Research Design

Learning Objectives

In this chapter you will

1. Learn different types of research design
2. Learn the concept of validity
3. Learn about threats to validity
4. Learn how to align the research design to answer the research question

---

Identifying Research Design

Emily's Case

Emily, HR director at the city of Westlawn, brought her research team together to share what she discussed with Ahmed, the Community Foundation program officer. The team included training manager, Mei Lin, and a graduate student intern, Leo. Emily explained that she now had two research questions that would help them focus their evaluation of the training: "Does the training improve people's cultural competence?" And, "Does the training decrease workplace tension?" She shared her idea to measure cultural competence and workplace tension before and after the training to assess the impact. She mentioned the idea of splitting each department into two groups, so half would participate in the training before the others. Then the team could measure the level of cultural competence and workplace tension and compare the two groups at that point. She tasked Leo to find as much literature as possible that discusses training evaluation, measuring cultural competence, and workplace conflict. She asked Mei Lin to identify multiple scenarios for rolling out the training. They decided to have a weekly meeting to discuss how to implement the project.

Emily told Mei Lin and Leo, "It looks like this is going to be a lot of work, but I really want to do this right. I don't want to be doing the training for the sake of training without knowing what kind of impact it has on our employees. I believe focusing on the research design before we launch the training is important. I appreciate both of your help on this."

Mary's Case

Mary, volunteer manager at Health First, was thinking of her friend Yuki's advice to conduct a series of long interviews with her volunteers instead of administering a survey. She agreed that the idea of an in-depth interview was more likely to give her the information she wanted about recruiting and retaining volunteers, but she was concerned. Her experience in graduate school was with surveys, using quantitative data analysis and statistics. She knew how to interview people, of course, but she was not sure how this could qualify as research. She had always thought collecting a bunch of statements from people was too "soft" to be scientifically legitimate. She worried what the board members would think. There would be no numbers and charts to help her make an impressive presentation. "How do I convince them of anything?" she thought.

Later in the day, a package arrived at the office from Yuki. Two books were inside, and a jotted note: "Mary, knowing you, I'm sure you have millions of questions about qualitative research. Read these books first. Then call me. Enjoy!"

Mary was moved by Yuki's thoughtfulness and prompt attention. She chose one of the books and eagerly started reading.

Research Design: A Game Plan

Every research project needs a *game plan* to determine how an answer will be produced for the research question. The game plan is called a *research design.* In the research flow outlined in Chapter 2 (Figure 2.1), the research design is Step 3, following the research objective and research question. A research design will establish a plan that includes the following elements: (1) the structure of when the data are collected, (2) the time frame of the research, and (3) the number of data collection points.

There are numerous variations in research designs. Ethridge (2002) notes that "research designs are custom-made rather than mass-produced, and we will rarely find two that are identical" (p. 20). However, there are basic types of research, and it will be useful to understand the strengths and weaknesses of common types of research design applied to types of research. Some designs are suitable for a particular type of research and not others.

When choosing a research design, there are some key factors that need to be taken into consideration. Most important, the selected research design should match the purpose of the research (Kumar, 2011). It should allow the researcher to collect appropriate data that provides answers to the research question. Also, the selected research design should fit the research objective. A research design to describe and explore would be different from a research design to confirm and test a hypothesis. In other words, the research design needs to be in alignment with the research objective, research question, type of research, and the type of data required.

Figure 4.1 Types of Research Design Based on When the Data Are Collected
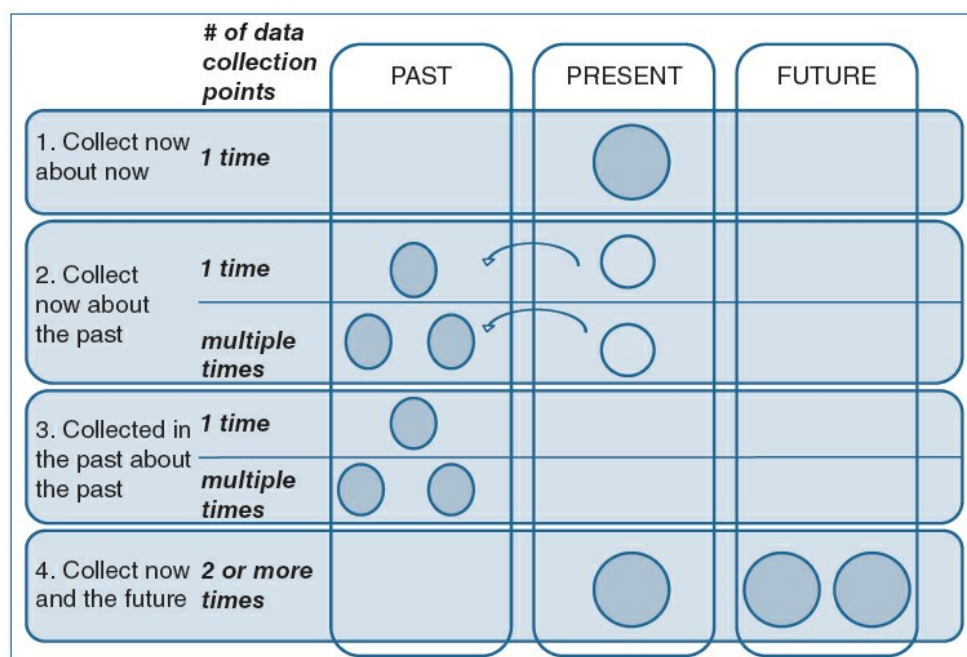
Types of Research Design

There are various ways to categorize types of research design. We have chosen to organize them by *when* and *how many times* data are collected. Figure 4.1 provides a schematic depiction of the organizing framework. In this framework, *when* the data collection occurs is represented along the horizontal axis in three categories: past, present, and future. Along the vertical axis, *when* refers to the character of the data, in four categories: collected now about now, collected now about the past, collected in the past about the past, and collected now and in the future. Along the vertical axis, we also took into account how many times the data are collected: one time or multiple times.

In this format, considering only *when* data are collected, four types of research design are distinguished:

Research Methods and Statistics for Public and Nonprofit
Administrators: A Practical Guide

1. *(1) Collect data one time now about now*. This research design is appropriate when you are interested in finding out how things are at the present moment. We see an example of this in Emily's case, with her interest in identifying the current level of cultural competence among city employees. If this is all she wanted to know, she could administer a one-time survey to obtain the information. This type of survey approach is referred to as **cross sectional survey design**.

2. *(2) Collect data now about the past.* In this research design, the data could focus on one event at a single time point or multiple events across multiple time points. We see an example of this research design in Mary's case, in her interest to ask volunteers why they volunteered, which refers to information about past events. Sometimes this kind of data can be collected in a survey. We saw, however, that Mary had difficulty finding a way to capture what she wanted to know in a survey. When collecting data about the past that stretches over a longer time period, not just one time point, a researcher may want to consider an in-depth interview or **oral history** to capture the information.

