LEARNING OUTCOMES

After reading this chapter, you should be able to:

• Provide operational definitions for key variables.
• Identify key design issues with two-group research.
• Identify factors involved in establishing causation.
• Explain how control is established in an experiment.
• Discuss the importance of group independence and random assignment.
• Write a results section for an independent samples t-test.

Something to Think About . . .

You have probably found yourself sitting in class dutifully trying to pay attention, but distracted by a window, people passing by the door, someone tapping a pen, a classmate with a cold, or some idiosyncrasy of the professor. (Really, how many times can a person say “alrighty then” in one class?) While these are surely all distractions, many people view technology as the biggest distraction of all. At this point, it is common knowledge that texting while driving is dangerous. If your syllabi are any indication, it seems professors are equally concerned about how technology can distract your learning. Whether you are checking Facebook on your laptop, tweeting from your tablet, texting from your smartphone, or doing all three, your professor probably does not condone this behavior. If you think about it, you can understand your professor’s perspective. These forms of technology can lead you to miss course material, or, worse, distract other students, leading them to miss things. Then again, it is all too easy to blame technology for an age-old problem. You play with your smartphone; your parents passed notes. Technology-based classroom distractions are not a “new” problem, but, rather, the latest version of a long-standing issue. Besides, before we worry too much about smartphones’ potential classroom harm, we need to see some proof that smartphone use is actually distracting.

In order to combat the perceived evils of technology, many professors craft highly restrictive course policies that treat using portable technology as a form of high treason. However, is it possible that by invoking an absolute ban, these
professors are creating more of a distraction? We will answer that question with another question: Have you ever been hungry and tried to avoid thinking about food? As soon as you tell yourself not to think about being hungry, what do you feel? Hungry. Clearly, this method does not work very well. By trying to avoid thinking about food, you tend to think about that exact thing even more. Are you still not a believer? Try this: Right now, no matter what, do not under any circumstance think of Donald Trump naked.

Yikes! See that? You thought about exactly what we asked you not to think about. Sorry about that! Next, we’d like you not to think about a cute puppy playing with a ball. That image should make it all better. This little exercise helps us demonstrate a point: Not thinking about something is really hard. If we are truly “addicted” to our technology, or just really attached to it, a rule that forbids us from using our devices essentially asks us not to think about them. Ironically, this could create a major distraction. It would be great if someone could design a study to test if well-intentioned, antitechnology professors are actually inhibiting your learning. That someone could be you.

Introduction to Our Research Question

Before we design a study to test our idea, we should state our key question a bit more clearly. Although smartphones allow us to do many different things, we will focus on texting because that seems to be the number one offender on professors’ lists. We want to answer the general question:

Can trying to ignore a text message lead to distraction?

Compared to our research questions from previous chapters, this question seems to make a more specific prediction about a type of behavior that leads to a specific outcome. Our focus has moved beyond merely describing what is taking place, to identifying how one factor (ignoring text messages) can lead to changes in another factor (distraction). To answer this type of question, we will need to use experimental research to determine cause and effect. But let’s not get too far ahead of ourselves just yet. If professors have antitexting classroom policies, there must be research on the detrimental effects of using technology. Let’s see some evidence.

Reviewing the Literature: What Do We Already Know About This Topic?

If we are going to conduct a well-designed study, we should familiarize ourselves with the existing research. It would be useful to know the answers to a few key questions: How prevalent is student texting in class? Does texting have any detrimental effects? What is the evidence that texting is distracting while driving? What about for learning or thinking, in general? Are there any benefits to texting? Has anyone looked at the consequences of preventing texting? Based on a literature search and reading through several articles, we have taken notes and written our own summaries of a few relevant studies that will help us understand the research in this area.
Prevalence of Texting


This study examined 269 college students to determine the extent to which students used their cellular devices in class. An overwhelming majority (95%) of students brought their cellular devices to lectures on a daily basis, though 91% said they had it set on “vibrate.” Of these students, most (92%) used their cellular devices in order to send SMS (short message service) messages during lectures at some point, 97% said they were aware of other students doing it, and a small number (10%) reported that they used their cellular device during a test. A majority of students (61.6%) agreed with the statement, “Yes, I see no problem with using a cell phone to text in class as long as I am not disturbing other students,” while some (31.5%) believed that sending texts negatively influenced students “through loss of attention and/or poor grades.”

The Influence of Texting on Driving


Researchers studied data on driving fatalities in the United States from 1999 through 2008 to determine how they related to texting and cell phone use. Their analysis indicated that distracted driving fatalities decreased from 1999 to 2005, but rose (up 28%) from 2005 to 2008. The researchers’ analyses led them to conclude that texting was responsible for an additional 16,000 fatalities over a 6-year period from 2001 to 2007.

The Negative Impact of Technology on Learning and Cognition


To determine whether texting has an effect on memory recollection of classroom material, researchers assigned 185 undergraduate students to one of three groups: no texting, medium texting (8–15 texts), and excessive texting (16 texts or more). Researchers manipulated texting by sending messages to participants while participants watched a 30-minute video of a class, and then gave participants a memory test based on the video. The excessive-texting group had the worst test scores of the three experimental groups. Those who typed more in their text messages and received longer messages also did worse than those sending/receiving shorter messages.


This study tested if distractions from a cellular device and texting could inhibit college students’ ability on a cognitive identification task (i.e., memory of words). Researchers randomly assigned 64 college students to one of three groups: no distraction, distraction (answered their phones if they rang), or different distraction (answered their phones using a text message). Groups reviewed two sets of 24 related words (e.g., diaper, crib), then took an identification exam to test their recall of the words. Results indicated that
participants in the no-distraction group performed best, while both distraction groups had worse performance.

**Potential Benefits of Texting**


It might be possible that text messages can promote students’ learning in educational settings through texting with tutors. To test this, researchers asked 1,121 students questions about texting behavior, then transcribed their responses. Students viewed text messages from tutors as intrusive, mainly because they viewed cell phones as being used for primarily personal reasons (i.e., texting friends and family). Students did acknowledge that texting kept them in constant contact with the tutor, which they saw as beneficial if it was consistent with students’ perceived needs.


To determine if texting and the use of textese (e.g., “LOL” or “IMO”) hindered literacy skills, researchers had 152 undergraduate students take literacy (reading and spelling) tests, then answer questions about texting behavior and use of textese. Students reported frequent use of texting and use of textese, though use was primarily in the context of texts or e-mails to friends, not professors. Results of the analysis showed that, contrary to previous studies, more self-reported texting was associated with greater literacy. However, participants who reported using more textese in e-mails to professors or on social networking sites had poorer reading accuracy scores.

**The Consequences of Preventing Texting**


This study examined texting behaviors, including the restriction of texting, among 23 psychology students (age 18–23). Results found that texting was the most frequent way students kept in touch. Both frequent and infrequent texters reported difficulty with limiting their texting and higher levels of anxiety when cell use was restricted. The researchers suggested that texting is a norm for this age group, and having texting limited throughout the day may hinder social relationships.

**From Ideas to Innovation**

We have to hand it to the antitexting professors of the world: Although it may be tough to admit, they seem to be right about texting’s negative consequences. However, just because texting on the road can make you a worse driver (Benedetto, Calvi, & D’Amico, 2012) and texting in class can lower test scores (Rosen, Lim, Carrier, & Cheever, 2011) does not mean that a total classroom texting ban is best. Remember that researchers have a number of strategies for coming up with interesting research ideas and hypotheses, one of which is looking for the exception to the rule. Texting has negative consequences, but a complete ban on texting in classrooms may introduce
problems as well. In spite of the downsides to texting, there is some research showing that texting may lead to greater literacy (Drouin, 2011), that it can have benefits in an academic context (Brett, 2011), and that restricting texting may provoke anxiety in some students (Skierkowski & Wood, 2012). Our literature search did not reveal any studies that directly examined how trying to ignore text messages may create a distraction, suggesting that our research question may be onto something new. The next time your professor gets annoyed by in-class texting, it would be nice to be able to provide scientific evidence that completely banning smartphone use can have negative consequences, too.

Research Spotlight

Are Smartphones Good for Friendships?

Smartphones have revolutionized the ways we stay in touch with other people. But have you ever wondered how they can affect our relationships, even when we are not using them? Recent research suggests that the mere presence of a cell phone can undermine relationships. In two studies, researchers asked randomly assigned pairs of strangers to have a conversation in a room with either a nondescript cell phone or an old-fashioned pocket notebook on a nearby desk (Przybylski & Weinstein, 2013). The presence of the cell phone (vs. the notebook) led participants to report less closeness with the partner and lower relationship quality following their conversation. Even when the researcher encouraged participants to talk about a meaningful topic, having a phone nearby still undermined the relationship by evoking less empathy and trust between the partners.

Defining Key Terms: What Do You Mean By ____?

Before we can develop a study, we need to define the key parts of our research question. Specifically, we should determine what we mean by “trying to ignore” and “distraction.” While at least one previous study looked at not using one’s cell phone (Skierkowski & Wood, 2012), that may not be the same as ignoring it. We all know the concept of ignoring something, but we need to determine how psychologists define “trying to ignore something.”

One way to ignore something is to engage in selective attention, where you purposefully focus on important information while screening out distractions (e.g., Wolfe, 2014). Selective attention is responsible for what psychologists call the “cocktail party effect,” or the ability to focus on one conversation in a noisy environment while ignoring everything else taking place (Bronkhorst, 2000). Selective attention requires something to focus on and some ongoing distraction to ignore. In a classroom setting, selective attention would involve your ability to pay attention to a lecture amid ongoing distractions. Receiving a text message does not really qualify as an ongoing distraction. Rather, it is an intrusion that enters your awareness when you hear a “ping” or feel a vibration. These experiences are short-lived, but they may trigger a cascade of thoughts about the text message. In this
context, to effectively ignore the text requires thought suppression, the process of deliberately trying to stop thinking about certain thoughts (Najmi, 2013). At the beginning of the chapter, when we asked you to avoid thinking about President Trump naked, we activated a thought-suppression scenario.

