DISCOVERING CAUSE AND EFFECT

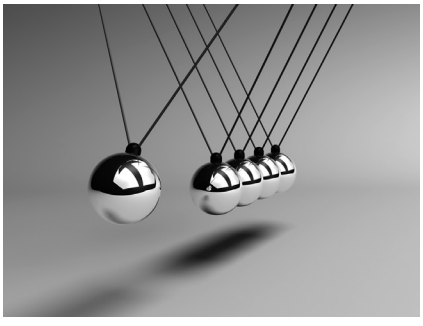
Life is a perpetual instruction in cause and effect.

—Ralph Waldo Emerson

LEARNING OBJECTIVES

- Explain how cause and effect work in an experiment.
- Compare the benefits of experiments to other methods.
- Identify the three basic criteria of experiments.
- Describe the elements of variation, confounds, control groups, and assignment.
- Develop a statement of the problem and answer the "so what" question for a study of your own.

This book is about experiments, the scope of which varies greatly. An experiment is a scientific test of some hypothesis or principle carried out under carefully controlled conditions in order to determine or discover something unknown. Experiments provide insights about the relationship between things where changes in one thing cause something to happen to another. We have all done informal experiments in everyday life without even knowing it, as the opening quote by Ralph Waldo Emerson illustrates. For example, if you change the amount of a certain ingredient in a recipe, does it taste better? If you drink lower-calorie beer, do you lose weight? As long as you only



iStock.com/terrymorris

change one ingredient at a time, or do not exercise more or eat lower-calorie everything, you probably assumed it was that one ingredient or that lower-calorie beer that caused the difference in taste or your weight loss.

CAUSATION

These examples of everyday experiments illustrate the concept of cause and effect. An "effect" is what happened. Better-tasting lasagna or weight loss are the effects in the earlier examples. The "cause" is the explanation for why these things happened—more garlic in the lasagna made it taste better, the lower-calorie beer helped you lose weight. We find the word *cause* used in everyday language, such as "the cause of death" or "the cause of an accident." The meaning is no different in experiments, but you will also see the terms *causality* or ***causation*** used.

Of course, for the purposes of this book, we are more interested in systematic experiments than the simple, everyday ones that are described. In medicine, these are called *clinical trials* or *randomized clinical trials* (RCTs; for more about RCTs, see More About box 1.1). In web design and market research, they may also be called *A/B testing*. The language is a little different from experiments in the social sciences, but the goal is the same—to discover what treatment (or cause) works best on a particular problem. In medicine, the problem is an illness or disease, the effect being a cure or improvement. In social science, the problems for which we are seeking solutions can be TV commercials that promote brand awareness, strategies for teaching students with Down syndrome, or interventions that help accountants be more honest.

To do this, social scientists use experiments as "the basic language of proof."[6] Experiments give us evidence of cause and effect by demonstrating what happens when something is changed while everything else remains the same. In this way, we have more assurance that the thing that changed is responsible for causing the outcome or effect we have observed.

Experiments are the most common kind of research conducted in the medical field, but in social science, they can be among the least often used. In communication journals, around 12% of studies use experimental designs.[7] In international relations, it can be as low as 4%.[8] In special education, experiments are promoted as the answer to calls for increased quality and rigor in an evidence-based profession.[9] And they are growing steadily in political science,[10] among other fields. As the benefits of this method become known throughout the social sciences, software to analyze data gets easier to use, and as technology makes subjects cheaper and easier to recruit, the use of experiments should

MORE ABOUT . . . BOX 1.1
Randomized Clinical Trials

RCTs are basically the same things as true experiments, where subjects are randomly assigned (the R) to either a treatment or a control group and given one or more experimental procedures or drugs. They are also referred to as randomized *controlled* trials, with language purists using controlled for studies that include a **control group** where subjects receive a placebo or no treatment. When *clinical* is used, it may or may not have a control group. Presumably, the clinical usage came into being because these studies were conducted on medical patients or others in a clinical setting. The term *trial* is used because the treatment or drug being studied is being tried out—that is, it is not approved for widespread use, and the study will determine if it is safe and effective.[1]

RCTs are considered the gold standard for medical studies, just as their counterpart is in social science—the true or lab experiment.

The history of RCTs dates back to 600 BC when Daniel of Judah compared the royal Babylonian diet to a vegetarian diet.[2] Others credit James Lind, who conducted the scurvy experiments in 1747 described in chapter 2.[3] The first modern-day RCT is usually recognized as the test of streptomycin's effects on tuberculosis in 1948.[4] The study begins by explaining how the preponderance of inadequately controlled clinical trials on tuberculosis had led to "exaggerated claims" about gold as a treatment. The study presents a "full description" of the methods because of the difficulty of planning such a rigorous trial, so that others may reproduce it.[5]

only increase. Thus, knowing how to do them well is all the more important. Already, many articles discuss the increased use of experiments for both academics and professionals in disciplines such as political science[11] and information systems.[12] Others discuss the growing importance of creating organized courses in experimental methods.[13] Some go so far as to say experiments are the most important method in their discipline.[14] Experiments are central to fields that use evidence-based practice such as social work[15] and education, where the randomized clinical trial is the gold standard.[16]

EXPERIMENTS COMPARED TO OTHER METHODS

The main benefit of the experimental method is that it offers a powerful tool to discover causation. Many other research methods, such as surveys, can identify **correlations,** or relationships, that vary together and are unlikely to have occurred by chance. This is not

necessarily the same thing as causation. Experiments, by contrast, can provide insights into how changing one thing leads to changes in another. Deliberately and systematically **varying**, or changing, something allows us to see a potential causal agent.[17] Following up correlational studies with experiments is a good way to give us more confidence that the relationships we find in these studies are actually causal. **Triangulation**, or the use of different methods to study the same phenomenon, is important to scientific inquiry because it helps give us confidence that what we see using one method can be **replicated**, or reproduced, using another. Of much concern is the overreliance on surveys and observational methods that show a correlation but with no follow-up using experiments that find evidence of a causal relationship.[18]