3. *(3) Collect data in the past about the past.* A researcher might be interested in data collected in the past only one time or multiple times over a period. Unlike the previous type of research design, this research design does not depend on the recall of an informant. We see an example of this research design in Jim's case, with his interest in response-time data since 2009. The times were recorded, so Jim can retrieve the data from archived records, ranging from 2009 to 2011, and analyze the trend (**trend analysis**) over multiple time points. This type of approach is referred to as **secondary data analysis**.

4. *(4) Collect data now and in the future.* This research design is typically used to assess change over time. Data collected at present as **baseline data** are compared to remeasurement at some point in the future. Remeasurement can occur multiple times, according to the resources of the researcher. Data collected multiple times in the future can be used to assess trends. This is similar to the previous research design, using secondary data from the past for trend analysis. Typically, though, this research design is used to assess the impact of an intervention and ascertain a cause-and-effect relationship. We see an example of this research design in Emily's case, in her intention to assess the impact of her cultural competence training. She is planning to conduct a baseline measurement with a survey, and repeat the same survey at a later time to observe any changes she could attribute to the effects of the training.

Conditions for Cause and Effec

When the objective of a research project is to confirm or test a hypothesized causal relationship, the research design requires special attention (Shadish, Cook, & Campbell, 2002).

The selection of the research design affects the level of rigor in making claims of causality based on study results. Generally speaking, in order to establish a causal relationship between *A* and *B* we need to meet the following three conditions:

- **Temporal precedence**: Changes in *A* precede the observed changes in *B*;
- **Covariation of the cause and effect**: Changes in *B* are related in a systematic way to changes in *A*;
- **No plausible alternative explanation**: No other factors are responsible for the observed changes in *B* (Trochim & Donnelly, 2007).

Figure 4.2 Temporal Precedence

Temporal Precedence
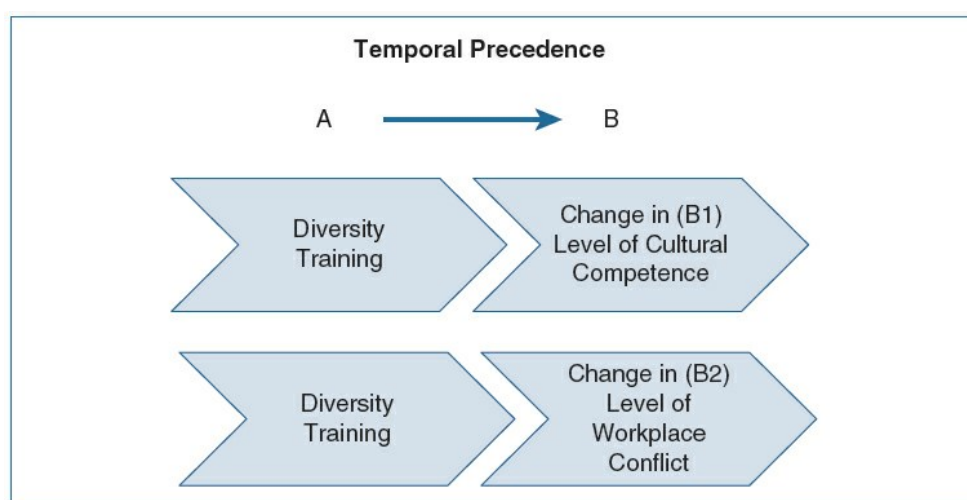
Temporal Precedence

When you attempt to establish that *A* causes *B*, one of the minimum conditions you need to meet is that the changes in *A* happened before the changes in *B*. A change that occurs prior to an event cannot be claimed to be caused by it. In Emily's case, the change in employees' experience due to the diversity training (A) needs to precede any observed changes in the level of cultural competence (B1) and level of workplace conflict (B2) observed in comparison to employees who did not receive the training.

Covariation of Cause and Effect

Another condition you need to meet the claim that *A* causes *B* is that changes in *A* are systematically related to changes in *B*. If the changes in *B* happen at random, regardless of the presence of changes in *A*, then you cannot make a claim that *A* caused *B*. In other words, *A* and *B* need to have a relationship.

If you observe that whenever *A* is present then *B* is present, or whenever *A* is absent, so is *B*,

there is a transitive relationship between *A* and *B.* This relationship is typically described as a **syllogism**:

If A, then B

If not A, then not B

In Emily's case, any changes in the level of cultural competence (B1) and the level of workplace conflict (B2) need to be systematically related to the change introduced by the diversity training (A). Putting the example of cultural competence in the syllogism illustrates why Emily needs a control group to establish a systematic relationship:

If diversity training (A) is offered, then there is an outcome in cultural competence (B1).

If diversity training (A) is not offered, then there is no outcome in cultural competence (B1).

A cause-and-effect relationship is not always binary (yes/no; present/absent). In some cases you may be looking for a situation where a different amount of *A* leads to a different amount of *B.* This relationship is described in a slightly different syllogism:

If more A, then more (or less) B

If less A, then less (or more) B

In Emily's case, she might later be interested in examining if any observed changes in cultural competence from the first training increased still more for employees who took additional diversity trainings. In this case, she would be looking for a relationship between the quantity of training and the quantity of improved cultural competence.

No Plausible Alternative Explanation

One you establish a relationship between *A* and *B* by temporal precedence and covariation, you then need to make sure the observed cause is not really due to some other factor *C* that is also systematically related to *A* and *B.* To be certain that *A* is the cause of *B*, you will need to eliminate all plausible alternative explanations for the changes in *B.* In Emily's case, she will need to show that no other factors other than diversity training (A) are responsible for any observed changes in the level of cultural competence (B1) and workplace conflict (B2).

When you claim a cause-and-effect relationship exists between *A* on *B*, the extent to which

your claim is valid is referred to as **internal validity.** The presences of plausible alternative explanations are threats to internal validity. Following are eight common threats to the internal validity (Campbell, Stanley, & Gage, 1963; Cook & Campbell, 1979; Shadish et al., 2002):

1. *History threat.* An external event can be a threat to the causal argument. Recall Emily's case in Chapter 2, when she was thinking of administering a survey and measuring the level of cultural competence and workplace conflict before and after the diversity training. She expected that any observed increase in cultural competence would be due to the training. Ahmed, the Community Foundation program officer, then countered that external events, such as the president giving a speech on race relations or a work picnic, might occur at the same time and influence the results. Ahmed was raising the possibility of a historical effect that could pose a plausible threat to validity.

2. *Maturation threat.* People change over time. They learn and mature from their daily experiences. They also grow physically and get older. This natural maturation can impact the outcomes you are observing in your research and could be a plausible alternative to your causal argument. This possibility makes it important to consider, for example, age or work experience differences among individuals in groups you are measuring.

3. *Instrumentation threat.* Researchers use a variety of instruments to measure the phenomenon of interest. Instrumentation threat refers to a case when the instrument itself could be influencing the result. In Emily's case, consider the survey she intends to use to measure cultural competence. If she uses a different survey before and after the diversity training, with different wording or order for the same questions, her results could be affected by differences in the instrument. A common example of the *instrumentation threat* to validity appears in face-to-face interviews on sensitive topics, where a respondent may shape answers to avoid a negative appearance. The *instrument* can also refer to the people who collect the data. Over time, an observer might get bored and pay less attention, or might learn from experience and change the way observations are interpreted.