This conceptualization fits in with our original idea about the distraction of ignoring texts. When we try to ignore a text message, we must suppress a variety of thoughts surrounding the text, such as who sent the text, if it is about an important matter that needs a response, or if the sender will be angry if we do not respond. And there is more to ignoring a text than simply not thinking about it. We have to avoid acting on it, too. Psychologists call this impulse control, a process that involves any attempt to suppress a desired but inappropriate behavior (Baumeister, Bratslavsky, Muraven, & Tice, 1998). When you receive a text in class or while driving, you need to suppress the urge to pick up the phone and read the message.

Based on what we were able to dig up from the psychological literature, we can combine existing conceptual definitions of thought suppression and impulse control to create the definition for “trying to ignore.”

We also need to determine a conceptual definition for “distraction.” Obviously, we are not the first to explore the topic of distraction, so again the literature can help us clarify what we mean in terms that concur with other psychologists. Distraction occurs when something captures our attention, drawing us away from the focal task or any task that requires our full attention (Parmentier, Elsley, & Ljungberg, 2010). Typically, we are trying to focus on one focal event or object (e.g., a class lecture), when another event or object (e.g., a text, phone call, etc.) grabs our attention and then requires us to divide our attention between the two. We could get distracted either because the focal event is not sufficiently interesting, or because the distractor is especially intriguing. This explanation of distraction fits perfectly with the notion that ignoring a text may be a distraction. After all, the sound or buzz of an incoming text can pull your attention from focal tasks, like listening to a dull professor, because the distracting text is potentially more appealing.

**Weighing Our Options: Picking a Design**

Next, we need to choose our study’s design. We should let our research question guide our design choice, so let’s refer back to our question: “Can trying to ignore a text lead to distraction?” From Chapter 2, we know that there are two main design types: nonexperimental designs, which focus on determining what happens, and experimental designs, which focus on determining why something happens. In our case, the research question makes a prediction about how one factor causes the other, rather than merely trying to describe what
takes place. As a result, the nonexperimental research designs (e.g., interviews, observations, and surveys) are not appropriate ways to answer this particular question. (If we wanted to know whether or not people believe they get distracted by texts, or if we wanted to watch them in class to see the extent to which a text might distract, those designs would work.) Our question requires an experimental design.

Benefits of Experimental Designs
A major benefit of experimental designs is their ability to identify cause-and-effect relations between variables. That is, we can infer that one variable actually creates or enacts a change in another variable. For an experiment to establish causation, it must meet three specific criteria.

Let’s assume that you want to determine whether different body postures or poses can change levels of hormones, such as testosterone or cortisol (Carney, Cuddy, & Yap, 2010). To establish causation, you must first establish that body posture relates to hormone levels, or that there is covariation between these two variables. For covariation to exist, two variables must vary or change together in a systematic way. If participants’ body poses varied, but you found no difference in hormonal levels, then you would not have covariation. If hormonal levels systematically varied with changes in participants’ body positions (e.g., participants’ testosterone levels were higher when standing in a dominant position and lower when standing in a submissive position), then you would have covariation. A correlational study cannot establish causation, but it can establish covariation.

Covariation tells us that two variables relate to each other, but, unfortunately, does not tell us the direction of this relation. That is why we have the mantra that correlation does not equal causality. While it is possible that people have higher levels of testosterone after standing in a dominant position, it is also possible that having higher testosterone levels leads people to stand in a dominant position. If you believe that body posturing causes a change in hormone levels, you have to establish that body posture changes occur before hormone level changes. That is, you must demonstrate temporal precedence by showing that changes in the suspected cause occur before the changes in the effect or outcome. Correlational designs often cannot establish temporal precedence because both variables occur at the same time. However, it is easy to establish temporal precedence in an experimental design because you dictate the study’s order and can manipulate your independent variable (body posturing) before you measure the dependent variable (hormone level). Without temporal precedence, you cannot determine which variable is the cause and which is the effect.

Research shows that parental support and grades covary (Hamilton, 2013). When parents give more financial support to their college-enrolled children, students’ grades are lower, while students who receive less help have higher grades. Although there is covariation between parental support and grades, we cannot say that receiving more financial help from parents causes lower grades. (fstop123/Stock/Getty Images)
The final step of establishing causality is showing that covariation between variables is only due to the independent variable and not due to an extraneous variable, which is any factor separate from the independent variable that could account for variations in the dependent variable. To determine causality, you need to eliminate or control any extraneous variables that could serve as an alternative explanation for the observed changes in the dependent measure. The extent to which you can do this has implications for the internal validity of your experiment. Internal validity refers to the degree to which you can rule out other possible or alternate causal explanations for an association between the independent and dependent variables in your experiment. To show that body posturing causes changes in hormone levels, you must demonstrate that everything except for body posture is the same throughout your experiment.

Ensuring that our text-messaging study has good internal validity may be challenging. The fact is that some things may vary in a study even when we identify and subsequently attempt to control them. Fortunately, a strong experimental design and procedures can eliminate most of the problems with extraneous variables and help us establish a cause-and-effect relationship.

Operationally Defining the IV: Manipulating Ignoring Text Messages

As we design our text-messaging experiment to establish cause and effect, we should figure out which of our key concepts, “trying to ignore” or “distraction,” causes the other. Our research question implies that we want to see how ignoring a text message impacts our ability to avoid distraction. Therefore, trying to ignore the text is our independent variable (IV), as we want to examine if ignoring texts causes or alters the outcome. The result or outcome of ignoring the text is our dependent variable (DV), or, in this case, the amount of distraction. Based on this, we will eventually want to manipulate the independent variable and measure the dependent variable. Now that we have clearly defined both of our key concepts (“trying to ignore” and “distraction”) and identified our independent and dependent variables, we are ready to operationalize each variable in our study.

When you set out to define variables in an experiment, you can operationally define either the independent or dependent variable first. As the researcher, it is your judgment call based on what you learned from your literature search or how you approach the research question (e.g., which variable feels like the most natural place to start for you?). In our case, our research question emanates from our interest in the potential consequences of professors forcing students to ignore text messages. These consequences are a good place to begin. Based on the conceptual definition we established earlier, ignoring should involve deliberately trying to suppress thoughts about texting.

Because ignoring texts is our independent variable, we will need to devise a way of manipulating the “ignoring” part of our experiment. We ignore smartphones
all the time, whether we are ignoring others’ use of their phones when they are
texting or when we hear others’ phones signaling an incoming call or text. While
these are potential distractions, trying to ignore our own phone can be more
distracting. Whenever our phone makes a noise, we cannot help but wonder who
is trying to contact us and why. However, we also know that there are some sit-
uations in which it is not permissible to check our phone, such as while driving
or sitting in a classroom. In fact, we know that many professors are not pleased
when they see students texting during class and that some have course policies
that strictly forbid phone usage during lectures.

The easiest way to manipulate our independent variable may be to mimic the
types of policies professors use to restrict phone use in their courses. That is, we
can create a “forbidden phone checker” group in which we force participants
to ignore their phones by strictly forbidding them to check their phones. Mim-
icking actual course policies with our manipulation has the benefit of increasing
our study’s mundane realism, or how closely our study parallels the real world.
The goal will be to make the participants’ task in our experiment as similar to a
classroom experience as possible.

Experiments are essentially about comparing sets of participants to deter-
mine how these groups may differ. For our study, we need to identify another
group to which we can compare our “forbidden phone checker” group. Given
that we are one of the first studies (or perhaps the first) to examine this
research question, we should keep things simple. Although it may be tempting
to design a study that tests many factors at once, it is often better to start with
a more focused study. That way we can see if ignoring one’s phone has any
effect on learning before we devote additional time and energy (our own and
our participants) to studying how other factors may relate to this research
question. We will use a two-group design, or simple experiment, which is an experi-
mental design that compares two groups or conditions and is the most basic way to
establish cause and effect.

Weighing Our Options: Identifying the
Best Groups for Our Study

Next, we need to determine the best two
groups to use for our simple experiment.
We refer to these groups as the experimental
and control groups. The experimental
group is whichever group gets the key treat-
ment, and the control group is the compar-
sion group that gets less of the treatment.
When creating groups, we can start with
the most obvious comparison, which is an
all-or-nothing comparison. Think of this as a compar-
ison where in one group a light switch is on (all) and
in the other group the light switch is off (nothing).
In the context of our study, we could have a high-restriction group, which is strictly forbidden to use or check their phones in any way, while the no-restriction group has free and complete use of their phones.

However, this all-or-nothing approach can lead to several differences between the groups. For example, compared to those in the high-restriction group, participants in the no-restriction group have full use of their phones and may keep them in view, respond to text messages or e-mails, check Twitter, use Snapchat, book a trip for spring break... you get the idea. In fact, there are so many differences between the two groups that if we were to compare groups and find that the no-restriction group was more distracted, we would not know why. Was it because their phones were visible and usable, because they sent more messages, or because they received more messages? With these two groups, it is impossible to know, which makes it difficult to have any confidence in a conclusion. Ultimately, the quality of the manipulations we use to create groups plays a major role in determining the quality of our study.