One of the most famous illustrations of the problem with inferring causation from correlation is the long-ago pronouncement that storks brought babies. A study done in Copenhagen in the 1930s documented that the years with larger stork populations also saw more babies born—a high correlation of .85.[19] But just because these two variables are highly correlated does not mean one caused the other. Instead, there were **plausible alternative explanations**, or other possible causes, that were not studied. This was right after World War I, and all those soldiers returning home after so long led to more babies being born. In addition, people were migrating from the country to the city where the jobs were, so more people to have babies equaled more babies. As the population increased, more houses were built, which led to more places for storks to nest, leading to more storks.[20] The cycle continued.

iStock.com/zuzulicea

BASIC CRITERIA FOR EXPERIMENTS

This example illustrates the three basic features that all experiments must have; that is, the cause must precede the effect—in this case, storks did come before the babies—but the cause must also be related to the effect, which it was not. There was no logical or theoretical reason and no empirical evidence to suggest that storks were related to babies. This is one reason why experimentalists do not test mere hunches but instead find linkages in the form of theoretical, logical, or existing evidence to test. This helps ensure that the cause is actually related to the effect. In the storks and babies example, the third feature of an experiment also was missing—there must be no other plausible alternative explanation for the effect. A real experiment must contain all three: Cause must come before the effect, be related to it, and there must not be any other plausible alternative explanations. If these three conditions are met, experiments give us a powerful way to have more

STUDY SPOTLIGHT 1.2

Discovering Effects and Explaining Why

SAGE Journal Article:
study.sagepub.com/
coleman

Hay, Carter, Xia Wang, Emily Ciaravolo, and Ryan C. Meldrum. 2015. "Inside the Black Box: Identifying the Variables That Mediate the Effects of an Experimental Intervention for Adolescents." *Crime & Deliquency* 61 (2): 243–270.

This study is a good example of an experiment that not only finds effects but also explains what caused them. The authors start by reconfirming that the treatment by a particular program actually did reduce juveniles' risk for delinquency. Then they examined the mediating variables that intervene, or come between, participation in the program and reduced juvenile delinquency. They examined a total of eleven risk factors that had been suggested as the reasons why the program worked, which researchers call *causal mechanisms*. They say, "In short, if we lack insight on the precise mechanisms by which a program reduces delinquency, then efforts to build on its strengths, replicate it elsewhere, and use it to inform public policy are necessarily hindered" (p. 248).

Out of the eleven possible variables that could have explained why the program worked, the researchers found only one that was significant: "Reduced association with peers who engaged in deviance and pressured them to do so as well" (p. 263). In other words, hanging out with a bad crowd.

They explain the importance of this in terms of the time and money spent by such programs pursuing these ten other variables that, it turns out, did not actually make a difference. Put in nicer terms than the wasting of time and money, the researchers say, "Our analysis—a rare test that has considered mediating variables—suggests that many of the risk factors targeted by these programs may be unresponsive to program services" (p. 264).

confidence that one thing led to, or caused, another, not just that there is an association that was unlikely to have occurred by chance.

While being able to test cause and effect is a powerful tool, it is still merely a description; it does not explain why something occurred. For example, in the field of social work, many studies have identified programs that successfully reduce juvenile delinquency, but few have undertaken an examination of why they are effective.[21] Hay and colleagues noticed the gap and designed an experiment to discover the mechanisms that explain why participating in a program reduced delinquency rates (for more on this study, see Study Spotlight 1.2). In my own work, I hypothesized that seeing photographs would cause journalists to use better-quality ethical reasoning when making news decisions. I found the effect I was looking for,[22] but that was a description, not an explanation. That finding alone did not say anything about why it happened. To provide a causal explanation, experiments need to build in potential mechanisms that explain these effects. If photographs do improve ethical reasoning, it is important to know why. **Mediators** and **moderators** such as this will be discussed in more detail in another chapter. Suffice it

to say here that good experiments should also include mechanisms for explaining why a certain effect occurred.

With this basic discussion of what experiments can do compared to other methods, next we turn to some specific elements of experiments. All of these will be elaborated in more detail in later chapters but are introduced here to provide some basic fundamental understanding of experiments.

ELEMENTS OF EXPERIMENTS

Variation

Varying or changing things is the first key to a good experiment. Obviously, if nothing varies, then there is nothing to study. Something has to change. In experiments, variation is achieved by **manipulating** the **independent variable**, or IV—the thing researchers think will cause a change. This is also called the *manipulation, treatment,* or *intervention*. The researcher should carefully control these. For example, many studies use professional actors to portray a political candidate rather than real politicians who may bias subjects' responses.[23] In one study, the actor was tasked with displaying positive nonverbal behaviors in one interview, negative ones in another, and neutral body language in a third. The study authors researched exactly what those looked liked—for example, crossed arms was negative, leaning forward and looking the interviewer in the eye was positive—and made sure the actor displayed those behaviors. It is not enough to merely ask an actor to behave positively or negatively but for the researchers to know exactly what that should look like and ensure that it is properly demonstrated.

Another key to good variation is that only one thing at a time can vary; otherwise, it is impossible to tell which of the several things that varied caused the outcome. Actually, many experiments do vary more than one thing at a time (more on that later), but the key is they do not allow those things to **covary**—that is, they cannot vary together. For example, one researcher noticed that studies showing negative political advertising were more powerful than positive advertising and had allowed the tone of the ads to covary with the amount of information in them.[24] There was always more information in the negative ads than in the positive ads; therefore, the more powerful effects could be due to more information, not necessarily to its negative tone. This provided a plausible alternative explanation for the effects found in these other studies. To determine if this was the case, the researcher did a study that held the amount of information **constant**—that is, there was the same amount of information in both the negative- and positive-toned ads. He found that the tone was not causing the effect at all—it was the amount of information;

negative advertising typically had more information in it than positive advertising, and that was the source of the greater levels of campaign knowledge, interest, and turnout that the other studies had documented.

In reality, cause is rarely **univariate**, or caused by one variable. Social science researchers seldom expect one thing alone to be the direct cause of another. There are usually many things that are responsible. In experiments, this is managed by holding constant other things than the one purposely being varied to make sure those things are not responsible for the effect.

Confounds

Things that could provide plausible alternative explanations are called **confounds**. It is important that there are no confounds—that is, anything that would harm the accuracy of the experiment. Key to a successful experiment is controlling for extraneous influences that might have caused the outcome. For example, a researcher studying the effect of photographs that were placed above or below the fold in a newspaper used as stimuli a vertical photo that showed a person up close above the fold and a horizontal photo showing people far away below the fold. He used *only* the horizontal far-away photo below the fold and *only* the vertical close-up photo above the fold. The problem with this is that it could have been the vertical or horizontal format of the photo, or the close-up or far-away distance of the people in them, that was responsible for the effect he found. Those were not the variables he had deliberately varied. The pictures varied on many levels, not just the ones he was interested in, which was placement in the newspaper. This represents a confound—a plausible alternative explanation that was not controlled for.