4. *Testing threat.* Similar to the instrumentation threat, the testing threat operates when measurement takes place more than once. Here, instead of the instrument itself, the issue relates to a learning effect by the subjects being measured. In Emily's case, if she uses the same survey before and after the diversity training, there is a possibility that some employees will have thought about their earlier responses and decided to change answers to a "right" answer that they think Emily wants to hear. Similarly, if students are given the same math exam a second time, they might show improvement that reflects experience with the particular questions on the exam, not improved skill in the math involved.

5. *Mortality or Attrition threat.* During a research study participants will often drop out. The term *mortality* is used metaphorically (usually) to refer to the attrition of the study

participants. Participants may drop out for particular reasons, perhaps because they performed poorly in a baseline assessment or for other reasons that distinguish them from participants who stay in the study. If attrition is random, there may be no consequence, but a systematic change in the people in your study is likely to affect your results. It will at least make it plausible that there is a threat to validity, and you will need to address the issue to avoid criticism of the results. In Emily's case, suppose a number of employees refuse to take the cultural competence survey following the training. If the results of the survey show improvement, she will need to consider the possibility that those who dropped out do not endorse the idea of embracing diversity in the organization and were responsible for lower average scores on the initial survey. Dropping out can be a form of protest among individuals who are systematically different from those who continue to participate.

6. ***Regression threat.*** The *regression threat* is also known as **regression artifact** or **regression to the mean.** It refers to a statistical phenomenon that occurs when the mean (scores of the data) from a nonrandom sample is measured twice, it will move closer to the population mean. (Note: we will discuss more about *nonrandom sample* and *population* in Chapter 5.) Scores from a nonrandom sample may include extreme scores in the first measurement. However, when the same sample is measured twice, it is less likely that the extreme scores will persist. In other words, even if you do nothing to the sample, an extreme score is likely to move closer to the mean when measured a second time. This threat to validity is dependent on the level of variation possible in the value being measured.

7. ***Selection threat.*** Comparing two groups is a common procedure researchers take to establish causality for an intervention offered to one of the groups. If an outcome changes for the group with the intervention (called the **experimental group**), but the outcome does not change for the group with no intervention (called the **control group**), then you have a basis to argue that the intervention caused the change in outcome. This approach is called an **experimental design**. We will discuss this kind of research design in more detail in the next section of this chapter. When two or more groups are compared, attention needs to be given to the possibility that the composition of the groups are systematically different from each other and are not comparable. If the groups are different to begin with (**selection bias**), then any difference in results observed following an intervention could be due to the original difference in the groups and not from the intervention. Notice in Emily's case, when she decided with Ahmed to offer the training first to half of the employees in each department to compare to the other half who would not take the training, the issue arose how the employees would be selected. If Emily allows people to sign up for the training (self-selection), then a selection bias could occur. Very possibly, those employees most

interested in the diversity training would sign up first. In that case, improved survey results on cultural competence following the training could be due to their interest and predisposition to be influenced by the training. Emily and Ahmed agreed that the employees would need to be randomly assigned to take the training to avoid this kind of selection bias in the composition of the groups. Note that even when groups are selected by random assignment, researchers usually examine the resulting composition of the groups by age and other factors to assess any differences that might have occurred in the selection process.

8. ***Selection interaction threats.*** The selection threat to validity can also interact with other threats to internal validity. Variations are described below:

- A *selection–history* threat could occur if individuals in two groups experience an external event differently; for example, due to differences in a preexisting attitude or different reporting of the event.

- A *selection–maturation* threat could occur if two groups mature differently; for example, due to gender or socioeconomic differences.

- A *selection–instrumentation* threat could occur, for example, when responses from two groups are measured with two different survey instruments.

- A *selection–testing* threat could occur if differences between two groups influence the way they respond or learn from exposure to repeated testing; for example, due to perceived burden and inattention or learning to avoid stigma by finding the "right" answer.

- A *selection–regression* threat could occur, in one example, whenever an undetected selection bias occurs in the composition of two groups. At first measurement, the groups could appear similar and only later appear different, due to regression toward different original conditions. Random assignment should control for this possibility. The threat is more common in situations where researchers select extreme cases for an intervention and then find improvement occurs. If variation is possible in the value that was used to select the sample, then part of the improvement could be attributable to regression to the mean (Barnett, van der Pols, & Dobson, 2005).

Key Elements of Experimental Research Design

When the purpose of your research is to test if there is a cause-and-effect relationship, you must develop a research design that meets the three conditions elaborated in the above section: (1) temporal precedence, (2) covariation of the cause and effect, and (3) no plausible alternative explanation. Specifically, you will need an *experimental design* or **quasi-experimental design** (Fisher & Bennett, 1990; Shadish et al., 2002). In an experimental

design, data are collected before and after an intervention or treatment (i.e. pretest/posttest) with an experimental group and a control group, both randomly assigned. This design meets all three conditions for causality and is considered the most rigorous research design for making a causal argument. The quasi-experimental design has the same kind of group comparison before and after a treatment or intervention, but group assignment is not random. In Figure 4.1, both of these research designs belong to the type c*ollect now and the future*.

There are five key elements in the experimental design: (1) observations, (2) treatments or interventions, (3) groups, (4) assignment to group, and (5) time. In this section, we will explain each element and introduce notations that are frequently used. Figure 4.3 shows how the notations are used to illustrate a research design:

1. ***Observations.*** Observations are your measurement results, focused on the outcome or effect you are testing in your study. For example, in Emily's case, she is hypothesizing that the diversity training will have an outcome in *cultural competence* and *workplace conflict*. She has two observations: one before the training and one after the training. The notation *O* is typically used to refer to the observations. Subscripts, such as O_1, O_2, and so on, are used to distinguish a different set of results for the measures or different types of measure used for the observation.