To create quality groups, we want to be sure we have a high degree of experimental control, or the ability to keep everything between groups the same except for the one element you want to test. This element, sometimes called the treatment, is the factor you think makes a difference in the outcome variable. This approach makes intuitive sense. For example, if you want to determine if 8 hours of sleep helps students’ alertness the next morning, you would tell one group to go home and sleep 8 hours that night, and tell the other group to sleep 5 hours. Although the groups certainly seem similar, with 3 hours of sleep being the only difference, assuring experimental control is harder than it seems. If you allowed participants to sleep at home, there could be a near endless number of differences in the study that could affect alertness, besides number of hours slept. For example, the participants’ beds and pillows might be different, some might have roommates while others do not, and some might sleep in a warm room while others prefer cold, not to mention that what the participants do before going to bed would vary considerably. For instance, even if everyone watches TV, what they watch might be different and might influence sleep and next-day alertness. The most common solution to these types of issues is to bring participants into a laboratory where experimenters can control many factors to help keep participants’ experiences similar.

Rather than strictly forbidding participants to look at their phones, we could allow participants to have their phones out, face down, and in “silent” mode, where they would not make any noise or vibration. Participants would not be allowed to use their phones or pick them up to check them. We could keep the same “no restriction” control group. Although this scenario is better because we have now controlled the presence of the phones across groups, the groups are still different in a way that makes it difficult for us to draw conclusions about distraction. Compare the two groups:

Group A: Not allowed to know about incoming messages; Not allowed to use the phone

Group B: Allowed to know about incoming messages; Allowed to use the phone
In this setup, we have two key differences between the groups: knowing about incoming messages and being allowed to use the phone to respond to the messages. However, our research question focuses more on the implications of participants ignoring their phones and less about participants’ use of their phones. Consequently, these two groups are too dissimilar. As we said, the goal of experimental control is to have only one element vary between groups, and that element should be whatever the researcher considers the “key ingredient,” the factor the researcher wants to test.
We see now that the differences between groups in an all-or-nothing approach can be drastic. To have more control in our study, we can use the *a-little-more-versus-a-little-less* approach, in which the experimental group gets a little bit more of the treatment (in our case, smartphone restrictions), while the control group gets a little less. Before we set up these groups, we should consider whether our groups will vary by the “key ingredient” according to our research question. We want to know whether trying to ignore text messages creates a distraction in the classroom. Thus, the key factor to manipulate is the degree to which participants have to ignore text messages on their smartphones. We will require one group to ignore their phones by forbidding them to check their phones when they hear a text, while the other group can check their phones to see who the text is from and what the message says. Ultimately, we will operationally define our groups to have the following differences:

**High-Restriction “Forbidden Phone” Group:** No Phone Checking, No Phone Use  
**Low-Restriction “Phone Checker” Group:** Phone Checking Allowed, No Phone Use

Our groups will vary in terms of phone checking, but we must make sure these groups are otherwise similar. We already decided that both groups of participants should have the phones out of sight and set on silent. Students use their phones for other potentially distracting things, like e-mail, games, calls, and checking social media. Although these are potential distractions, they are not our central interest at this point. Yet, we cannot ignore them. We will keep these aspects consistent in both groups by telling everyone those uses are not allowed and that the only notification they should enable on their phones is the vibration for texts.

Ultimately, there should only be one key difference: phone checking forbidden versus allowed. In looking at our groups, the experimental group should be the one that we think will be more distracted. Our exploration of this topic started with the premise that overly restrictive classroom phone policies are actually more distracting than less restrictive policies. Based on this, the “forbidden phone” group is more restrictive, and since we think being overly restrictive is the key factor in distraction, that makes it the experimental group, while the less restrictive “phone checker” group is our control group.

**Operationally Defining the DV: Measuring Distraction**

We think that overly restrictive phone policies can create distraction, but we still need to determine how we will know whether a participant becomes distracted. That is, we need to find a way to capture or measure our dependent variable. It sounds simple, but distraction is a mental process that we cannot see directly. Because the participant is the only person privy to that information, we could ask our participants to self-report their levels of distraction. Although there are many times when a self-report is appropriate, this is not one of them. Due to social desirability concerns, participants may be reluctant to admit to succumbing to distraction.

What Do You Think?

What other groups could you compare to test the influence of smartphone restrictions?
It is also possible that part of experiencing distraction is not realizing you are distracted. Even if we could ensure that participants were completely forthcoming and accurate about their levels of distraction, asking them to reflect on their distraction might lead them to wonder why we are asking about distraction, all of which might ultimately make them less distracted. Given the limitations of self-report for this particular dependent variable, we need to consider alternatives. As we discussed in Chapter 4, mental processes can have behavioral manifestations. Feeling sad coincides with certain facial expressions and, on occasion, crying. Similarly, feeling distracted will manifest in certain behaviors, so it is possible that a behavioral measure is more appropriate for our study.

**Research Spotlight**

**Driven to Distraction**

Driving is full of distractions. There are pedestrians, other drivers, music, and making sure we finish our drive-thru fries before they get cold. Then there are cell phones. Nearly every American driver owns one and can be distracted from the roadway. While some states have yet to pass any laws about cell phone use, others have restricted cell phone use to situations that are “hands free.” These laws imply that the danger of cell phone use while driving is more about not having your hands on the wheel and less about the distraction of talking. Does hands-free cell phone use eliminate the danger of using your phone while driving? To find out, researchers had 30 participants use a driving simulator (Benedetto et al., 2012). During the simulated driving experience, participants used a cell phone while encountering a critical stopping decision. Participants repeated each road scenario four times without calls and while answering a call in one of the following ways: hands-free voice activated device, hands-free cell phone, or handheld cell phone. The conclusion was that using a cell phone of any kind slowed drivers’ reaction times and decreased driving performance equally across all cell phone modes. States that encourage the use of hands-free phones may not be making their roads any safer than those without cell phone restrictions.

**Weighing Our Options: Identifying Key Behaviors**

If we are going to examine behaviors that indicate distraction, we need to be relatively sure that the behavior actually indicates distraction and not some other mental process. Otherwise we run the risk of inferring too much based on the observed behavior. For example, if a participant plays with her hair, is that an indicator of distraction? It could just as easily indicate nervousness or boredom. The easiest way to avoid inferring too much is to use previous research to identify behaviors that other researchers have already used as indicators of distraction. By mirroring previous research as much as possible, we will have more confidence in our measurement.

Since our goal is to test the idea that bans on texting can be as distracting as texting, showing similar effects on measures used in texting studies would be convincing. Although previous research has not focused specifically on distraction
resulting from restrictive phone policies, texting research has examined behaviors that texting influences. As we learned from our literature review, texting can influence driving ability (Benedetto et al., 2012), memory of a video (Rosen et al., 2011), and memory for a word list (Smith et al., 2011). Of these, the studies on memory are most similar to the type of distraction students might experience in class.

Memory tests relate to learning in a classroom, but we may be able to improve on them. We could, for example, measure distraction by observing students’ natural reactions in class. However, it would be hard to know exactly what we observe. Some people might stare or appear to “zone out” when distracted, while others might do the exact same thing when they are thinking deeply about a difficult problem. Our measure of distraction should be more direct. Given that other research used tests to measure distraction, we could use students’ test performance as our measure, under the assumption that greater distraction will lead to poorer performance.

**Focus on Ethics: Should We Really Do That?**

There are both practical problems and ethical issues to consider with this proposed measure of distraction. If our study undermines students’ learning or their actual quiz or test performance in the class, there are ethical implications. In fact, our literature search revealed that cell phone use while studying can undermine performance (Rosen et al., 2011; Smith et al., 2011). As a result, we should not expose our participants to a potentially harmful situation. In a real classroom setting, poor test performance could hurt students’ overall course grade, which could then hurt their GPA, which could then influence their acceptance into a desired graduate school, and so on. While this may sound extreme, we need to consider the ethical implications of our design decisions when conducting research in a real-life setting.

Even if we are able to effectively resolve these ethical issues, there are too many factors in a natural classroom environment that are out of our control. These include things like the day’s lecture topic, where the student sits, other students’ behaviors that day, and so on. Any of these could lead to more or less distraction in a way that would make it hard to tell what type of influence a phone policy would have. Because these factors are not part of our intended manipulation, we will control them by keeping them constant across our two groups. To do that, we will want to measure distraction in a way that parallels how it occurs in a classroom setting, increasing our study’s mundane realism. We will also want to conduct our study in a laboratory setting where we can minimize outside influences. Thus, to assess participants’ distraction, we can measure test performance on material presented via a video lecture and a reading in a controlled situation. This exercise addresses the ethical issues we identified, closely parallels a classroom setting, and borrows methods from previous research.

If our dependent variable is test performance, we want to be sure that we have a good way to measure it. Our test should be long enough that we can collect a variety of scores. That is, if we only have 5 questions, the top and bottom scores will only differ by 5 points, but if we have 40 questions (20 for the video and 20 for the reading), scores will be more varied. This feature makes our measure more sensitive, which means we will be more likely to see potential differences between the two groups. We can also help ensure variation in our scores by making our
questions moderately difficult. If the questions are too easy, we may have a problem with a ceiling effect, in which most participants get high scores regardless of whether or not the phones were distracting. Finally, we need to determine the quiz format. Should we use multiple-choice or short-answer questions? Although the latter may provide more information, multiple-choice questions are more objective (i.e., it is easier to identify correct and incorrect answers), and it is more likely that participants will have had previous experience with multiple-choice questions. To keep our testing situation controlled, we will have participants record their answers on a Scantron bubble sheet. To make sure we have a high-quality test, we will pilot test the questions ahead of time. Ultimately, we will operationally define distraction as participants’ performance on the 40-question multiple-choice test, with lower scores indicating more distraction.