To avoid confounds in one study designed to determine how the race of the people in photographs affected journalists' ethical judgment, the researcher took the exact same picture and digitally altered the skin tone, hair, and facial features of the people in order to manipulate their race. That way, everything was the same except race—the thing designed to vary. The backgrounds were the same, the distance from the camera, the people's attractiveness, and everything else was the same. This avoided any confounds. Researchers know, for example, that close-ups make people feel more comfortable with the people in the photographs than long-shots,[25] and that attractive people are evaluated better on a variety of characteristics than less attractive people, including trustworthiness and electability in political candidates.[26] So this study needed photographs that were exactly the same in every way except the race of the people in them.

In experiments, it is helpful to be as Type A as possible. For example, a researcher might measure the distance from the TV to the chair that subjects sit in before every new

subject comes in. All subjects should be exactly the same distance from the TV because being even a little bit closer can make a difference.

Researchers discover potential confounds two ways: by using common sense and by reading other studies. The literature abounds with evidence showing, for example, that political party and ideology affect many things.[27] Thus, many experiments control for these potential confounds by creating stimuli that do not implicitly or explicitly state the political party of a fictitious candidate. This represents an **experimental control**, or carefully managing the variables in play.[28] Another example of avoiding potential confounds is not using real issues in the news at the time of an experiment held during an election.[29] When a researcher cannot control for a potentially spurious relationship between the independent and dependent variables, another approach is to measure it and **statistically control** for it—that is, take it out of the equation before examining what effect the manipulation had. Measuring people's political party and ideology, and then using them as covariates, is a form of statistical control.[i]

It is important to read the literature to discover things that need controlling. For example, in studies of moral development, researchers have found a connection between being liberal or conservative and quality of moral judgment.[30] This is not something that is intuitively obvious, so reading the literature to discover this is key. Age and education are also related to better moral judgment,[31] so experiments that use moral judgment as the outcome or **dependent variable** typically measure participants' political ideology in order to incorporate them in the statistical analyses as covariates should random assignment not have already made the groups equivalent on that variable. Covariates work by taking out the effect of the potentially confounding variable so researchers can see the true effect of whatever variable is being manipulated. It is important to know the literature in the domain being studied, because not everything one might suspect as having a potentially spurious relationship between a dependent and independent variable really does. For example, gender does not usually matter in moral judgment,[32] but it does in a great many other things. As most effects are rarely caused by simply one thing, it is important to know what else might be affecting the outcome.

It is also important to control for potential confounds that your own common sense tells you might affect the outcome. Just because you do not see something in the literature, if you think it might cause an effect, build it into the experiment so you can

[i]Covariates are not always necessary in experiments the way they are in observational studies because random assignment is designed to eliminate the effects of such variables. Random assignment will be discussed in more detail in chapter 7.

test it and see. Presumably you are studying a phenomenon that you know a lot about. If you are an expert, you should have some idea of what kinds of things affect the phenomenon you are studying. This is one way that knowledge is created and the literature develops.

Control Groups

Another important requirement of experiments is not just knowing what happens to people who receive the treatment, but also knowing what would happen if they had *not* received the treatment. Having a group of people who do not get the treatment, known as the *control group*, gives us a way to infer what the outcome would be had there been no treatment. The effect is the difference between what *did* happen when people got the treatment and what *would have* happened had they not gotten it. This allows us to isolate and test the effects of one variable at a time, and provides greater certainty that the effect is due to that variable and not something else. The people who are exposed to the treatment or intervention are called the *treatment group, experimental group,* or *manipulation group*. The people who are not exposed to the treatment are the control group. This group of people who are not given the manipulation is used as a comparison for what "normal," "neutral," or "no manipulation" would be like.

In medical research, the control group is often given a placebo, which looks like a pill, injection, therapy, or some other treatment but really is only a sugar pill, saline injection, or something else that leads people to perceive they had something done to them. (See More About box 1.3 on placebos.) Because of the power of suggestion, it is not good enough to actually do nothing to the people in the control group; they must perceive that they had some sort of treatment. For social scientists, that can sometimes present a problem. For example, when testing different types of marketing messages, what constitutes a control group? No message, of course, but then what do subjects do instead? You cannot have them just come in and answer the questionnaire without having them be exposed to something or they will perceive the study as too artificial. Control groups will be discussed more in the coming chapters; the point of the discussion here is that experiments need some way to compare what happens to people who get the treatment versus what happens to people who do not.

Assignment

Last in this discussion, but probably most important to experiments, is the topic of how people are assigned to the different interventions in an experiment. This topic is so important that it rates its own chapter (chapter 7). Briefly, the gold standard is

MORE ABOUT . . . BOX 1.3
Placebos

In medical research, the use of placebos is much more complicated than a simple injection of saline, a sugar pill, or the pretense of surgery to mimic an actual treatment on a group of control subjects. In fact, these inert or ineffective but harmless treatments actually have been shown to have effects. Called the *placebo effect*, people receiving them actually report improvement in whatever condition they were supposedly being treated for.

The word *placebo* is Latin for "I shall please."[33] The first controlled trial to use placebos has been traced to 1801.[34] The perception of placebos as fraudulent, deceptive, and unethical arose because up until the middle of the twentieth century, many practicing physicians would administer them to patients under the guise of actual medicine.[35] Questions about the ethicality of using placebos in research continue to this day.[36]

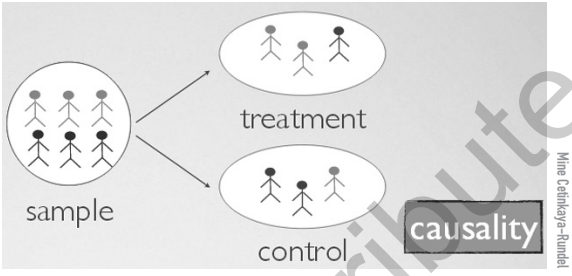
The first report on the placebo effect found that four out of five patients receiving the placebo reported relief of their symptoms; results were the same when they received the actual treatment.[37] The report cites the power of hope, faith, and the imagination. Ever since, research on the placebo effect has continued to show that patients improve after receiving the inert placebo to varying degrees, including equal or better results than the active drug.[38] A review of fifteen studies showed that, on average, placebos performed as well as the active treatment 35% of the time.[39] Flaws in this review have been pointed out, but research continues to show that placebos "work" about a third of the time.[40] In 1961, Beecher conducted another review and found a placebo effect in 37% of patients. [41] The term placebo effect has been defined as the positive effects of an inert substance due to the power of suggestion,[42] or "the difference in outcome between a placebo treated group and an untreated control group in an unbiased experiment."[43]

In social science research, a placebo is akin to a control group that receives something designed to look like a treatment or manipulation[44]—for example, subjects who read stories about a new movie rather than the episodic and thematically framed crime stories that are the real focus of the study.