Figure 4.3 Experimental Research Design Illustration

Research Methods and Statistics for Public and Nonprofit
Administrators: A Practical Guide
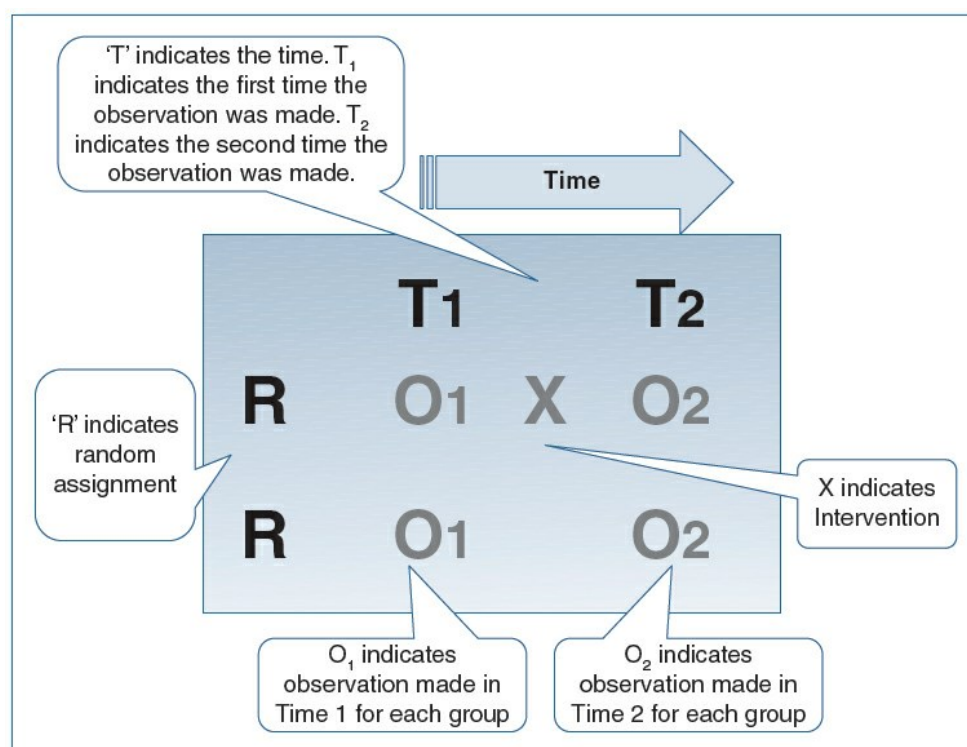
2. ***Treatments or intervention.*** Treatments or interventions are the hypothesized cause that is supposed to lead to a desired outcome (Judd & Kenny, 1981). In Emily's example, diversity training is the hypothesized cause for a desired increase in the employees' cultural competence and reduction in workplace conflict. A notation X is typically used to refer to an intervention or treatment.

3. ***Groups.*** When multiple numbers of groups are involved in a study, each group is given a line in the description of the design. For example, if the notation of the research design has three lines, that means the study involves three different groups. In Emily's case, if she decides to split the city employees into two groups—one group of employees to take the diversity training and another group to not take the training—then the description of her research design will have two lines. In this kind of experimental or quasi-experimental research design, the group that has a treatment or intervention is called the *experimental group* and will have an X in the line. The group that does not have any treatment or intervention is called the *control group* and will not have an X in the line.

4. ***Assignment to Group.*** When there are multiple groups involved in a study, you will need to decide how to assign study subjects to the groups. There are two ways to assign study subjects to the groups: random assignment and nonrandom assignment. **Random assignment** refers to the case when all study subjects are given an equal chance to be assigned to one of the groups in the study. Nonrandom assignment refers to the case when the assignment of the study subject is not randomized. Random assignment is preferred to assure the groups are roughly equivalent and comparable. Many factors, both known and unknown, could make the individuals in one group different from a second group. Even deliberately matching certain characteristics to make the groups appear comparable could still leave a selection bias in the composition of the groups. Random assignment is designed to overcome any selection bias by giving each study subject an equal chance to be assigned to one of the groups. **Nonrandom assignment** of study subjects to groups creates what is called **nonequivalent groups** (Fisher, 1970). We will discuss different methods for randomly or nonrandomly selecting study participants in Chapter 5. The notation R is used to denote random assignment, and NR is used for nonrandom assignment.

Figure 4.4 Quasi-Experimental Research Design Illustration
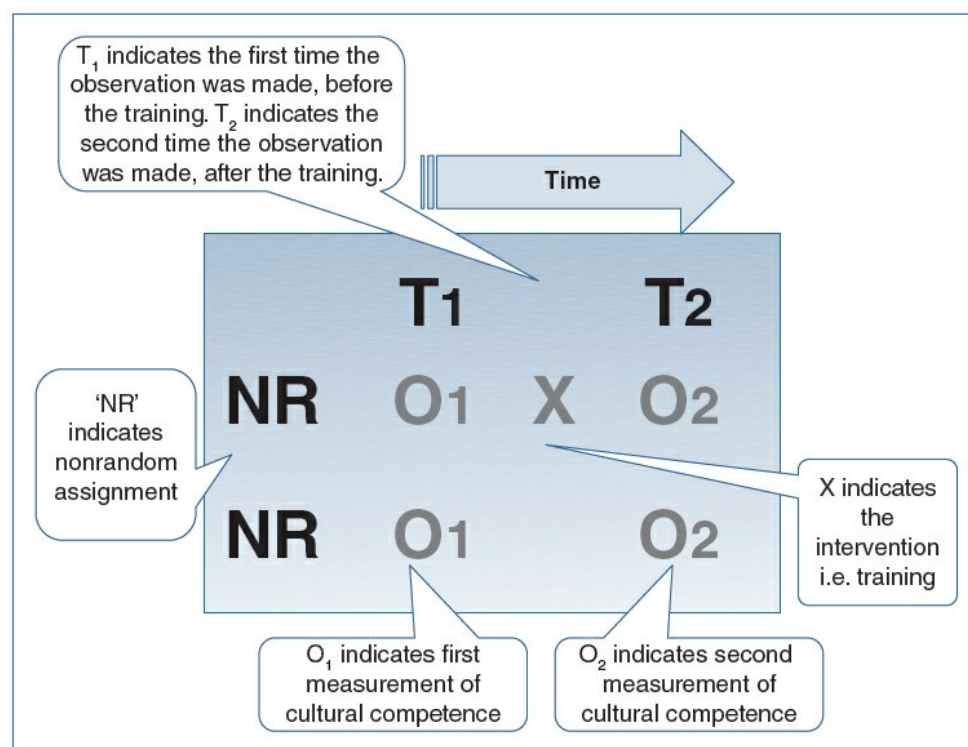
5. **Time.** One of the conditions in establishing causality is the temporal precedence of cause before effect. If a treatment or intervention is hypothesized as the cause of a certain effect, then it must occur prior to the effect. A researcher must be careful about the timing of observation and intervention periods to make sure the temporal order is maintained. In the description of the research design, time moves from left to right; elements listed on the left take place before the elements listed on the right. The typical notation of time is T. When the outcome is measured multiple times, subscripts, such as T_1, T_2 and so on, are used to distinguish different time points or times for certain measures in an observation.

Variations of Quasi-Experimental Research Design

In the social sciences and applied research, it is frequently not possible to randomly assign participants to groups. For example, people are already residents of certain geographic areas, children are already assigned to classrooms, programs may already be in operation, and policies already implemented. In addition, in public service work, ethical and legal constraints may prevent randomly exposing a particular group of people to a specific service. Whatever the reason, the given reality sometimes makes it impossible for a researcher to randomly assign study subjects into groups. Considering such factors in applied settings, a cause-and-effect research design may need to use a quasi-experimental design as the only feasible choice. In

the following sections, concluding this chapter, we will introduce a variety of quasi-experimental approaches. Jim's case will allow us to describe practical examples of an **after-only design** (or posttest only design), a *before-and-after design* (or pretest/posttest design), and a *before-and-after two group design* (or pretest/posttest two group design). In the final section, we will describe other variations that use additional groups or observation time points.