**Our Hypothesis**

Our original research question has led us to think through a variety of issues in order to identify the best way to test our idea that overly restrictive smartphone-use policies may be a distraction. While others may choose to tackle this question in a different way (which is the beauty of science), we have settled on a simple experiment where we will manipulate our independent variable of smartphone use policy. We will have two groups, as detailed in Table 8.1. The experimental group will have a high level of restriction. Participants in this “forbidden phone” group cannot use their phones in any way. The control group, or “phone checkers,” will have restrictions but will still be able to check their phones. To compare how distracting these two policies are, we will measure our dependent variable of distraction by examining test performance.

| TABLE 8.1

<table>
<thead>
<tr>
<th>Our Two-Group Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smartphone Policy</td>
</tr>
<tr>
<td>High Restriction</td>
</tr>
<tr>
<td>Low Restriction</td>
</tr>
</tbody>
</table>

Based on these, we can formally state our experimental hypothesis, where we make a clear and specific prediction of how the independent variable will influence the dependent variable. We could test the following hypothesis:

*Those in the highly restricted smartphone-use group will perform differently on the test than those in the group who are allowed to check their phones.*

With this prediction, we are making a nondirectional hypothesis that we expect the groups to differ, but without saying how they will differ. We make this type of prediction when we think that restricted smartphone use will cause a difference in how well one does on the test compared to the nonrestricted group, but we are unsure if the restriction will improve or hurt test performance.

Based on what we already know about impulse control and thought suppression, we have a theoretical reason to believe that restricting smartphone use will be more distracting than allowing phone checking. Therefore, we do not
just want to predict a difference between our groups; we want to predict the
direction of that difference. Specifically, we will test the following hypothesis:

Those in the highly restricted smartphone-use group will perform worse on the
test than those in the group who are allowed to check their phones.

You will notice that we predicted a specific outcome by stating which of the two
groups would perform worse. This is called a directional hypothesis because we
specified which group would have higher (or lower) scores on the dependent
variable.

Before we move on, let’s be sure we understand the experimental hypothesis,
because it seems a bit counterintuitive. One group is not allowed to check their
phones, while the other group is. Common sense would suggest that the phone
checkers would be more distracted. However, science suggests that trying to sup-
press or ignore thoughts (in this case, thoughts about whether you are missing
messages on your phone) is distracting. Thus, our experimental hypothesis pre-
dicts that being forced not to think about a smartphone will leave participants
more distracted than if they were allowed to check it. This, in turn, will impede
learning as measured by performance on the test.

YOUR TURN 8.1

1. Professor Dunphy conducts a study where two groups listen to a 50-minute
audio recording of a lecture. The first group listens to a lecture on art
history and the second group listens to a lecture on Greek philosophy.
After the lecture, each group completes a questionnaire that measures
intellectual curiosity. In this study, the independent variable is _____, while
the dependent variable is _____.
   a. length of the lecture; lecture topic
   b. length of the lecture; intellectual curiosity
   c. lecture topic; intellectual curiosity
   d. intellectual curiosity; lecture topic

2. A school psychologist believes that preschoolers who eat chicken nuggets in
the shapes of letters show a greater interest in books. For the experimental
group, preschoolers will eat five chicken nuggets shaped like the letters A,
E, I, O, and U. What would be the best control group to test the school
psychologist’s belief?
   a. A group that eats 5 ounces of macaroni and cheese with noodles shaped
      like the letters A, E, I, O, and U.
   b. A group that eats a bowl of cereal with alphabet-shaped bits.
   c. A group that eats five chicken nuggets shaped like the letters D, N, R,
      S, and T.
   d. A group that eats five regularly shaped chicken nuggets.

3. Cognitive-behavioral therapist Dr. Johnson believes her type of therapy
is best for treating nocturnal enuresis (nighttime bed-wetting). To test her
assumption, she assigns one group of bed-wetters to cognitive-behavioral
therapy for 3 months, while the other group undergoes psychodynamic therapy for the same time period. Both groups’ therapy sessions are led by an independent therapist, Dr. Jack, who is not aware of Dr. Johnson’s prediction. At the end of treatment, Dr. Johnson has a neutral therapist assess symptoms in both groups. In this study, the experimental group is the _____, while the control group is the _____.

a. cognitive-behavioral therapy group; psychodynamic therapy group
b. psychodynamic therapy group; cognitive-behavioral therapy group
c. group led by Dr. Jack; group led by Dr. Johnson
d. therapist; symptoms

How did you do? Turn to the end of the chapter to check your answers.

Design in Action

We know what we want to have in our study in terms of the manipulation and measures. We even have some of the basics in place for establishing experimental control and using appropriate measurements. Next, we need to turn our attention from the “what” of our study to the “how,” as in, “Just how are we going to carry out the study in a way that meets our goals?” The first decision we have to make is from where to gather our participants.

Weighing Our Options: Obtaining Participants

Generally speaking, it is best to use a random sample of the overall population as the participants in an experiment. In our case, that would mean randomly selecting from the entire population of college students who have smartphones. In practice, obtaining a sample of that entire population is nearly impossible, not to mention prohibitively expensive. Instead, a convenience sample of students at our own school makes the most sense. Convenience sampling has the benefit of being easier to carry out, and, since students may need to participate in research as part of their classes, the sample is not expensive to obtain. Although we may lose some ability to generalize to other populations if students at our school react differently than students at another school would, we need to remember that no study is perfect. Your job as the researcher is to know enough about the relevant issues to make the choices that will best answer your research question. First, we need to figure out whether or not restrictive smartphone policies actually create distraction. If we establish that this happens, then we can conduct follow-up studies to see if we can replicate our result in other groups of students.

Next, we need to determine how we will obtain our convenience sample. We only want students who regularly use smartphones, so we should be sure to screen out those who may not have a phone or may not use their phone often. Thus, when we recruit participants, we will tell them, “We are conducting a study of smartphone habits,” and then ask them to text us to set up an appointment. This method will ensure that the participants are highly familiar with texting, and will allow us to ask them to bring their phones to the study without arousing suspicion.

What Do You Think?

Is sampling only college students and not sampling other parts of the population, such as older people who are not in college, a problem? Why or why not?
Weighing Our Options: Placing Participants in Groups

We want to make sure we have enough participants to reasonably determine whether or not the smartphone restrictions make a difference, given all the other random factors that may influence their ability to concentrate. Based on our previous planning, we know that we will have two groups (the “forbidden phone” group and the “phone checkers” group), and we should aim for about 50 participants in each group. Collecting data from 100 participants will definitely be time-consuming.

The Importance of Independent Groups

One way to speed up data collection is by testing several participants at a time. Our room holds up to eight people, so we can sign up that many participants for each of the study’s time slots. Once the eight participants arrive, we can tell them all which group they are in, making sure that we have equal numbers in each group. But this approach has a downside. It is possible that participants will stop acting like eight individuals and start behaving as a single group. In research terms, this can compromise our study’s independence, which is the assumption that each participant represents a unique and individual data point. As we know from social psychology and the minimal groups paradigm (Struch & Schwartz, 1989; Tajfel, 1970), as soon as people know they are part of a particular group, they favor that group. As a result, in our study, participants in Group 1 will like their fellow Group 1 members more than they like those in Group 2, behaving less like autonomous individuals. Participants may even take cues from other group members in ways that make each individual participant act less naturally. For example, the behavior of others in the room may have more influence than experimental manipulation over whether or not participants look at their smartphones. In addition, those in the experimental condition may start complaining to each other about being forbidden to use their smartphones. Thus, it is important to minimize interactions between the participants. The simplest way to solve this problem is to test one participant at a time. Although this approach will take up more time, it minimizes the possibility that interactions between participants will undermine our study.

Problems with Nonrandom Assignment

Even when collecting data from one participant at a time, we need to be careful about how we assign each participant to the experimental or control group. We could place participants in either group simply by allowing them to decide whether they prefer the high-restriction or low-restriction group. However, doing this creates two problems. First, in order to choose, participants will need to know what both groups will do in the study, which increases the chances that participants will accurately guess the purpose of the study and the researcher’s hypothesis. Second, participants who pick the high-restriction group may do so because they naturally care less about monitoring their smartphones. Thus, if we found that high restriction led to less distraction, it would be impossible to know if the finding was due to the restrictive policy or if it was due to the high-restriction group having more participants who cared less about monitoring their smartphones.
Perhaps it would be better if you, the researcher, placed participants in groups. As the person who knows the study best, you seem like a natural choice. However, although you may believe you are assigning groups in a fair and ethical manner, there is always the chance that you may bias your assignments. For example, what if you unknowingly tend to put participants with iPhones in the low-restriction group? That may not seem like a big deal, but it is possible that iPhone users are more addicted to their phones. If this were the case, you could find differences between your two groups that had nothing to do with your restriction manipulation.

The Importance of Random Assignment

To avoid these problems, we need to eliminate human choice and potential bias from the group assignment process. We can do this by using random assignment, which is any method of placing participants in groups that is nonsystematic and nonbiased, and that ensures each participant has an equal chance of being in any group. We could determine random assignment with a simple flip of a coin. When the coin is heads, we could assign that participant to our “forbidden phone” group. When it is tails, we could assign the participant to the “phone checkers” group. We could roll a die and designate odd rolls as one group and even rolls as the other. Or, we could place the name of each group into a hat and draw one out each time we have a new participant in our study. Better yet, we could use an online resource. There are websites such as http://www.randomizer.org that generate random numbers we could use to assign participants to groups.