A placebo is the inert or fake treatment given to subjects; the placebo effect is subjects' response to it. It is different from the Hawthorne effect, described in the More About box in chapter 2, which describes how subjects' performance changes because they are being observed. In the placebo effect, subjects experience change because of their expectations, beliefs, or hopes that the treatment will work rather than the treatment itself.[45] In order to be effective, a treatment has to pass the placebo test—that is, the treatment must be significantly better on the outcome than the control group that gets the placebo.

Research in the social sciences also finds a placebo effect—for example, with products that promise better athletic performance,[46] factors that affect financial decisions,[47] drink labels and perceptions of intoxication,[48] and well-known brands.[49] Placebos are also found in daily life—for example, the buttons at intersections that lead pedestrians to feel a sense of control but don't really affect the light.[50]

random assignment, which is a way of placing subjects[ii] into the groups in such a way that individual differences are evenly distributed across the different groups. That feature is important to ensure that people's individual characteristics are not confounds of the study. Men and women should be evenly distributed across the groups, as should the young and old, for example. This book devotes an entire chapter to assigning subjects to conditions of experiments, including discussing what to do when it is not possible to assign people this way.



STARTING A STUDY OF YOUR OWN

Now that we have some basic understanding of what an experiment can do compared to other methods, this chapter turns to the first practical step in starting your own experiment: writing a clear and concise statement of the research problem. Clear and concise writing is important in all kinds of research but especially, in my experience, when writing up experiments. Because they can be complex, and many social scientists may not be as familiar with experimental methodology as they are with surveys or other techniques, experiments seem to be harder for readers to follow. Thus, one key focus of this book is to help readers write experiments in plain language and terms that anyone with moderate knowledge of social science can understand. That can be harder than you think. We begin with the statement of the research problem. This is more overarching than specific **hypotheses**, although they are related. Writing hypotheses will be developed in a later chapter (chapter 3). To begin a new experiment, one must come up with an idea to test. It helps to think about it in terms of cause and effect, testing if one thing causes another. For example, here are three (rather oversimplified) ideas for studies:

[ii]This textbook uses the term *subjects* to refer to the people studied in an experiment. The latest (6th) edition of the *Publication Manual of the American Psychological Association* (APA) says both terms, *subjects* and *participants*, are appropriate, noting that "subjects" has been in use for hundreds of years. The history of objections to "subjects" began in 1994 with the fourth edition of the APA manual, when "participants" was preferred. When the current edition appeared in 2010, subjects and participants were on equal footing. The entry on page 73 says, "Indeed, for more than 100 years the term subjects has been used within experimental psychology as a general starting point for describing a sample, and its use is appropriate." Because subjects is as appropriate as participants, this text uses subjects in order to maintain consistency with other terms in experimental design language, including *between-subjects designs, within-subjects designs,* and *human subjects* used by IRBs. For more on this topic, read the essay "What Should They Be Called" by Roddy Roediger in the APS Observer, April 2004, 17, no. 4, available at http://journal.sjdm.org/roediger.html.

1. Does seeing a photograph improve ethical reasoning?
2. Does voice pitch affect the credibility of a radio announcer?
3. Does the height of female politicians affect voters' assessments of their qualifications and the likelihood of voting for them?

In the first study, the treatment is showing subjects a photograph versus no photograph in the control condition. In the second study, the voice of a radio announcer was pitched high in one treatment condition and low in another versus normal in the control condition. In the third, pictures of short and tall women were the two treatment groups versus average-height women as the control. Of course, all three experiments ended up being more complex than this, but this simple "does A cause B" approach was the genesis of all the studies.

Mine Cetinkaya-Rundel

Writing a Statement of the Problem

The next step is to write a clear and focused statement of the problem to be studied. A good way to start this sentence is with "The purpose of this study is . . ." or something similar. Here are some examples of good, clear, and focused statements of the problem from experiments:

- "This article investigates how media use of the microblogging tool Twitter affects perceptions of the issue covered and the credibility of the information."[51]
- "We experimentally study the common wisdom that money buys political influence."[52]
- "We assess the extent to which communication setting (i.e., face-to-face versus online chat room discussion) affects individuals' willingness to express opinions."[53]
- "A first question is whether providing general information on the welfare properties of prices and markets modify attitudes toward repugnant trades."[54]

- ". . . the goal of this study is to demonstrate that moral convictions and moral judgments in politics are causally affected by harm associations and moral emotions."[55]

All of these examples have three things in common: They say what the intervention, manipulation, or cause is, and what the effect or outcome is, usually in that order. And they have a verb. In the first example, Twitter is the intervention or cause, and the effect or outcomes are perceptions of the issue and credibility. In the next one, money is the cause, and political influence is the effect. In the last example, the order is reversed, with the effect—moral convictions and judgments—listed first, and the causes—harm associations and moral emotions—given last. You may also recognize these as independent variables (the cause, interventions, or manipulations) and dependent variables (the effect or outcomes). The third thing in common is that there is a verb in each of these, some word that describes the action the cause is expected to have on the outcome, or what one thing is expected to do to the other. In these examples, the verbs are "affects," "buys," "modify," and "causally affected." For experiments, it is advisable to stay away from less precise words such as "explore," "understand," and "examine," and instead to use more specific, causal language such as this. Other good words to use include "differs," "improves," and similar words. For example, a statement might say, "The purpose of this study is to test the idea that photographs *improve* ethical reasoning," or "The purpose of this study is to see if assessments of credibility *differ* with the pitch of a radio announcer's voice."