Jim's Case

Jim, deputy fire chief at the city of Rockwood, felt more confident with his research projects after meeting with his professor friend, Ty. He decided he needed to work on the alternative service model project first, because Chief Chen wanted to submit a budget proposal to the city council. Originally, he intended to call a few other jurisdictions to see what experience they had with alternative service models, but Chief Chen pointed out that no other jurisdiction in the state had adopted the alternative model he had in mind: sending a physician's assistant and a firefighter in a car to medical calls. Also, Ty had stated that the research objective was to "test a hypothesis," which made Jim think of setting two cars side-by-side on a race track. He needed to test the model in Rockwood, or make it apply to Rockwood. Other jurisdictions were different in size and population, and he wasn't sure results from somewhere else would be applicable. The trouble was he still didn't know how to start. He called Ty and asked if he had time to meet again.

In the fire station conference room, both men sat across from each other at the long table. Ty asked Jim what he had so far.

"I want to test this alternative service delivery model," Jim started, "but I really don't see how, unless we just do it and track those things we talked about last time, track the cost of the operation for efficiency, and see if the mortality rate goes down for effectiveness."

Ty stood up, gesturing to Jim that he was going to follow up on that idea. He wrote on the board:

After-OnlyDesign

X   O

Ty turned back to Jim and said, "If that's the only way you can implement the program, that's one

way to do it. This is called an 'after-only design.' The 'X' here represents the new model, and the 'O' represents your measurement of cost and mortality sometime after you implement it." He wrote the text in parentheses under the symbols: "alternative service program" and "cost/mortality rate.'"

"But wait," Jim interjected. "We have data on the cost of operation and the mortality rate for medical calls under the current service model. So if we introduce an alternative model, we can compare before and after we introduce the alternative model."

Ty grinned, turned around and continued writing on the board. "Good. Now you have a 'before-and-after model'." He started a new line:

Before-and-AfterDesign

O   X   O

"This is better than the 'after-only' design," Ty said as he wrote. "You compare service data before and after the implementation of the alternative model. In research design language, you have a pretest and a posttest, with an intervention in between."

Jim was glad Ty liked his idea. But then Ty asked another question.

"Jim, do you see any problem with this approach?"

Jim thought, "Problem? What's the problem?" He stared at the symbols on the board. Then something occurred to him. "Actually, there could be a problem," he said. "I know the number of medical calls changes during the year, and even from year to year, depending on the weather, certain holidays, like July fourth and New Year's, and I don't know what else, but I do know the numbers go up and down. And the severity of the incidents can be different, too. So if we start the alternative model and find good results, we still can't be sure that it's due to the model, or due to fluctuations in what's happening."

Ty looked pleased again. "Exactly. We need to rule out all other plausible explanations for any improvement we observe. Any thoughts on how to do that?"

"It would help, I guess, if we ran the model for a whole year, so it covers the same holidays, but I'm not sure that would account for everything. Plus, I don't think we could run a test that long without knowing if it's working. So, I don't know." He looked mischievously at Ty, "You tell me, professor."

Ty laughed. "All right. First let me tell you that you're right in everything you just said. What we

need here to solve the problems you mentioned is a control group. We need to start the alternative model with one group—call that the experimental group—and continue with another group that keeps the existing model of service delivery—call that the control group. Then we set them in operation with the same external circumstances over the same period of time."

Ty started to write a new set of lines on the board. "This way," he said, "you can compare the results between the two groups and decide if the alternative model had an effect." When Ty stepped aside from the board, Jim could see the new lines:

Before-and-After   Two Group design

O   *X*   *O*

O   *O*

Staring at the notation, Jim got the idea. "So, you want me to have some stations adopt the alternative service model, and some stations continue with the existing model, and measure them both sometime before and after we implement the alternative model?"

"That could be one way to do it," Ty answered. "Is that feasible?"

"I guess," Jim replied. "We have eight stations in Rockwood, so four could adopt the alternative model, and the remaining four could continue usual practice." As he formulated this idea, the advantages became apparent. "Actually, the council may like that idea. We won't have to change everything at once, just test the alternative model on a smaller scale for awhile. That will be cheaper."

Then a new problem occurred to him. Jim knew the different stations served neighborhoods with different numbers and kinds of medical calls. "Wait a minute," he said suddenly. "This doesn't solve anything. We still have one group with a set of external circumstances that are different from the other group. You can't compare these groups either."

Ty took the challenge in stride. "Good point. That's exactly what I was going to ask you next. How do you think we should select who uses the alternative model and who uses the current model? Ideally, we would toss a coin whenever a medical call comes in to any of the stations, and depending on whether we get heads or tails, send out either four firefighters and an engine or a physician's assistant and one firefighter in a car. Then add up our observations for cost and mortality for the two groups at the end."

*Ty drew a line down the center of the board and at the top of a new column wrote: **Group Assignment**.*

Research Methods and Statistics for Public and Nonprofit
Administrators: A Practical Guide

"A coin toss?" Jim muttered. "How can that be scientific?"

*Ty picked up on Jim's unease. "A coin toss would assure that all the medical calls have an equal chance to be assigned to one model or the other. This kind of **random assignment** would eliminate bias in selecting which calls go into each group, and the two groupings of calls would be, in theory, as equal as possible in their characteristics.*

Jim thought a moment about the different kinds of medical calls coming in—some severe and life threatening, some with urgent injuries or health problems, and some more frightening than anything else—and thought of a coin toss sending each one randomly to one model or the other. Sure, that might make a fair distribution. But the practical issue concerned him more. "That's impossible," he said firmly. "We can't equip every station with bothmodels and then keep everyone on call, waiting for the dispatcher to toss a coin and send only some of them out."

"I suspected that," Ty said. "So, tell me about the stations. At first you thought you could assign the stations to one model or the other, but then you decided the medical calls coming into the stations might be too different from each other to be comparable. What are you thinking?"

Jim answered, "We have four stations closer to downtown, and four stations in suburban and rural areas. The four stations downtown overlap to some degree and cover areas that are probably comparable, but the rural stations cover independent areas that are a little different from each other, but are more like each other, I think, than the urban stations. How do we decide which stations go into each group to make the groups equal? With only eight stations, I don't see how a coin toss will help us assign the stations to one model or the other. What if we end up with all four stations in the rural area assigned to the alternative model and all four stations near downtown assigned to the current model, just by chance?"

Ty listened to Jim and responded carefully. "You make a good point. Let's think this through. You say the location of the stations matters, because the kinds of medical calls some of them receive, overall, are different from other stations. The urban stations are more like each other, and different from the rural stations. The rural stations are also more like each other than they are to the urban stations. With these differences, you are worried that when we select which stations adopt the alternative model, there are too few stations to be confident that random assignment will give us equal groups in terms of the kinds of medical calls they receive. For example, one group might get all the urban stations. Is that right?"

Jim nodded.

"So, let's try something else," Ty said. He turned and wrote under the "Group Assignment"

heading, first one line saying, "1. Random Assignment: coin toss," and under that, "2. Nonrandom Assignment—matching." Then he scribbled in boxes and lines and arrows underneath. (See Figure 4.5.)