Now you may be thinking, “What if there are some participants who are complete texting addicts, while there are others who just got their first smartphone?” As the researcher, you would want to ensure that all of the texting
Two-Group Design

addicts are not in one group, with all of the novices in the other, which theoretically could happen through random assignment. In cases where you can easily identify important differences like this, you can use a matched-pair design, in which you create a set of two participants who are highly similar on a key trait (e.g., two texting addicts), then randomly assign one to the experimental group and the other to the control group. Doing so makes sure that both groups have participants who are similar, or matched, on that important trait. This technique can work well if there are participant characteristics that you as the researcher know are important and can measure in a reliable and valid way. The problem is that researchers cannot possibly know every variable that can influence the dependent variable. Thus, a matched-pair can sometimes create a false sense of security where you think you have your groups matched and highly similar, when, in reality, some other unmatched participant trait influences the outcome.

The list of potential influences in any study is endless. For example, in our study, participants’ age, sex, ethnicity, socioeconomic status, past experience with a smartphone, data plan, number of friends, and so on might influence how our restriction manipulation affects participants. It is also possible that what happened earlier that day, current mood, typical attention span, tiredness, general disinterest, and so on can influence the participant’s distraction level. Thankfully, random assignment accounts for all of these potential influences, plus all of the ones we did not think of. Notice that we say “accounts for” and not “eliminates.” There will always be factors that are unaccounted for that may have an influence, so the best we can do is have these factors influence both groups similarly. Provided you have enough participants (which is why we want 100 participants in our study rather than 10), random assignment will even out any variable’s influence. Because participants have an equal chance of being in any group, one group should not end up with a disproportionate number of tired, disinterested, or text-obsessed participants.

So, we will randomly assign our participants to either the “forbidden phone” group or the “phone checkers” group. Because we want to know how these two groups compare on test performance, we will use a between-subjects design, where we expose our participants to only one of the two groups. This means we will also only assess the dependent variable of test performance once for each participant.

Developing a Protocol

Now that we know how we are going to find participants and assign them to either the experimental or control group, we can make decisions about our procedure. As we do, we will want to maintain experimental control by keeping the procedure identical for everyone. We know that using a protocol will help us accomplish that goal. We decided earlier that having participants come to a controlled, “laboratory” setting would be better than using an actual classroom setting, so we will plan on collecting data in a small classroom on campus.

When they arrive for the study, we will first ask our participants to read an informed consent form. Our informed consent will include enough information for our participants to make an educated decision about whether they want to be
in our study, without revealing too many details that could cause participants to act unnaturally. In our case, we will tell participants, “We are conducting a study of smartphone habits that will require participants to watch a video, read a short passage, and complete a brief quiz.”

Following informed consent, we will direct each participant to sit at a desk in the middle of the room. Once the participant takes a seat, we will explain that his or her task is to watch a 30-minute video lecture on the nature of clouds. The participant will then read a textbook chapter section that is not too long or boring on the same topic. (Clouds are a good topic because they should interest participants, while likely not being a topic that they are extremely familiar with.) Prior to putting on the video, we will ask the participant to switch his or her smartphone to vibrate and place it out of sight in a metal basket under the desk. This instruction might seem silly to add to the protocol, but smartphones can vary in terms of their vibration intensity, which makes some easier to hear than others. A metal basket will ensure that the vibration is consistently audible. Never underestimate the importance of these types of details when designing a study!

Everything up to this point has been the same for all of our participants. Now we need to manipulate our independent variable, giving participants slightly different instructions so that they know which “smartphone policy” to follow (allowed to check their phones vs. not allowed to look at them). Remember, we want to keep our instructions as similar as possible. As a result, everyone will hear the same first part of the instructions:

Please adjust your smartphone settings so that all sound effects are off. In case you receive any texts, please leave your phone on vibrate for texts only. Once you have the settings adjusted, please place your smartphone out of sight in the basket below your desk. You should not, under any circumstances, use your smartphone to make a call, check e-mail or social media, play games, or send any text messages.

These instructions help ensure participants all set up their phones the same way, and ensure that they all know not to actually use their phones. The second part of the instructions will vary depending on which group the participant is in:

**High-Restriction “Forbidden Phone” Group:** If your phone vibrates, you must ignore it and not think about it. Do not pick up your phone to check who the message is from or read the message. Leave your phone in the basket.

**Low-Restriction “Phone Checker” Group:** If your phone vibrates, you can check to see who the message is from and read the message. Once you are done, return your phone to the basket.

It is the second part of the instructions that establishes the key “ingredient,” or difference, between the groups: *No Phone Checking* versus *Phone Checking Allowed.*
We want to be sure the study feels as real as possible. In other words, we want to ensure that our study has **experimental realism**, such that participants become engrossed in the manipulation and feel like it influences them. Rather than having participants sit at a desk by themselves, we could have others sitting at desks as well. Certainly, in a class environment, there are other people sitting near you, but where are we going to find the extra people? One option would be to have other participants sitting in surrounding desks. But in this scenario, we will have less control over the situation. Participants may know each other, and different combinations of participants (e.g., all males or all females) may lead to different dynamics. A better solution would be to use confederates, or accomplices of the study, who will know to treat all participants the same. When the participant enters the room, we can have five confederates already sitting at other desks. We can be sure that our confederates always sit in the same spots and act in the exact same way for every participant.

In order to have participants either ignore their phones or check them, we need to be sure that the phones actually vibrate during the study session. The only way to ensure that each participant receives a text is for us to be the ones who send it. To do this, we will need the participant's cell phone number. Since participants signed up for the study via text message, we can ask participants, following informed consent, to provide their phone numbers so that we can allegedly check off who showed up to the study from our list of numbers.

During the 30-minute video lecture and subsequent reading, we will text the participant several times. In order to standardize the experience for everyone, we will text a total of eight times, spaced out in a preplanned way. We could space the texts out so that we send one every 6 minutes, but that may seem too obvious. Instead, we will plan on sending a text at the following minute marks during the video: 2, 9, 15, 20, and 26. We will then text participants three additional times while they are doing the reading, at the 1-, 3-, and 4-minute marks. When we send the texts, the participant should not know it is us, which means they will not recognize the sender. Rather than sending the participant five texts from the same unknown number, we should mix it up by sending texts from three different phone numbers. The content of our texts should be general so that it applies to all participants (e.g., “heyyy how are you?? :)”), should be informal in terms of grammar and punctuation in order to seem more authentic, and should be the same for each person in our study.

There is one potential problem: Some participants may get real texts while watching the video or reading the selection. Because we cannot do much about this problem, we will keep **researcher notes**, in which we keep track of anything out of the ordinary that happens during the study. For example, we would note if any of the participants in the high-restriction/forbidden smartphone group looked at their phone in violation of our “policy.”

After watching the video, the participant will start the reading portion of the lesson. To be sure that this portion of the procedure is also controlled, we will give all participants 5 minutes to read the short selection. If they finish early, the researcher will ask them to reread the passage until the 5 minutes are up. As we walk away, we will start a timer and stop the participants after 5 minutes. Once the participants complete the reading, we will administer the knowledge test.

---

**What is a confederate?**

See Chapter 3, p. 68

---

**experimental realism**

The degree to which a study participant becomes engrossed in the manipulation and truly influenced by it.

**researcher notes**

A place to keep track of anything out of the ordinary that happens during a study.
Following the test, we will give each participant a manipulation check, which is a measure that helps determine whether the manipulation effectively changed or varied the independent variable across groups. Because we were manipulating whether participants could check their smartphones, we need to see if the two groups differed in their perceptions of the permissiveness to check their phones during the study. We can simply ask them to rate on a scale of 1 to 7 how permissible it was to check their phones. If we find that there is no difference between the high-restriction and low-restriction groups on this question, then we cannot be certain we adequately manipulated our independent variable. This is important information to have if we find no differences between the experimental and control groups on our dependent variable. It may help us decide whether a smartphone ban is distracting or whether there was a flaw in our study because the manipulation did not work. We will also want to check if participants noticed the phones vibrating by asking them to report the number of times they detected it going off. This is important because if participants in the highly restricted group did not hear the phones, then they were not actually ignoring them as we hoped.

Next, we will ask participants to answer a few demographic questions and several additional questions related to our study’s topic. For example, we should ask questions about the participant’s typical smartphone use, texting habits, GPA, and knowledge about clouds prior to the study. In each case, these variables may help us understand potential results. Finally, we will debrief participants and thank them for participating. Because we are using a number of different materials in our study (i.e., Scantron testing sheet, demographic sheet, manipulation check), we want to be sure that we place a participant number on each piece of material. That way we can easily match all the materials associated with a particular participant when it comes time to enter our data.