To write a clear and focused statement of the research problem for an experiment, first determine what the cause (or intervention, manipulation, independent variable) is, then say what it is expected to do (differ, affect, modify, cause—the verb) to some outcome (the effect or dependent variable). Here, I offer my fill-in-the-blanks template for writing a statement of the research problem:

"The purpose of this study is to see how (one thing, insert the cause, intervention, manipulation, treatment, or independent variable) (does something, insert a verb—differs, affects, modifies, causes, changes, etc.) *to* (something else, insert effect, outcome, dependent variable)."

Some of the terms in these examples, such as *harm associations*, might not be familiar if they are outside your discipline, so it helps to examine some experiments in your own field for other examples that may be more commonplace. Also notice how all of these are

only one sentence long. Writing such a clear, focused, and easily understandable statement that describes the experiment in one sentence is not easy. It is not surprising to have to write, rewrite, and edit it many times. Have someone familiar with the discipline read it and see if he or she can understand it. As the project evolves, you may have to rewrite this statement, maybe even a few times.

Many experiments have more than one purpose, so two or three statements of the problem may be written. If that is the case, put them together in the paper and link them with phrases like "This study also seeks to . . ." For example: "This study tests whether photographs improve ethical reasoning. If this effect is found, this study also seeks to determine the causal mechanism for this improvement." This way, one has the primary purpose of the experiment followed by additional purposes all in one place rather than spread out around the paper. Readers will appreciate having everything the study intends to do all listed together rather than reading a slowly evolving purpose of the study, waiting for it to unfold like a murder mystery.

Once satisfied with the statement of the problem, authors will need to remind readers periodically throughout the paper what the mission is, so be sure that the statement of the problem stays consistent. One frequent problem I see when reviewing experiments for journals is how the stated goal of the study changes as the paper progresses. Finally, a good practice is to write the statement of the problem on a sticky note and paste it on your computer where you can see it as you work. This will help keep you focused and consistent.

Answering the "So What" Question

After formulating this clear and focused statement, the next step is to articulate why this study is important. This is commonly referred to as answering the "so what" question. Some journals even have a highlighted box that is devoted to this—for example, see the "Significance" box in *Proceedings of the National Academy of Sciences*.[56] This is another area where researchers frequently think the importance of their study is obvious and should not have to point it out; in fact, this is one of the most crucial aspects of a study. For this task, I tell students they need to state the obvious. It may be abundantly clear to you, but it will not necessarily seem so to other readers. To do this, think about the reason your study is important on three different levels: (1) to other academics, (2) to professionals in the field, and (3) to society or people in general. For the first, you can point out some gap in knowledge or some obstacle that the study overcomes. It should

also contribute to theory in some way or help uncover any of the mechanisms or reasons for some phenomenon. One reason that is never acceptable alone is because a study has never been done before. That is a start, but you should always go further to say why it is important that it be done beyond never having been done before; otherwise, perhaps it has never been done before for a good reason.

Frequently, one sees researchers point out the importance to other scholars but overlook the other two groups: professionals and the public. For any kind of study, being able to articulate why your findings are meaningful beyond academia is crucial to research that makes a difference. R. Barker Bausell's book *Conducting Meaningful Experiments*[57] is predicated on that premise. For the second step in considering the "so what" question, ask if your study will be of interest to those in the professional arm of your discipline. Will it help accountants, campaign managers, teachers, public relations professionals, or anyone else do their jobs better? Will the findings of the study give professionals more insight into their own subconscious decision making? Will it help them overcome some obstacle or give them evidence they need to change the way they practice their craft? Perhaps it will show them which of their efforts are paying off in the outcome they desire and which are not, whether that is more engaged citizens or more customers. Making this kind of concerted effort to solve real-world problems helps bridge the gap between scholars and the profession they serve.

Finally, being able to say why all of society will benefit results in more meaningful science. Too often, the public considers academic researchers to be "eggheads in ivory towers" writing about things that have no basis in reality in order to get another publication. There are even awards that make the news for the most wasteful research.[58] Having one's research ridiculed in this way might be avoided if studies better articulated why seemingly silly or obvious findings are important to someone other than our scholarly colleagues. This kind of attention does nothing to help advance the cause of research or increase funding for it—conducting research that is meaningful to ordinary people does. So answer the "so what" question; will it help anyone or improve any social ills? Not every study will broker world peace, but it might help lessen racial profiling, change a morally repugnant practice, or give policy makers information needed to pass a law. Not every study articulates all three, but thinking through the benefits to these different publics can help experimentalists design studies that are truly meaningful. For some examples of statements that answer the "so what" question, see How To Do It box 1.4.

HOW TO DO IT 1.4

Examples of Answers to the "So What" Question

From: Neil, Nicole, and Emily A. Jones. 2015. "Studying Treatment Intensity: Lessons from Two Preliminary Studies." *Journal of Behavioral Education* 24: 51–73.

"There is only a recent and small literature examining treatment intensity, and the research on treatment intensity focused on specific disorders is even more limited. It may be that etiology and characteristics associated with specific etiologies impact the effects of intervention intensity. Many children with Down syndrome display poor task persistence and inconsistent motivational orientation . . . For some learners with Down syndrome, it is possible that there is an optimum moderate level of intensity, past which learners engage in greater levels of escape-motivated problem behavior and there are diminishing gains in acquisition rates."[59]

From: Coleman, Renita. 2011. "Color Blind: Race and the Ethical Reasoning of African Americans on Journalism Dilemmas." *Journalism and Mass Communication Quarterly* 88 (2) (Summer): 337–351.

"This study is of value because it provides important information to evaluate one of the solutions offered to the problem of stereotypical media portrayals—hiring and promoting more minority journalists. Newsrooms across the country are staffed primarily by whites; incorporating more minority viewpoints should lead to more equal coverage of minorities, according to the argument. . . . it is important to examine whether minority journalists do in fact exhibit more tolerant attitudes toward minorities in their cognitive processing. To date, there is no empirical evidence that black journalists have more favorable perceptions of blacks in the news. . . . This study also fills that void by exploring how race influences the ethical reasoning of blacks when blacks and whites are in news stories."[60]

From: Aday, Sean. 2006. "The Framesetting Effects of News: An Experimental Test of Advocacy Versus Objectivist Frames." *Journalism and Mass Communication Quarterly* 83 (4) (Winter): 767–784.