Turning back to Jim, Ty explained what he had in mind. "The trouble here is that once we gather a number of calls together by fire station, as a matter of convenience, then each call no longer has an equal chance of being randomly assigned to one group or the other, because it's dependent on the selection of any other call at that fire station. If one call is assigned to a group —the alternative model or the current model—then all the other calls at that station go, too. We need to correct for that. What we can do is use a form of nonrandom assignment called **matching.** *We'll group the urban stations on one side and the rural stations on the other, representing two different populations of calls. The calls are fairly similar within the matched groups, urban station or rural station, but different between the groups. You see?" Ty pointed to the drawings of boxes on the board, representing the urban stations on one side and rural stations on the other. "This is a* **nonrandom assignment***, because we are choosing. Once we've matched the stations this way, we can make a separate random assignment of stations within each matched group, so we will be sure to get two urban stations and two rural stations for each of the service delivery models."*

Jim took in the drawing. It made sense. "I can do that."

Figure 4.5 Jim's Design Options
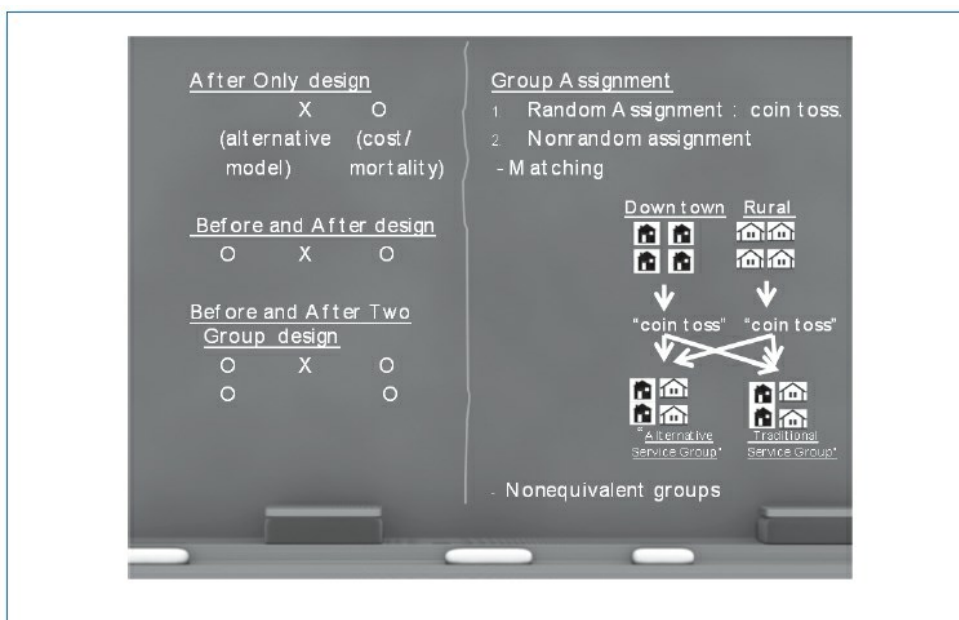
"Great," Ty replied. "I was worried that you might say something like, you will need to let each

station decide whether to try out the alternative model or not. If you do that, you could end up having two groups that look very different. Allowing the participants to choose the group they are in is likely to result in **nonequivalent groups.** *It's a good compromise in a lot of situations, if you have to do it, but not an ideal design."*

Making a Causal Argument Based on the Experimental Design

Jim's case illustrates the development of a quasi-experimental approach to a research design. The intent is to make a causal argument about a particular intervention, in this case, the alternative service delivery model. Jim and Ty discovered that making a random assignment of service calls to the experimental group (alternative model) or control group (current model) was not feasible. Consequently, they determined that a combination of **matching** and randomized assignment would be the most likely method to make the groups comparable. To complete the research design for a causal argument, Jim now needs to determine how his data will be collected.

Jim's Case (continues)

"We've made good progress," Ty continued. "Now we have an idea how to implement the alternative service delivery model so we can compare it to a control group before and afterthe implementation. Since we couldn't rely completely on random assignment for the two groups, we ended up with a quasi-experimental research design. That's OK in applied research like this. We still have a strong basis to assess the effect of the alternative service delivery model. What we need to do now is figure out how you will collect the data to measure cost and mortality. Why don't you walk me through your data collection process step-by-step?"

"All right," Jim said, and got up to go to the chalkboard. He had been thinking about this part of the research. "We collect data on operating cost for every station as well as the mortality rate, so I was thinking I would compile the data from the last six months."

On the left side of the board Jim wrote: "Compile cost & mortality rate for the last 6 months Jan–June FY 00." He drew a box around the text and over it wrote: "FY 00." To the right, he wrote "FY 01."

Turning to Ty, Jim said, "I figure I can collect data from January to June as baseline data. The next fiscal year starts in July. We use the next six months, July to December, to set up a system for the alternative model, then run it for six months next year, January to June"—he pointed at the "FY 01" at the top of the board—"during the same time of the year, you see, because I think that's important. Then I collect data again and see how it works."

Jim wrote in more information for the planning phase and the idea of four stations adopting the new model. Ty was impressed. He moved to the board next to Jim.

"Let's add the notation for a research design we talked about earlier," Ty said. Underneath what Jim had written on the board, Ty added notation to illustrate the research design:

4 stations	O1 (A)	X	O2 (A)
(Alternative model)			
4 stations	O1 (T)		O2 (T)
(Traditional model)			

Ty explained, "You are going to collect data from all eight stations while they are operating with the traditional model, but four of the stations are going to be the experimental group, and will adopt the alternative model." He pointed to the "X" in the middle of the board in line with the top row of notation. "The other four stations will continue with the traditional model."

"I see," Jim said. "The 'O' is an observation period, the 'X' is the start of the alternative model."

"That's right," Ty said, and moved to another board on the wall. He drew a graph, and along the bottom wrote in the two time periods, "This year Jan~June" and "Next year Jan~June." On the vertical axis he wrote in numbers, from 0 to 9. "This is just for illustration," Ty said as he chalked in a heavy dashed line for the traditional model, and a heavy solid line for the alternative model. The dashed line was almost flat as it moved from the first time period to the second. The solid line started in about the same place as the dashed line, and then dropped dramatically.

