the same research question. In another study, they again interviewed subjects on the suspension bridge just as before. As a control, they interviewed subjects that had already crossed the suspension bridge at least 10 minutes earlier and were sitting or walking in the nearby park.

13. Describe the confounding variable they are trying to eliminate by completing this new study.

Let’s see how this second study turned out. Of those on the suspension bridge, 13 of the 20 men called the female interviewer. Of those that had already crossed the bridge, 7 out of 23 men called the female interviewer.

14. In this study, what proportion of those on the suspension bridge called the interviewer? Of those who had already crossed the bridge?
The hypotheses are modified slightly for this new study.

Null: The long-run probability of calling the researcher from men in the study is the same in both groups

Alt: The long-run probability of calling the researcher from men in the study is larger in the group of men who are on the Capilano suspension bridge than the in the group of men who recently crossed.

15. Write a conclusion about the strength of evidence in this new study based on obtaining a p-value of 0.024.

16. Generalization. Have the researchers addressed the concerns about generalization with this new study?

17. Causation. Is a cause-effect conclusion possible in this new study? Why or why not?

Research Article: Impacting People’s Willingness to Pay

Read “The Effect of Red Background Color on Willingness-to-pay: The Moderating Role of Selling Mechanism” and then answer the following questions.

Step 1. Ask a research question

1. In no more than 1-2 sentences state a broad research question (or related questions) that the researchers were trying to investigate in this study.
2. Identify at least two (cited) points of evidence the researchers cite as to why the research question(s) are important. Include the citation in your response.

The following questions (3-12) refer to Study #1.

Step 2. Design a study and collect data

3. In the first part of study 1, the researchers sold 28, identical, Nintendo Wii bundles on eBay. There were over 900 bid jumps measured across the 28 auctions, in your own words explain what a bid jump is.
Step 3. Explore the data

4. Explain how the mean (20.10) and median (10.00) of the bid jumps data across the 28 auctions confirm the researchers' conclusion that the data was right-skewed.

Step 4. Draw inferences

5. The average bid jump for Wii's sold on red backgrounds was $20.82 and was $19.22 on blue backgrounds, for a two-sided p-value of < 0.02. Complete the following sentence “We have ______________________ that the average bid jump is different (higher) for Wii’s sold on a red background compared to blue.

Step 5. Formulate conclusions

6. In order to be a randomized experiment observational units (bid jumps) must be randomly assigned to the treatment groups (red or blue background). Is this study a randomized experiment? What are the implications on the conclusions from this study?

7. Why was it important to keep seller and product budget identical across auctions?

8. The researchers state that “Although the results support our expectations, we acknowledge that the red and the blue auctions may have differed on other factors (for instance, the number of individuals who viewed the listing or self-selection).” What is the term for these factors?


9. The researchers followed up by conducting a randomized experiment on 78 undergraduate students. Provide a one sentence summary of the strength of evidence in this study based on the p-value (p < 0.01), comparing the average bid jump for red ($63.17) to the average bid jump for blue ($35.13).

Step 5. Formulate conclusions

10. Because this study involved random assignment what should be true about how similar the 39 students are in each of treatment groups (red and blue)?

11. Because of the random assignment, what type of conclusion is possible?

12. How were the participants for this study selected? What are the implications on the conclusions that can be drawn from this study?

13. In study #2, the researchers randomly assigned 89 individuals to make offers on a hotel where the pictures are placed on a blue or red background. Make a statement about the strength of evidence for a difference in the average bid price between the red and blue groups.

**Step 5. Formulate conclusions**

14. The researchers state that they measured how much participants liked the vacation (1 to 7 scale). How should the number of people who like the vacation a lot (6 or 7) compare between the two treatment groups (red and blue backgrounds)? Why?

15. Explain the impact of random assignment on the conclusions that can be drawn from study #2.

16. How were the participants for this experiment selected? What are the implications on conclusions you can draw from this study?

**Step 6. Look back and ahead**

17. Identify at least two things that, if you were running the study and had a reasonable amount of financial resources, you would do differently. You must justify your answers.