Focus on Open Science: Preregistering Your Hypotheses, Materials, and Data Analysis Plan

Before we collect any data, we will be sure to preregister our study on an open science website (e.g., http://www.socialscienceregistry.org). There, after our study, we will share study details such as how we manipulated our independent variable (including the exact instructions participants heard) and our study materials, including the video participants saw, the article they read, and the quiz they completed. Based on our hypothesis, we will also preregister our plan for analyzing the data we will collect. For this study, we will want to make sure our manipulation worked by seeing if our two groups (“Forbidden Phone” and “Phone Checker”) differ on how permissible it was to check their phone, as well as their self-reported distraction. We will also test our hypothesis by comparing the two groups’ test scores to see if the “Forbidden Phone” group had lower test scores. By having a preregistered plan, we can focus only on the analyses that help us test our preregistered research hypotheses. Once we have collected all of the data, we will post the data set on an open science website where other researchers can conduct their own analyses.
**YOUR TURN 8.2**

1. Murray, a server in a restaurant, wonders if writing a simple “thank you” on the check before he delivers it will increase his tips compared to not writing anything at all. What he is most worried about is the individual difference between his patrons in tipping habits, as some people naturally tip at a higher percentage than others. How can Murray reduce this concern?
   a. By having patrons report their average tipping tendencies
   b. By using random assignment
   c. By using random sampling
   d. By only using patrons that are known to be bad tippers

2. Emily wants to know if mood impacts cooperation. She has half of her participants watch a slide show that shows sad images and half watch a slide show of neutral images. Afterward, she gives each participant the opportunity to cooperate on a game with a player in another room. At the end of the study, she collects demographic information and administers a measure of mood. Which element of her study was included as a manipulation check?
   a. The sad slide show
   b. The neutral slide show
   c. The opportunity to cooperate
   d. The mood measure

3. As an animal trainer, Melaine wants to test which type of reward (crab vs. fish) will work best in training otters for her new amusement park show. She creates sets of subjects based on both age and the level of previous training before she begins the training. What technique is Melaine using?
   a. random assignment
   b. random sampling
   c. matched pairs
   d. independence

*How did you do? Turn to the end of the chapter to check your answers.*

**Statistics: In Search of Answers**

After we have run the study and collected information from 100 participants, we need to enter our data into the computer for analysis. Because we conducted an experiment, it is crucial that we identify each participant’s group or condition prior to entering their test scores, manipulation checks, and demographics. In our researcher notes, we should have a list of participant numbers that includes their groups and any notes that we made regarding each participant’s experience.
Prior to entering data, we will need to have the Scantron sheets scored so that we have the participants’ test scores. Once we have all of the information, we can enter the data into a statistical computer program, as shown in Figure 8.2. Even though participants will be in different groups, any time we enter data, each participant’s information must appear on one row. The key difference with an experiment compared to a survey is that in one column we will have a variable for the independent variable (“Condition”) with numbers identifying each group (1 = Experimental Condition, 2 = Control Condition).
With the data entered, we can focus on answering our research question, “Can trying to ignore a text message lead to distraction?” Given our design, we will want to see if our high-restriction group had test scores that were significantly lower than the low-restriction group. Did the participants in the high-restriction group do worse on the test as hypothesized, and did they do worse enough for us to be relatively certain that it did not happen by chance?

Selecting the Proper Tool

To answer our research question, we need to accomplish three tasks during our data analysis:

1. Make sure our groups were similar on variables that may provide alternate explanations.
2. Make sure our manipulation was effective.
3. Test our hypothesis.

Our first set of analyses should focus on making sure that alternate explanations (e.g., number of actual texts received, GPA, familiarity with clouds, etc.) are not responsible for any potential findings. Random assignment should have ensured that those variables were controlled for or kept similar in both groups, but, as the saying goes, “Trust, but verify.” We can verify that our two groups are similar by using statistics, and we can determine the right statistic to use by considering key features of our study. We have a between-subjects design, two levels of the independent variable, and a continuous dependent variable. Based on these qualities, we should run a \textit{t-test for independent means} (also known as an \textit{independent samples t-test}), which is, generally speaking, a test to see if the two groups or conditions are different. The \textit{t}-test, the anatomy of which is shown in Figure 8.3, determines this by comparing each group’s mean to see if they differ from one another.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
\textbf{t-test Symbol} & \textbf{t Score} & \textbf{Significance Level} & \textbf{Calculated Effect Size} \\
\hline
\hline
\textbf{t} (98) & \#\#\# & \#\# & \#\# \\
\hline
\end{tabular}
\caption{Anatomy of a t-test for Independent Means.}
\end{table}
to a degree that could not have just happened accidentally. Remember here that statistical significance (represented by $p$) is the probabilistic indication (really just a percent likelihood) of how much confidence we have that the two groups differ. If the $t$-test for independent means is significant ($p < .05$), we can be fairly confident that our results represent a real difference between the groups. If the $t$-test for independent means fails to reach significance ($p > .05$), there is not enough evidence to suggest that the groups are different.

A $t$-test only indicates the likelihood of there being a difference. Regardless of whether the difference is significant, we often want to know the effect size (represented by $d$), or the magnitude of the difference, between the groups. The effect size allows us to evaluate the practical significance of our findings. For example, suppose we find that forbidding smartphone checking leads to students performing significantly worse on a test. Before we start advocating for large educational reforms, we should evaluate how large of an impact this policy has on test performance. If the effect is small, we might want to direct our efforts to improve student learning elsewhere. If it is large, we will want to consider how to minimize such distractions in the classroom.

Second, in order to be sure our manipulation did what we intended it to do, we need to make sure participants knew the instructions. We can do this by comparing the two groups on whether they thought smartphone checking was permissible. Related to the effectiveness of our manipulation, we should also verify that participants heard the texts, because if they did not hear the texts, there would not have been any thoughts for them to suppress. Similarly, we should verify that the high-restriction group reported more distraction, because the logic of our hypothesis depends on restriction creating more distraction. We can run another set of $t$-tests for independent means to see if the groups were different in how well they heard the texts and in how much they felt distracted by their smartphones.

Finally, we need to test our hypothesis that those in the highly restricted smartphone-use group will have worse test performance than those in the group who can check their phones. Since we are looking for differences in a between-subjects design with two groups on a continuous dependent variable, we will again use a $t$-test for independent means. Also, because we have a directional hypothesis, we are predicting how our groups will differ. In experiments, we rarely predict that there will be no difference. A “no difference” hypothesis, also known as a null hypothesis, is most often the hypothesis that we are trying to statistically reject. That is, we start out with the assumption that there is no difference between groups. Our goal is then to show that there is a difference between groups, by demonstrating that the assumption of no difference is unlikely given our results. The null hypothesis is an important part of the logic behind the statistics we use to test or analyze our hypothesis, but it is something that researchers naturally assume, and thus do not explicitly state when writing up the results.
CHAPTER 8  Two-Group Design

Writing the Results in APA Style: *t*-test for Independent Means

**Results**

Means and standard deviations for the two experimental groups on key variables appear in Table 8.2 (all analyses are two-tailed).

**Alternate Explanations**

To determine if the high- and low-restriction groups differ on several variables that could potentially serve as alternate explanations when testing our hypothesis, we conducted a series of *t*-tests for independent means. The results of those analyses were as follows: typical smartphone use, *t*(98) = 1.31, *p* = .19, *d* = .27; texting expertise, *t*(98) = 0.29, *p* = .78, *d* = .16; number of real texts during the study, *t*(98) = 0.31, *p* = .75, *d* = .06; grade-point average (GPA), *t*(98) = 0.18, *p* = .86, *d* = .04; and previous cloud knowledge, *t*(98) = 0.67, *p* = .50, *d* = .14. These analyses suggest that our two groups were not significantly different on these variables.

**Manipulation Check**

To determine if the high- and low-restriction groups differed on how permissible it was to check their smartphones, we conducted a *t*-test for independent means comparing the groups. The analysis was significant and had a moderate effect size—*t*(98) = 51.19, *p* < .001, *d* = 10.34—suggesting that the manipulation was effective.

We also wanted to see if smartphone restriction influenced participants’ self-reported distraction. To do this, we conducted a *t*-test for independent means comparing the high- and low-restriction groups on distraction scores. The analysis was significant and had a moderate effect size: *t*(98) = 4.72, *p* < .001, *d* = .95. This suggests that those in the high-restriction group were significantly more distracted than those in the low-restriction group. We also conducted another *t*-test for independent means to determine if the groups differed on how often they heard the smartphone ring. Results of that analysis, *t*(98) = 0.98, *p* = .33, *d* = .20, suggest that the groups were not significantly different.
Hypothesis 1

To determine if smartphone restriction influenced participants’ quiz scores, we conducted a $t$-test for independent means comparing the high- and low-restriction groups on quiz scores. The analysis was significant and had a moderate effect size: $t(98) = 2.32, p = .022, d = .47$. As hypothesized, those in the high-restriction group had significantly lower quiz scores (an average of 28.88 points out of 40 total points, or 72.20% correct) than those in the low-restriction group (an average of 32.10 points out of 40 total points, or 80.25% correct). Results also appear in Figure 8.4.

Don’t Just Tell Me, Show Me: Using Tables and Figures

In addition to presenting their findings in the text of a research report, researchers also depict the results of their experiment in tables and figures. These visuals provide another way to summarize the data. Table 8.2 shows the differences between our restriction conditions on key variables, and Figure 8.4 displays the differing quiz percentages between our high- and low-restriction participants.

<table>
<thead>
<tr>
<th>Variable</th>
<th>High Restriction</th>
<th>Low Restriction</th>
<th>$t$</th>
<th>$p$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical Smartphone Use</td>
<td>6.12 (0.82)</td>
<td>6.34 (0.85)</td>
<td>1.31</td>
<td>.19</td>
<td>.27</td>
</tr>
<tr>
<td>Texting Expertise</td>
<td>6.36 (0.69)</td>
<td>6.32 (0.71)</td>
<td>0.29</td>
<td>.78</td>
<td>.16</td>
</tr>
<tr>
<td>Real Texts During Study</td>
<td>1.18 (1.24)</td>
<td>1.26 (1.31)</td>
<td>0.31</td>
<td>.75</td>
<td>.06</td>
</tr>
<tr>
<td>GPA</td>
<td>3.27 (0.58)</td>
<td>3.29 (0.64)</td>
<td>0.18</td>
<td>.86</td>
<td>.04</td>
</tr>
<tr>
<td>Cloud Knowledge</td>
<td>5.04 (2.59)</td>
<td>4.68 (2.76)</td>
<td>0.67</td>
<td>.50</td>
<td>.14</td>
</tr>
<tr>
<td>Distraction</td>
<td>5.54 (1.47)</td>
<td>4.08 (1.61)</td>
<td>4.72</td>
<td>&lt;.001</td>
<td>.95</td>
</tr>
<tr>
<td>Times Heard Phone</td>
<td>7.52 (1.07)</td>
<td>7.28 (1.37)</td>
<td>0.98</td>
<td>.33</td>
<td>.20</td>
</tr>
<tr>
<td>Quiz Score</td>
<td>28.88 (7.12)</td>
<td>32.10 (6.74)</td>
<td>2.32</td>
<td>.02</td>
<td>.47</td>
</tr>
</tbody>
</table>

Note: $n = 100$. Higher scores indicate a greater magnitude of each variable. All analyses are two-tailed.
Our Research Plan at a Glance

What Is Our Research Question? Can trying to ignore a text message lead to distraction?