"Specifically, we still do not know enough about why effects are found in some cases and not others, and too little work has been done exploring the cognitive basis of the effect that would allow us to develop a theory for when and why some attributes would have a second-level effect and others would not."[61]

From: Elias, Julio J., Nicola Lacetera, and Mario Macis. 2015. "Markets and Morals: An Experimental Survey Study." *PLoS ONE* 10 (6) (June 1): 1–13. Public Library of Science.

"Prohibiting some of these transactions has costs. Life insurance contracts, for instance, were once illegal because they were seen as gambles against God; they now create value for millions of people, and are viewed as a form of 'institutionalized altruism.' Similarly, the idea of an all-volunteer paid army was long rejected in the United States, despite arguments showing its efficiency. The prohibition of payments to people who give their organs contributes to the growing gap between organ demand and supply. Banning some trades may also lead to the formation of illegal markets, which, in turn, entail further costs such as violence . . ."[62]

From: Grober, Jens, Ernesto Reuben, and Agnieszka Tymula. 2013. "Political Quid Pro Quo Agreements: An Experimental Study." ***American Journal of Political Science*** **57 (3) (July): 582–597.**

"There are good reasons to suspect there is some truth behind the common belief that money in politics is undesirable. First, in spite of being banned, political quid pro quo can occur outside publicly observable channels. Second, for economically powerful special interests, most of which are large corporate firms, giving as an investment that increases profits is a more plausible explanation than political participation. Moreover, returned favors to such interests, such as specific tax breaks, subsidies, and regulations, can be easily concealed as an economic necessity and are therefore hard to quantify. Third, collusion between major candidates may also take the form of an agreement on a common view with regard to a given political issue . . . Finally, even if the impact of money in politics is overestimated by the public, this belief can affect the public's political trust and behavior."[63]

Common Mistakes

- Not clearly stating the purpose of the research, and not keeping it consistent throughout the paper
- Not putting all the things the study intends to do in one place, but having the goals of the study unfold slowly through the article
- Not putting the statement of purpose up high in the paper, in the introduction and before page 3
- Failing to say why the study is important to theory, other researchers, the profession, and regular people. "Because it has never been done" is not a reason why a study is important by itself.

Test Your Knowledge

1. Experiments need to show that the cause precedes the effect but not that it is necessarily related to it. If there is a statistically significant relationship between two things, that is all that matters.
   a. True
   b. False
2. A researcher studied the effects of attractiveness on how well students liked a teacher. The attractive teacher was twenty-five years old; the unattractive teacher was forty-five years old. The problem here is that:

a. Age is confounded with attractiveness
b. It is hard to define attractiveness
c. The cause did not precede the effect
d. The researcher did not control for political ideology

3. You assign fifteen employees to go to a one-day seminar on stress management. Another fifteen are assigned to a one-week seminar. At the end of the month, you measure each employee's perceived level of stress. What is the treatment or manipulation in this study?
   a. How stressed out employees are
   b. How long the seminar is
   c. How you chose the thirty employees
   d. The quality of the stress management teacher

4. A study measures students' arousal level before they take a test. It finds that as arousal increases, performance decreases. This finding shows:
   a. Causality
   b. Correlation
   c. A plausible alternative explanation
   d. A confound

5. Which of the following is NOT one of the three basic criteria for an experiment?
   a. Cause must precede the effect
   b. The effect must be unlikely to have occurred by chance
   c. Cause must be related to the effect
   d. There are no plausible alternative explanations for the effect

6. Variation is achieved by:
   a. Holding everything constant
   b. Using demographics as covariates
   c. Systematically changing something
   d. Having a control group

7. Things that could provide plausible alternative explanations are called:
   a. Covariates
   b. Confounds
   c. Independent variables
   d. Causal mechanisms

8. In experiments, control groups serve the purpose of:
   a. Allowing us to know what happens to people who receive the treatment
   b. Allowing us to generalize to more people
   c. Allowing us to know what would happen to subjects if they had not received the treatment
   d. Allowing us to say some effect occurred with a specific degree of certainty
9. In an experiment, "assignment" is:
   a. The task that subjects must complete
   b. How authorship order is calculated for the paper
   c. The way researchers ensure subjects believe the experiment is real
   d. How subjects are put in the different interventions or control group
10. The gold standard in experiments is to assign subjects:
   a. Representatively
   b. Purposively
   c. Randomly
   d. Haphazardly

Answers:

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

Application Exercises

1. Use scholar.google.com or your school library's database to find studies that use experimental designs in your discipline. Include the word *experiment, experimental design,* or *controlled experiment* in the search terms. Read three of the experiments that interest you most and look for the concepts covered here. Specifically, identify the treatment group or groups. Is there a control group? If so, what is used to represent "no treatment"? Identify the statement of the problem and the answer to the "so what" question.
2. Think up three distinct studies that you would like to do with an experiment. That is, something should be manipulated or changed in order to see what effect it has on some outcome. Write a clear and focused statement of the problem. Explain why it is important to academics, the profession, and the world at large (the "so what" question). Use 250 words for each. These should not be straight replications but new ideas, or they may be replications with substantial extensions to the study you are replicating.

Suggested Readings

The introduction and Chapters 1 and 2 of Bausell, R. B. 1994. *Conducting Meaningful Experiments: 40 Steps to Becoming a Scientist*. Thousand Oaks, CA: Sage.

Chapter 1 in Shadish, W. R., T. D. Cook, and D. T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Belmont, CA: Wadsworth Cengage Learning.

Thorson, Esther, Robert H. Wicks, and Glenn Leshner. 2012. "Experimental Methodology in Journalism and Mass Communication Research." *Journalism and Mass Communication Quarterly* 89 (1): 112–124.