Figure 4.6 Jim's Time Line
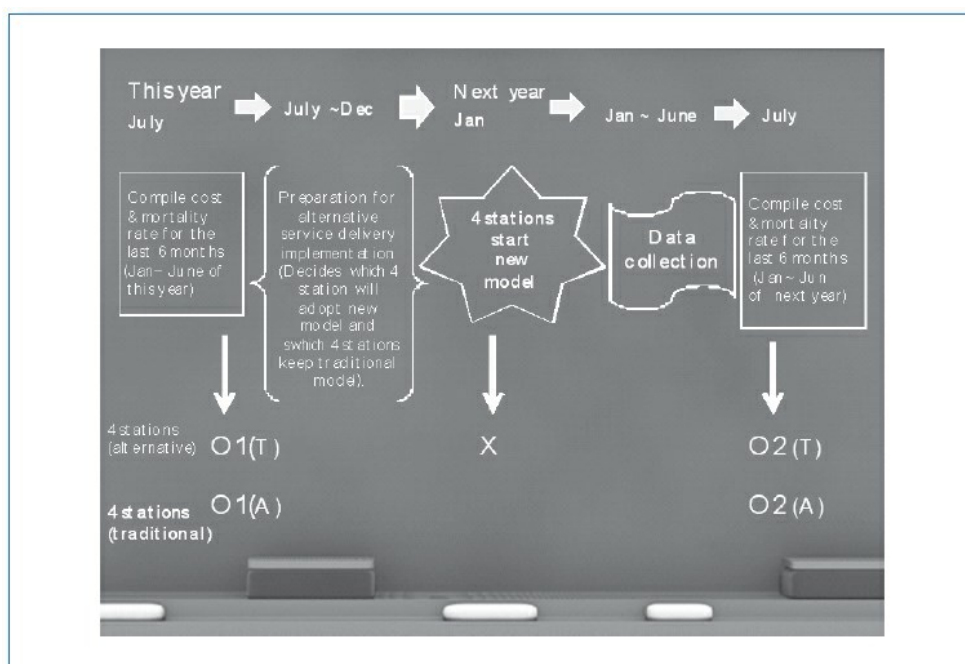
"If the matching works when you select your two groups of stations," Ty explained, "then the cost and mortality values you get during the first observation period should be about the same for both models." He pointed to the starting point for the lines. "If the alternative model really reduces cost or mortality—you could use a graph like this for either one—then you will see the difference at the second observation period." He pointed to the wide gap at the end points. "If it works, this could be a good way to make your argument to the city council."

Jim nodded.

"The real reason I wanted to show you this graph," Ty continued, "is to get you to start thinking about how you are going to get the numbers for your results. Notice I just made up the numbers here from zero to nine. We don't know yet what your numbers are going to look like. If you have good data, like you say, then calculating a number for mortality rate or the cost should be pretty straightforward."

Ty tilted his arms up in surrender and smiled, indicating he was done. Jim looked around the room at all the writing on the chalkboards and said in a low voice, "I think I can do this." The two friends joked and gathered up their things.

In the foyer outside the conference room, they shook hands, and Jim looked straight at Ty. "You are really a boring guy," he beamed, "but thanks to you, I know I can make a good proposal to the council. And I know how to get this project going."

Figure 4.7 Graph of Change in Outcome: Suggesting Causation
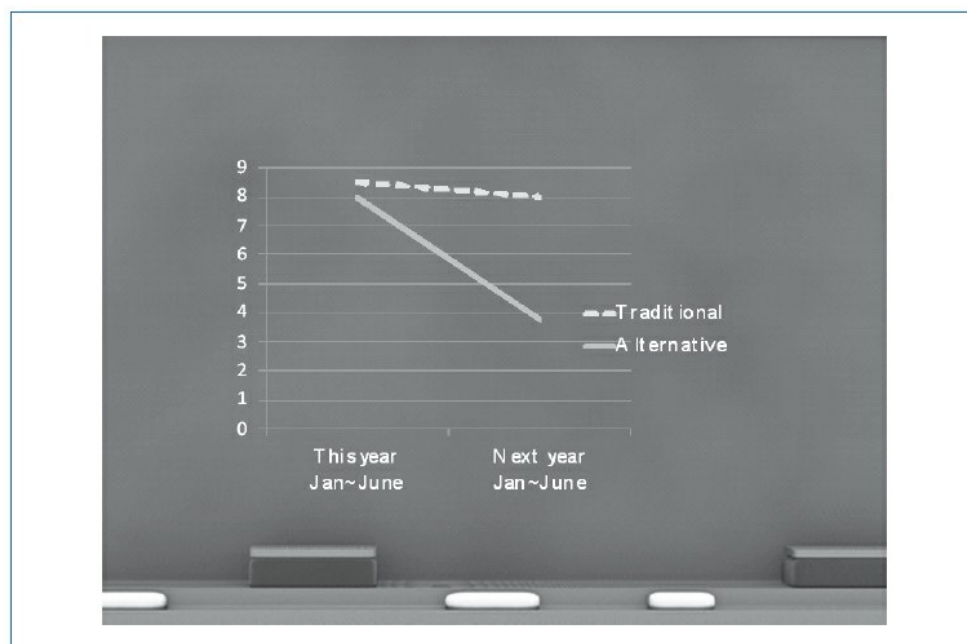
Other Variations of Experimental and Quasi-Experimental Design

So far we have looked at the experimental designs and quasi-experimental designs with one experimental group and one control group. There are other ways you can structure and design the research. For example, it is possible to have more than one treatment group. One of the most influential and widely cited policy experiments, *Minneapolis Domestic Violence Experiment* (Sherman & Berk, 1984), used three different interventions that were compared to each other. The focus of the study was to determine which strategy was most effective at reducing domestic violence assaults. Among a pool of offenders where probable cause for an arrest existed, officers were directed to randomly choose how to proceed by opening instructions at the scene that were sealed in an envelope. Three different instructions could be in the envelopes: (1) arrest the suspect, (2) separate the parties for 8 hours, or (3) advise and mediate. The notation would appear as follows:

Arrest	R	O_1	X_1	O_2
Separate	R	O_1	X_2	O_2
Mediate	R	O_1	X_3	O_2

The researchers observed police records for subsequent assaults six months later and calculated the percentage of repeat offenses for the three different interventions: arrest 19%, separate 33%, and mediate 37%. Among the options police officers had available to them, represented by the three interventions, arrest was shown to be the most effective. Comparing

the randomly assigned groups to each other provided a clear result.

Another variation is found in the *placebo* design, commonly used in clinical trials for pharmaceuticals. In medical interventions, it is known that when patients believe they are receiving treatments, they may improve even when the treatment has no therapeutic benefit. This psychological effect is called a **placebo effect.** To control for this possible result, medical researchers have learned to imitate an intervention with a placebo that appears just like the intervention, so the subjects (and usually the researchers) do not know if they are getting the real treatment. The research design has three groups: a treatment group, a placebo group, and no treatment. With this design, an experimental treatment needs to demonstrate not only that it is better than no treatment, but also that is it better than a placebo. The notation would appear as:

Treatment group	R	O_1	X_1	O_2
Placebo group	R	O_1	X_2	O_2
Control group	R	O_1		O_2

Limitations in applied research can sometimes determine the research design. For example, pretest information may not always be available, especially in program and policy evaluations where a decision was made to assess effectiveness only after the fact. In this situation, researchers could use an *after-only design* (introduced in Jim's case). This research design includes an experimental group and a control group, but with only one measurement after the implementation. The results can suggest the effectiveness of an intervention, but the design is not ideal in terms of rigor, and it limits the ability to make a causal argument. The notation would appear as in the example:

XO

  O

The after-only design is used as a kind of control in a more complex research design, called the **Solomon Four-Group Design**. This design utilizes four groups in a hybrid experimental design: the first group (A) has a pretest and posttest with intervention; the second group (B) is a control group to Group A, with a pretest and posttest, but no intervention; the third group (C) receives an intervention like Group A, and a posttest, but no pretest; and the fourth group (D) is a control group for Group C, with a posttest, but no pretest and no intervention. The notation would appear as shown:

Group A	R	O	X	O
Group B	R	O		O
Group C	R		X	O
Group D	R			O

This Solomon four-group design is useful to control for a possible testing threat to validity, where a subject's exposure to the test or measurement at the pretest may have affected the posttest scores. There are a number of possible comparisons built in. First, the researcher can compare the difference in posttest scores between groups A and B versus the difference in posttest scores between groups C and D. If the difference score between A and B is similar to the difference score between C and D, then the researcher can rule out the testing threat. A comparison can also be made between Group A and Group C for posttest scores, as both groups received the treatment, and a comparison can be made between Group B and Group D, as both groups did not receive the treatment. If Group A and Group C have similar scores, and Group B and Group D have similar scores, then the researcher can rule out the testing threat.