What Is Our Design? We are using a two-group between-subjects experimental design.

Why Are We Using This Design? This design allows us to establish causality by manipulating one variable (ignoring text messages), while keeping other factors controlled (i.e., the same) so that we can determine how it affects outcomes (distraction).

What Are Our Variables?

Independent Variable(s): Ignoring Text Messages

High-Restriction “Forbidden Phone” Group: If your phone vibrates, you must ignore it and not think about it. Do not pick up your phone to check who the message is from or read the message. Leave your phone in the basket.

Low-Restriction “Phone Checker” Group: If your phone vibrates, you can check to see who the message is from and read the message. Once you are done, return your phone to the basket.

Dependent Variable(s): Distraction

Participants watch a video lecture and complete a short reading about clouds, followed by a 40-question multiple-choice test. Test performance (score out of 40) serves as an indicator of distraction (lower test performance = more distraction).

What Are Our Hypotheses? Those in the highly restricted smartphone use group will perform worse on the test (i.e., display signs of more distraction) than those in the group who are allowed to check their phones.

Who Are Our Participants? Convenience sample of college students who regularly use a smartphone.

What Ethical Considerations Do We Need to Keep in Mind?

• Since smartphone use could potentially hurt actual test performance, we would not want to run this study in a real class where it could lower students’ actual test scores.

What Is Our Data Analysis Plan?

1. Make sure our groups are similar on variables that may provide alternate explanations. We compare the two conditions (Experimental/“Forbidden Phone” vs. Control/“Phone Checker”) on variables such as number of actual texts received, GPA, and familiarity with clouds using a t-test for independent means.

2. Make sure our manipulation is effective. We compare both groups to see if they report a difference on how permissible it was to check their smartphones and self-reported distraction using a t-test for independent means.

3. Test our hypothesis. We compare the two conditions (Experimental/“Forbidden Phone” vs. Control/“Phone Checker”) of our independent variable with a t-test for independent means, this time to determine if the groups had different quiz scores. If the data support the hypothesis, the “Forbidden Phone” group will have lower quiz scores, indicating more distraction.

Want to practice the analyses for this research yourself? Ask your instructor about the data set that accompanies this study.
Let’s Discuss What Happened

Our experiment attempted to determine how course policies concerning students’ smartphone use influence students. Based on the premise that trying not to think about something makes you think about it more, we hypothesized that overly restrictive smartphone policies may create more distraction and undermine student learning. To test this, we conducted a two-group experiment where we randomly assigned students to a group that was forbidden to check their phones or to a group that was allowed to check their phones. With the results of our experiment in hand, we must now interpret our findings. As we do this below, notice how we place our findings in the context of previous research and theory, and how we identify the limitations of our design and make concrete suggestions for future research.

What Did We Find?

A series of analyses testing potential alternate explanations found no significant differences between the groups on typical smartphone use, texting expertise, the number of real texts the participant received during the study, grade point average (GPA), or previous cloud knowledge. The results of our study suggest that our manipulation was effective. Participants in both groups were aware of their group’s restrictions, both groups heard the smartphone a similar number of times, and, most importantly for the assumption underlying our hypothesis, the high-restriction group reported feeling more distracted than the low-restriction group. The test of our hypothesis revealed a significant difference, such that participants in the high-restriction group had test scores that were 8% (nearly a full letter grade) lower on average than participants in the low-restriction group.

Why These Findings?

Participants who were forbidden to check their phones performed worse on the test and reported experiencing more distraction. Although people believe they have the ability to control their thoughts, our findings are consistent with previous research on thought suppression showing that attempting to control or suppress a thought is futile (Najmi et al., 2010). In the context of a restrictive smartphone policy, the participant needs to enact a negative implementation intention or plan not to act (e.g., “If I get a text, I will not check my phone”) in order to follow the restriction (Adriaanse, van Oosten, de Ridder, de Wit, & Evers, 2011). However, research on these strategies shows that planning not to act on a habit can actually strengthen the habit by increasing thoughts of the target behavior, which is consistent with Wegner’s (1994) ironic process theory. By asking our high-restriction group to plan on not checking their phones (a negative implementation intention), we likely led them to think about their phones more than they would have.

These thoughts are likely distracting because the person receiving the text wonders about the message’s content (e.g., “Is it important?” “Is there something wrong?”). Ironically, the text’s actual content is probably much more benign. In fact, teens tend to text others for mundane reasons like wanting to say hi or to chat rather than for important personal matters (Lenhart, Ling, Campbell, & Purcell, 2010).
Rather than restricting smartphone checking, which conjures anxiety (Skierkowski & Wood, 2012) and thoughts the student futilely tries to suppress, having course policies that permit students to check their phones is not only something students want (Tindell, 2012), but may be something that allows them to spend more time focusing on class material.

**What Could Be Improved?**

A major benefit of the present study was that we looked at participants’ actual behavior rather than self-report, which is important since previous research shows that self-reported smartphone use and actual smartphone use are not significantly correlated (Underwood, Rosen, More, Ehrenreich, & Gentsch, 2012).

There are a few potential limitations to our study. First, our sample focused entirely on college students, ignoring older segments of the population and those who do not attend college, which could undermine generalizability. However, given the nature of our research question (how course policies affect college students), we must have a target population of college students. Some may also question whether participants in our study really cared about the texts since they came from us. Although this limits realism to some degree, many of our participants still received “real” texts during the study, it would not have been immediately obvious that the texts were from us, and responding to texts may be such an automatic response that the sender’s identity is not a key influence (and definitely was not for those in the group who were not allowed to check their phones). We cannot rule out the possibility that the lack of realism influenced the results, but the fact that we got an effect with this limitation suggests that in real-life settings, texts may be even more distracting.

In any simple experiment, there is always room for improvement, especially in terms of measuring additional factors that contribute to the dependent variable. In our case, though we found that placing restrictions on phone checking was problematic, we cannot be sure if students who could respond to texts (i.e., an even lower level of restriction) would do better or worse on the test. The truth is that there are numerous other variables that a researcher could test or additional factors that a researcher could control. In each case, it will become useful to add additional groups to the research design, a technique we will discuss more in Chapter 9.

**What’s Next?**

Importantly, our study is silent in terms of how checking one’s phone can potentially distract others. Thus, future research could use our same paradigm, but have a confederate be the person with the phone and see how others nearby react. It is also possible that the difficulty of the lesson or course content influences the smartphone’s ability to distract a student. Research on cognitive load, the extent to which a person has working memory available, suggests that in more demanding classes where load is higher, smartphone use could make it more difficult for students to pay attention and ignore their smartphone (Lavie, 2010).
Let's Discuss What Happened

The present findings can also apply to the workplace in terms of how interruptions influence job performance. Given the present findings regarding the negative effects of trying to ignore an interruption, it may be best to allow employees to check their smartphones. In fact, research shows that interruptions hurt task performance, but that there is a learning effect that takes place such that later interruptions are less damaging (Altmann & Trafton, 2007). Research on restrictive smartphone policies in workplace settings would help establish the true cost of smartphone use on worker productivity.

YOUR TURN 8.3

1. Fernando wants to determine if a person’s mood influences belief in God. He randomly assigns participants to either a positive or negative mood group, then administers a “Belief in God” measure, along with several demographic questions. To see if there is a difference between groups, which statistic should he use?
   a. $t$-test for dependent means
   b. Pearson $r$
   c. Effect size
   d. $t$-test for independent means

2. Amie reads a magazine story that says that people who undergo life coaching are more successful than those who do not. As a savvy statistics and research student, Amie wants to know the study’s _____, which will tell her whether the difference is significant, and the _____, which will tell her how large the difference was between the means.
   a. $t$ score; degrees of freedom
   b. $p$ level; $t$ score
   c. $p$ level; effect size
   d. effect size; $p$ level

3. In the discussion, we noted that several variables that could serve as potential alternate explanations did not differ between the groups. Based on this, which of the following statements is true?
   a. Participants’ GPA and cloud knowledge did not have any influence on test performance.
   b. The groups were relatively similar on these variables, meaning that we kept them relatively controlled.
   c. Random assignment was not effective.
   d. All students who were participants in our study had a high degree of texting experience.

How did you do? Turn to the end of the chapter to check your answers.

What Do You Think?
Given the research findings, what study would you conduct next to continue our understanding of smartphone use in classrooms?
CHAPTER 8 Two-Group Design

Research in Action

To Multitask or Not to Multitask?

Do you ever feel like your to-do list just keeps getting longer and longer with no end in sight? What is the best way to get everything done on your list? Should you go through the list item by item, or would it be better to do several things at once? Maybe you should first study for your exam, then do your laundry, and then text your friends to confirm your weekend plans. Alternatively, you could do these three things at the same time. After all, it could be just as effective as completing one task before moving on to the next.

Fortunately, we can rely on science rather than intuition to see which strategy is better. In the online activity To Multitask or Not to Multitask?, you will design your own simple experiment to solve this very question. Who knows—what you learn could help you actually complete all of those things on your to-do list!