Notes

1. Markus MacGill, "What Is a Randomized Controlled Trial in Medical Research?" *Medical News Today*, http://www.medicalnewstoday.com/articles/280574.php. Accessed February 10, 2017.
2. Harald O. Stolberg, Geoffrey Norman, and Isabelle Trop, "Fundamentals of Clinical Research for Radiologists: Randomized Controlled Trials," *American Journal of Roentgenology* 183, no. 6 (2004): 1539–1544.
3. Marcia L. Meldrum, "A Brief History of the Randomized Controlled Trial: From Oranges and Lemons to the Gold Standard," *Hematology/Oncology Clinics of North America* 14, no. 4 (2000): 745–760.
4. S. Shikata et al., "Comparison of Effects in Randomized Controlled Trials with Observational Studies in Digestive Surgery," *Annals of Surgery* 244, no. 5 (2006): 668–676; Stolberg, Norman, and Trop, "Fundamentals of Clinical Research for Radiologists."
5. Streptomycin in Tuberculosis Trials Committee, "Streptomycin Treatment of Pulmonary Tuberculosis: A Medical Research Council Investigation," *British Medical Journal* 2, no. 4582 (1948): 769–782.
6. D. T. Campbell and J. C. Stanley, *Experimental and Quasi-Experimental Designs for Research* (Chicago: Rand McNally, 1963), 3.
7. Esther Thorson, Robert H. Wicks, and Glenn Leshner, "Experimental Methodology in Journalism and Mass Communication Research," *Journalism and Mass Communication Quarterly* 89, no. 1 (2012): 112–124.
8. R. McDermott, "New Directions for Experimental Work in International Relations," *International Studies Quarterly* 55, no. 2 (2011): 503–520.
9. Matthew C. Makel et al., "Replication of Special Education Research: Necessary but Far Too Rare," *Remedial and Special Education* 37, no. 3 (2016): 1–8.
10. Rebecca B. Morton and Kenneth C. Williams, *Experimental Political Science and the Study of Causality: From Nature to the Lab* (New York: Cambridge University Press, 2010): 5–6, 16–17; Gerry Stoker, "Exploring the Promise of Experimentation in Political Science: Micro-Foundational Insights and Policy Relevance," *Political Studies* 58 (2010): 300–319; Cengiz Erisen, Elif Erisen, and Binnur Ozkececi-Taner, "Research Methods in Political Psychology," *Turkish Studies* 13, no. 1 (2013): 13–33; James N. Druckman et al., "The Growth and Development of Experimental Research in Political Science," *American Political Science Review* 100, no. 4 (2006): 627–635.
11. Ulrich Hamenstädt, "Teaching Experimental Political Science: Experiences from a Seminar on Methods," *European Political Science* 11, no. 1 (2012): 114–127; Erisen, Erisen, and Ozkececi-Taner, "Research Methods in Political Psychology"; James N. Druckman and Arthur Lipia, "Experimenting with Politics," *Science* 335, no. 6 (March 9, 2012): 1177–1179; Luke Keele, Corrine McConnaughy, and Ismail White, "Strengthening the Experimenter's Toolbox: Statistical Estimation of Internal Validity," *American Journal of Political Science* 56, no. 2 (2012): 484–499; Sharon Crasnow, "Natural Experiments and Pluralism in Political Science," *Philosophy of the Social Sciences* 45, no. 4/5 (2015): 424–441.
12. Yair Levy, Timothy J. Ellis, and Eli Cohen, "A Guide for Novice Researchers on Experimental and

Quasi-Experimental Studies in Information Systems Research," *Interdisciplinary Journal of Information, Knowledge and Management* 6 (2011): 151–161.

13. Susan D. Hyde, "Experiments in International Relations: Lab, Survey, and Field," *Annual Reviews of Political Science* 18 (2015): 403–424. Hamenstädt, "Teaching Experimental Political Science."
14. Erisen, Erisen, and Ozkececi-Taner, "Research Methods in Political Psychology."
15. William J. Reid, Bonnie Davis Kenaley, and Julanne Colvin, "Do Some Interventions Work Better Than Others? A Review of Comparative Social Work Experiments," *Social Work Research* 28, no. 2 (2004): 71–81.
16. Joshua B. Plavnick and Summer J. Ferreri, "Single-Case Experimental Designs in Educational Research: A Methodology for Causal Analyses in Teaching and Learning," *Education Psychology Review* 25 (2013): 549–569.
17. William R. Shadish, Thomas D. Cook, and Donald T. Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (Belmont, CA: Wadsworth Cengage Learning, 2002).
18. K. Imai, G. King, and E. Stuart, "Misunderstandings Between Experimentalists and Observationalists About Causal Inference," *Journal of the Royal Statistical Society* Series A 171 (2008): 481–502.
19. Gustav Fischer, "Ornithologische Monatsberichte," *Jahrgang,* 44, no. 2 (1936); Gustav Fischer, "Statistiches Jahrbuch Deutscher Gemeinden," *Jahrgang,* 48, no. 1 (1940): 27–33, http://pignottia.faculty.mjc.edu/math134/classnotes/storks.pdf
20. http://faculty.vassar.edu/lowry/ch3pt2.html
21. Carter Hay et al., "Inside the Black Box: Identifying the Variables That Mediate the Effects of an Experimental Intervention for Adolescents," *Crime & Deliquency* 61, no. 2 (2015): 243–270.
22. Renita Coleman, "The Effect of Visuals on Ethical Reasoning: What's a Photograph Worth to Journalists Making Moral Decisions?" *Journalism and Mass Communication Quarterly* 83, no. 4 (2006): 835–850.
23. H. Denis Wu and Renita Coleman, "The Affective Effect on Political Judgment: Comparing the Influences of Candidate Attributes and Issue Congruence," *Journalism and Mass Communication Quarterly* 91, no. 3 (2014): 530–543.
24. Daniel Stevens, "Tone Versus Information: Explaining the Impact of Negative Political Advertising," *Journal of Political Marketing* 11, no. 4 (2012): 322–352.
25. Arthur Asa Berger, "Semiotics and TV," in *Understanding Television: Essays on Television as a Social and Cultural Force*, ed. R. R. Adler (New York: Praeger, 1981): 91–114.
26. Zhao Na et al., "Face Attractiveness in Building Trust: Evidence From Measurement of Implicit and Explicit Responses," *Social Behavior and Personality: An International Journal* 43, no. 5 (2015): 855–866; Leslie A. Zebrowitz, Robert G. Franklin Jr., and Rocco Palumbo, "Ailing Voters Advance Attractive Congressional Candidates," *Evolutionary Psychology* 13, no. 1 (2015): 16–28.
27. Amy E. Lerman, Katherine T. McCabe, and Meredith L. Sadin, "Political Ideology, Skin Tone, and the Psychology of Candidate Evaluations," *Public Opinion Quarterly* 79, no. 1 (2015): 53–90; Sarah Reckhow, Matt Grossmann, and Benjamin C. Evans, "Policy Cues and Ideology in Attitudes toward Charter Schools," *Policy Studies Journal* 43, no. 2 (2015): 207–227; Jonathon Schuldt and Adam Pearson, "The Role of Race and Ethnicity in Climate Change Polarization: Evidence from a U.S. National Survey Experiment," *Climatic Change* 136, no. 3/4 (2016): 495–505.
28. Stoker, "Exploring the Promise of Experimentation in Political Science."
29. Kevin J. Mullinix et al., "The Generalizability of Survey Experiments," *Journal of Experimental Political Science* 2 (2015): 109–138.
30. J. R. Rest and Darcia Narvaez, eds., *Moral Development in the Professions: Psychology and Applied Ethics* (Hillsdale, NJ: Erlbaum, 1994).
31. James. R. Rest, "Morality," in *Handbook of Child Psychology, Vol. III Cognitive Development,*, ed. P. H. Mussen (New York: Wiley, 1983).
32. Stephen Thoma, "Estimating Gender Differences in the Comprehension and Preference of Moral Issues," *Developmental Review* 6, no. 2 (1986).
33. A. J. De Craen et al., "Placebos and Placebo Effects in Medicine: Historial Overview," *Journal of the Royal Society of Medicine* 92, no. 10 (1999): 511–515.
34. Ibid.
35. Ibid.; David B. Elliott, "The Placebo Effect: Is It Unethical to Use It or Unethical Not To?" *Ophthalmic & Physiological Optics* 36, no. 5 (2016): 513–518.