Another variation of the research design is called **time series design**, which takes measures or observations of a single variable at many consecutive periods in time. These designs are sometimes also referred to as **interrupted time series designs**. In this version, several observations are conducted before a treatment is introduced, and then there are another series of observations. The notation would appear as written below:

O   O   O   O   X   O   O   O   O

This research design has an advantage over before-and-after observations, because it controls for history and any immediate effects the treatment may have that could possibly dissipate as time progresses. In Jim's case, this design could be adopted to track cost and mortality before the implementation date of the alternative model, and then after implementation, to detect any changes that may be occurring due to external factors.

To increase the rigor of this design, a control group can be introduced. The advantage here is that the researcher can get more precise information on the trends that lead up to the intervention, and how things change afterward even when there is no intervention. The notation would appear as the following:

O   O   O   O   X   O   O   O   O
O   O   O   O   O   O   O   O

Ethical Considerations in Experimental and Quasi-Experimental Design

In designing an experimental or a quasi-experimental study, researchers need to consider its ethical implications on subjecting study participants to a treatment, or not providing a certain group an opportunity to benefit from the experimental treatment. In a placebo study, is it ethical for a researcher to subject study participants to treatments that are known to have no effect on the outcome, though the study participants believe they are receiving a treatment? In Emily's case, is it ethical for her to randomly assign a group of employees to benefit from diversity training, and not allow another group to take the training? In Jim's case, is it ethical to introduce an alternative model of service delivery when the impact on residents is unknown?

These are the kinds of important considerations that a researcher needs to weigh before finalizing the research design. One way to address some of these ethical concerns is to obtain **informed consent** from the study participants. In a placebo study, participants should be informed prior to their participation to the study that they may be receiving a treatment that may not be effective, and they are taking that chance. In Emily's case, she could make sure that employees who did not originally take the training received the opportunity later. In Jim's case, he might inform the residents of the City of Rockwood that the fire department is implementing the experimental alternative service model, discuss possible pros and cons of the alternative service model, and get citizen consent. Researchers need to consider these issues and be aware that there may be some instances where experimental or quasi-experimental approaches may not be appropriate, due to ethical implications.

Chapter Summary

In this chapter we introduced different types of research design. Research design is a game plan for your research. You will need to identify your research design in Step 3 of your research process after you have determined your research objective (Step 1) and research questions (Step 2). Research design can be categorized based on when data were collected and what information the data captured. The four types of research design we identified are research that (a) collects data now about now, (b) collects data now about the past, (c) uses data already collected in the past about the past, and (d) collects data now about now and again in the future.

We also discussed key principles that the research design needs to meet in order to establish a causal argument: (1) temporal precedence, (2) covariation of the cause and effect, and (3) no plausible alternative explanation. In ruling out plausible alternative explanations in the research design, researchers can eliminate threats to validity. The eight threats we discussed are: history, maturation, instrumentation, testing, mortality (attrition), regression to the mean, selection, and interaction with selection. As a way to address these threats, we introduced the

basic idea of experimental and quasi-experimental design. Jim's case illustrated the development of a quasi-experimental design. We also introduced some variations on experimental and quasi-experimental designs. Finally, we introduced ethical implications researchers need to consider, with a few examples of situations that could impact study participants.

There are many things to think about when deciding what type of research design most suits your research. There are also some practical aspects that need to be taken into account, such as availability of personnel, funding, time, and existing data. Your role as a researcher is to make the final determination on what type of research design is most appropriate for the research question you are pursuing, and is also balanced with practical constraints.

Review and Discussion Questions

1. Review the approaches Emily, Jim, and Mary are considering for their study. How would you classify their approach in terms of the four types of research design introduced in this chapter?
2. Consider yourself as a research consultant (like Ty). Suppose Emily came to you to get help deciding the details of her research. Imagine your conversation with Emily and develop a research design for her. What insights would you offer, and what would be your rationale for the approach chosen?
3. How would you describe the primary difference between experimental and quasi-experimental designs? What are the implications of adopting an experimental design versus quasi-experimental design in an applied setting?
4. How does random assignment in a research design assist in increasing internal validity?
5. Discuss a possible internal threat to validity if Jim adopts an after-only design.
6. A municipality has had a problem with crashes in some intersections due to motorists running red lights. To combat this problem, the city decided to install red light cameras that photograph a violator in the intersection and send a citation through the mail. To evaluate the effectiveness of this program (if any) and determine if it was due to the intervention, why might a time series design be beneficial? Is there a threat to validity?
7. Find a research-based article for a topic that you are interested in. After reading the author's description of the research methods, categorize the approach into one of the four types of research design. What other research design approaches can you think of to address the research questions?

References

Barnett A. G., **van der Pols J. C.**, & **Dobson A. J.** (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology*, 34(1), 215–220.

Campbell D. T., **Stanley J. C.**, & **Gage N. L.** (1963). *Experimental and quasi-experimental designs for research.* Chicago, IL: Rand McNally.

Cook T. D., & **Campbell D. T.** (1979). *Quasi-experimentation: Design & analysis issues for field settings.* Chicago, IL: Rand McNally.

Ethridge M. E. (2002). *The political research experience: Readings and analysis.* Armonk, NY: M. E. Sharpe.

Fisher R. A. (1970). *Statistical methods for research workers.* Darien, CT: Hafner.

Fisher R. A., & **Bennett J. H.** (1990). *Statistical methods, experimental design, and scientific inference.* Oxford, UK: Oxford University Press.

Judd C. M., & **Kenny D. A.** (1981). *Estimating the effects of social interventions.* Cambridge, NY: Cambridge University Press.

Kumar R. (2011). *Research methodology: A step-by-step guide for beginners.* Los Angeles, CA: Sage.

Shadish W. R., **Cook T. D.**, & **Campbell D. T.** (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton Mifflin.

Sherman L. W., & **Berk R. A.** (1984). *Minneapolis domestic violence experiment.* Washington, DC: Police Foundation.

Trochim W. M. K., & **Donnelly J. P.** (2007). *Research methods knowledge base.* Mason, OH: Thomson Custom.

Key Terms

After-Only Design With Comparison Group 59

Baseline Data 51

Control Group 55

Covariation of the Cause and Effect 51

Cross Sectional Survey Design 50

Experimental Design 55

Experimental Group 55

Group Assignment 57

History Threat 53

Student Study

Visit the Student Study Site at **www.sagepub.com/nishishiba1e** for these additional learning tools:

- Data sets to accompany the exercises in the chapter

http://dx.doi.org/10.4135/9781544307763.n4