LaunchPad To complete this activity, visit LaunchPad at launchpadworks.com

Final Thoughts

In this chapter, we moved beyond correlational research designs to begin establishing how one factor causes another. The simple two-group experiment allowed us to establish cause and effect by randomly assigning participants to groups and by implementing high degrees of experimental control. Our research question also demonstrated the potential benefits of studying counterintuitive research questions. That is, even though instructors adopt well-intentioned course policies, it may be foolish to blindly trust in their effectiveness. Rather, we should put science to use doing what it does best: testing assumptions.

Who knows—if a smartphone policy could unknowingly undermine learning, could other common course practices or policies be doing the same? Do pop quizzes help or hurt students’ scores on larger exams? If a professor requires class attendance, does that impose extrinsic motivations on students, when intrinsic motivation may serve them better? When professors give out their notes ahead of time, do students learn more or less? Does group work help students consider new perspectives, or does groupthink take over and ultimately narrow the range of ideas? You could conduct the research to find out. But before you set out to design a study, be sure to do a thorough literature review. You never know; it could be that others have already researched some of these questions, leaving it up to you to take the next step in broadening our understanding.
Review Questions

1. The local driving school owner, Lorraine, wants to see if her driving school students learn to drive better while listening to jazz music. To test this, she has half of her students listen to loud jazz music, while the other half listen to loud bluegrass music. Everyone learns the driving laws, parallel parking, and the meaning of road signs with music playing. Lorraine then measures the students’ performance on their driver’s license test. In this study the independent variable is ___, while the dependent variable is ___.
   a. music volume; driving test performance
   b. music volume; type of music
   c. type of music; driving test performance
   d. the driving skills in the lesson; music volume

2. As a TV producer, you are responsible for creating a reality TV show where 20 contestants from a variety of educational and experiential backgrounds and representing a variety of ages compete on physical and mental tasks like solving logic puzzles, building a pyramid, shooting a bow and arrow, completing a trivia quiz, and navigating an obstacle course. To create the two teams, you decide to put everyone’s name on a Ping-Pong ball and then draw teams like numbers in the lottery. Which of the following would present a problem to the fairness of your teams?
   a. Four of the contestants are math students from Ivy League schools.
   b. A majority of the contestants have been part of Habitat for Humanity and have building experience.
   c. Eight contestants consider themselves “trivia buffs.”
   d. None of the above.

3. Adele thinks the key to weight loss is eating expensive specialty dark chocolate bars that contain 65% cacao, pomegranate extract, and macadamia nuts. In particular, she thinks the pomegranate extract is the key ingredient. She wants to test her idea by comparing a group that eats 4 ounces of this specialty chocolate each day for a week to another group. What would be the best control group to test Adele’s belief?
   a. A group that eats 4 ounces of a white chocolate bar.
   b. A group that eats 4 ounces of dark chocolate containing 65% cacao, grape extract, and macadamia nuts.
   c. A group that eats 4 ounces of dark chocolate containing 85% cacao, raspberry extract, and peanuts.
   d. A group that eats macadamia nuts and drinks 4 ounces of pomegranate extract.
4. Eerikki works in human resources and believes that employees at his telemarketing company would be more productive if they took a 30-minute nap each day at 2:30 p.m. at their desks. To test this, he has half of the employees take a nap, while the other half sit quietly for 30 minutes at 2:30 p.m. at their desks. To determine productivity, Eerikki examines log sheets to measure the number of sales each group made. In this study the experimental group is the ___, while the control group is the ___.
   a. nap group; quiet rest group
   b. quiet rest group; nap group
   c. time of day; number of sales
   d. log sheets; number of sales

5. Chip is doing a study on gift-giving around the holidays. He finds that people who received lots of gifts for their birthdays during childhood expect more gifts as adults. In order for Chip to conclude a causal relation, he still needs to establish which of the following?
   a. Covariation
   b. Correlation
   c. Temporal precedence
   d. Elimination of extraneous variables

6. Royale is a tattoo artist at The Ink Spot. He believes that getting a tattoo makes people become more outgoing, brave, and confident. To test this, he finds 60 people who want to get a tattoo and randomly gives half of them a tattoo on their shoulder blade, while the other half get a temporary tattoo in the same spot. Royale asks everyone to keep a daily diary that he will use to determine how outgoing, brave, and confident participants act over the next 3 months. Once the study is underway, Royale becomes worried that some of his participants were already outgoing and finds out that about a dozen already had tattoos. Should Royale be concerned? Why or why not?
   a. Yes, if people were already outgoing, then a tattoo cannot have any effect on them.
   b. Yes, having people who already have tattoos ruins the whole premise of the study.
   c. No, random assignment should balance these types of differences out between groups.
   d. No, a dozen people isn’t enough to worry about.

7. Pastora is a club promoter who was hired to increase profitability at The Groove Lounge. Pastora decides to make the lights even dimmer and install oil drums with dry ice to make the entire club foggy. She thinks these changes will lead patrons to experience deindividuation, which will lead them to buy more food and drinks. As people leave the club at night, Pastora asks everyone demographic information, has them complete a deindividuation measure and a self-esteem measure, and asks the club owner how much
money was made each night. Which element of her study was included as a manipulation check?
   a. The deindividuation measure
   b. The self-esteem measure
   c. The amount of money the club made
   d. The number of dry ice blocks the club uses each night

8. Dr. Winston is doing a class demonstration where she has half of the class complete a series of 10 math problems, while the rest of the class completes a series of 10 analogies to determine which type of problem promotes greater alertness in her students. She randomly assigns students to groups to complete their problems quietly. Once everyone is finished, she asks everyone in Group 1 to report their current alertness level out loud. She then does the same for Group 2. Which of the following is a problem with Dr. Winston’s design?
   a. There was a problem with independence such that the groups started out being different.
   b. There was a problem with independence such that individuals knew their groups.
   c. Participants were not matched up based on math and verbal ability.
   d. Some participants may be math or vocabulary geniuses.

9. As a speech pathologist, Flavia always wants to make sure her clients get the best treatment. One day a salesperson for a neurological diagnostic company comes to Flavia’s office to share information about the effectiveness of a new diagnostic tool for stuttering to be compared with the current leading technique. The bar chart comparing the two groups looks impressive, but Flavia wants to see the statistical results. The appropriate statistic for comparing two groups is a ____. In those numbers, if Flavia wants to know how big the difference is between the two diagnostic tools, she would focus on the ____.
   a. t score; effect size
   b. Pearson r; p level
   c. p level; Pearson r
   d. t score; p level

10. Arielle hypothesizes that the more often a mother interacts with her child, the less distressed the child will be when left with a stranger. This is an example of what type of hypothesis?
    a. A nondirectional hypothesis
    b. A directional hypothesis
    c. A null hypothesis
    d. A causal hypothesis

11. Describe the benefits of using confederates in an experiment.
12. Why is using random assignment in experiments important when trying to establish causality?
13. What role does a manipulation check play when interpreting the results of an experiment?

**Applying What You've Learned**

1. Find a published empirical article that describes an experiment. Evaluate the groups in terms of experimental control. (A) What did they control well? (B) What could they have controlled better?
2. Go to the list of studies on [http://psych.hanover.edu/research/exponnet.html](http://psych.hanover.edu/research/exponnet.html). Find a correlational study that interests you and design a two-group experiment to test a similar research question.
3. This chapter’s study focused on how overly restrictive smartphone policies can undermine test performance. Design a study to test this basic idea on driving behavior. Be sure your study carefully considers potential harm to participants.
5. Texting and technology influence our lives in many ways. Generate five new ideas for studies that have the same focus, but that do not focus on classroom settings or on restrictive smartphone policies. Pick your favorite idea and describe how you would test it with a two-group design.
6. Read the following study: Kay, A. C., Wheeler, S., Bargh, J. A., & Ross, L. (2004). Material priming: The influence of mundane physical objects on situational construal and competitive behavioral choice. *Organizational Behavior and Human Decision Processes, 95*, 83–96. doi:10.1016/j.obhdp.2004.06.003. As you will see, in Study 1, researchers placed participants into either a business materials group or a neutral control condition, then had them complete a word completion task. (A) For this study, identify the independent variable, levels, dependent variable, and hypothesis. (B) What did the researchers control well? (C) What could they have controlled better?
7. **THE NOVICE RESEARCHER:** It is important to have experimental control. Playing the role of a novice researcher, design a two-group study that has at least four problems with experimental control. At the end, identify each problem and discuss how a more experienced researcher would have handled it.
8. **DIG INTO THE NUMBERS:** We have provided your instructor with supplemental data for a two-group design. Analyze that data to build your skills in using the *t*-test for independent means in SPSS. Write an APA-style results section based on your analyses. If you would like even more practice, your instructor also has data that accompanies the study discussed throughout this chapter.
**Key Concepts**

- control group, p. 251
- covariation, p. 249
- effect size, p. 269
- experimental control, p. 252
- experimental group, p. 251
- experimental hypothesis, p. 257
- experimental realism, p. 264
- extraneous variable, p. 250
- independence, p. 260
- independent samples t-test, p. 268
- internal validity, p. 250
- manipulation check, p. 265
- matched-pair design, p. 262
- mundane realism, p. 251
- null hypothesis, p. 269
- random assignment, p. 261
- researcher notes, p. 264
- simple experiment, p. 251
- temporal precedence, p. 249
- t-test for independent means, p. 268
- two-group design, p. 251

**Answers to YOUR TURN**

**Your Turn 8.1:**
1. c; 2. d; 3. a

**Your Turn 8.2:**
1. b; 2. d; 3. c

**Your Turn 8.3:**
1. d; 2. c; 3. b

**Answers to Multiple-Choice Review Questions**

1. c; 2. d; 3. b; 4. a; 5. d; 6. c; 7. a; 8. b; 9. a; 10. b