36. Elliott, "The Placebo Effect."
37. H. K. Beecher, "The Powerful Placebo," *JAMA* 159, no. 17 (1955): 1602–1606.
38. De Craen et al., "Placebos and Placebo Effects in Medicine."
39. Beecher, "The Powerful Placebo."
40. De Craen et al., "Placebos and Placebo Effects in Medicine."
41. H. K. Beecher, "Surgery as Placebo: A Quantitative Study of Bias," *JAMA* 176 (1961): 1102–1107.
42. Beecher, "The Powerful Placebo."
43. P. C. Goetzsche, "Is There Logic in the Placebo?" *Lancet* 344, no. 8927 (1994): 904.
44. Alan Bryman, "Placebo," in *The Sage Encyclopedia of Social Science Research Methods*, ed. Michael S. Lewis-Beck, Alan Bryman, and Tim Futing Liao (Thousand Oaks, CA: Sage, 2011): 825.
45. Ibid.
46. Aaron M. Garvey, Frank Germann, and Lisa E. Bolton, "Performance Brand Placebos: How Brands Improve Performance and Consumers Take the Credit," *Journal of Consumer Research* 42, no. 6 (2016): 931–951.
47. Michael A. Kuhn, Peter Kuhn, and Marie Claire Villeval, "Decision-Environment Effects on Intertemporal Financial Choices: How Relevant Are Resource-Depletion Models?" *Journal of Economic Behavior & Organization* 137 (2017): 72–89.
48. Yann Cornil, Pierre Chandon, and Aradhna Krishna, "Does Red Bull Give Wings to Vodka? Placebo Effects of Marketing Labels on Perceived Intoxication and Risky Attitudes and Behaviors," *Journal of Consumer Psychology (Elsevier Science)* 27, no. 4 (2017): 456–465.
49. Carlos Alberto Alves, Evandro Luiz Lopes, and José Mauro da Costa Hernandez, "It Makes Me Feel So Good: An Experimental Study of the Placebo Effect Generated by Brands," *Journal of International Consumer Marketing* 29, no. 4 (2017): 223–238.
50. Gwen Sharp, "The Placebo Effect," *The Society Pages,* March 10, 2011, https://thesocietypages.org/socimages/2011/03/10/the-placebo-effect/.
51. Mike Schmierbach and Anne Oeldorf-Hirsch, "A Little Bird Told Me, So I Didn't Believe It: Twitter, Credibility, and Issue Perceptions," *Communication Quarterly* 60, no. 3 (2012): 317–337.
52. Jens Grober, Ernesto Reuben, and Agnieszka Tymula, "Political Quid Pro Quo Agreements: An Experimental Study," *American Journal of Political Science* 57, no. 3 (2013): 582–597.
53. Shirley S. Ho and Douglas M. McLeod, "Social-Psychological Influences on Opinion Expression in Face-to-Face and Computer-Mediated Communication," *Communication Research* 35, no. 2 (2008): 190–207.
54. Julio J. Elias, Nicola Lacetera, and Mario Macis, "Markets and Morals: An Experimental Survey Study," *PLoS ONE* 10, no. 6 (2015): 1–13.
55. Pazit Ben-Nun Bloom, "Disgust, Harm and Morality in Politics," *Political Psychology* 35, no. 4 (2014): 495–513.
56. For an example, see William Minozzi et al., "Field Experiment Evidence of Substantive, Attributional, and Behavioral Persuasion by Members of Congress in Online Town Halls," *Proceedings of the National Academy of Sciences* 112, no. 13 (2015): 3937–3942.
57. R. Barker Bausell, *Conducting Meaningful Experiments: 40 Steps to Becoming a Scientist* (Thousand Oaks, CA: Sage, 1994).
58. FoxNews.com, "Senate Report Finds Billions in Waste on Science Foundation Studies," Foxnews.com, http://www.foxnews.com/politics/2011/05/26/senate-report-finds-billions-waste-science-foundation-studies/.
59. Nicole Neil and Emily A. Jones, "Studying Treatment Intensity: Lessons From Two Preliminary Studies," *Journal of Behavioral Education* 24 (2015): 51–73.
60. Renita Coleman, "Color Blind: Race and the Ethical Reasoning of African Americans on Journalism Dilemmas," *Journalism and Mass Communication Quarterly,* 88, no. 2 (Summer 2011): 337–351.
61. Sean Aday, "The Framesetting Effects of News: An Experimental Test of Advocacy Versus Objectivist Frames," *Journalism and Mass Communication Quarterly* 83, no. 4 (2006): 767–784.
62. Elias, Lacetera, and Macis, "Markets and Morals," 2.
63. Grober, Reuben, and Tymula, "Political Quid Pro Quo Agreements," 582–583.