Chapter 14: Non-experimental Design

This chapter provides an overview of Non-Experimental Research including Correlational Research, Qualitative Research and Observational Research.

## Learning Objectives

1. Define non-experimental research, distinguish it clearly from experimental research, and give several examples.
2. Explain when a researcher might choose to conduct non-experimental research as opposed to experimental research.
3. Define correlational research and give several examples.
4. Explain why a researcher might choose to conduct correlational research rather than experimental research or another type of non-experimental research.
5. Interpret the strength and direction of different correlation coefficients.
6. Explain why correlation does not imply causation
7. List several ways in which qualitative research differs from quantitative research in psychology.
8. Describe the strengths and weaknesses of qualitative research in psychology compared with quantitative research.
9. Give examples of qualitative research in psychology.
10. List the various types of observational research methods and distinguish between each.
11. Describe the strengths and weakness of each observational research method.

# Non-Experimental Research

Researchers who are simply interested in describing characteristics of people, describing relationships between variables, and using those relationships to make predictions can use **non-experimental** research. Using the non-experimental approach, the researcher simply measures variables as they naturally occur, but they do not manipulate them. For instance, if I just measured the number of traffic fatalities in America last year that involved the use of a cell phone, but I did not actually manipulate cell phone use then this would be categorized as non-experimental research. Alternatively, if I stood at a busy intersection and recorded drivers’ genders and whether or not they were using a cell phone when they passed through the intersection to see whether men or women are more likely to use a cell phone when driving, then this would be non-experimental research. It is important to point out that non-experimental does not mean nonscientific. Non-experimental research is scientific in nature. It can be used to fulfill two of the three goals of science (to describe and to predict). However, unlike with experimental research, we cannot easily make causal conclusions using this method; we generally cannot confidently say that one variable causes another variable using this method.

What do the following classic studies have in common?

Stanley Milgram found that about two thirds of his research participants were willing to administer dangerous shocks to another person just because they were told to by an authority figure (Milgram, 1963).

Elizabeth Loftus and Jacqueline Pickrell showed that it is relatively easy to “implant” false memories in people by repeatedly asking them about childhood events that did not actually happen to them (Loftus Pickrell, 1995).

John Cacioppo and Richard Petty evaluated the validity of their Need for Cognition Scale—a measure of the extent to which people like and value thinking—by comparing the scores of university professors with those of factory workers (Cacioppo Petty, 1982).

David Rosenhan found that confederates who went to psychiatric hospitals claiming to have heard voices saying things like “empty” and “thud” were labeled as schizophrenic by the hospital staff and kept there even though they behaved normally in all other ways (Rosenhan, 1973).

The answer for purposes of this chapter is that they are not experiments. In this chapter, we look more closely at non-experimental research. We begin with a general definition of non-experimental research, along with a discussion of when and why non-experimental research is more appropriate than experimental research. We then look separately at two important types of non-experimental research: correlational research and observational research.

## What Is Non-Experimental Research?

Non-experimental research is research that lacks the manipulation of an independent variable. Rather than manipulating an independent variable, researchers conducting non-experimental research simply measure variables as they naturally occur (in the lab or real world).

Most researchers in psychology consider the distinction between experimental and non-experimental research to be an extremely important one. This is because although experimental research can provide strong evidence that changes in an independent variable cause differences in a dependent variable, non-experimental research generally cannot. As we will see, however, this inability to make causal conclusions does not mean that non-experimental research is less important than experimental research. It is simply used in cases where experimental research is not able to be carried out.

## When to Use Non-Experimental Research

As we saw in the last chapter, experimental research is appropriate when the researcher has a specific research question or hypothesis about a causal relationship between two variables—and it is possible, feasible, and ethical to manipulate the independent variable. It stands to reason, therefore, that non-experimental research is appropriate—even necessary—when these conditions are not met. There are many times in which non-experimental research is preferred, including when:

* the research question or hypothesis relates to a single variable rather than a statistical relationship between two variables (e.g., how accurate are people’s first impressions?).
* the research question pertains to a non-causal statistical relationship between variables (e.g., is there a correlation between verbal intelligence and mathematical intelligence?).
* the research question is about a causal relationship, but the independent variable cannot be manipulated or participants cannot be randomly assigned to conditions or orders of conditions for practical or ethical reasons (e.g., does damage to a person’s hippocampus impair the formation of long-term memory traces?).
* the research question is broad and exploratory, or is about what it is like to have a particular experience (e.g., what is it like to be a working mother diagnosed with depression?).

Again, the choice between the experimental and non-experimental approaches is generally dictated by the nature of the research question. Recall the three goals of science are to describe, to predict, and to explain. If the goal is to explain and the research question pertains to causal relationships, then the experimental approach is typically preferred. If the goal is to describe or to predict, a non-experimental approach is appropriate. But the two approaches can also be used to address the same research question in complementary ways. For example, in Milgram’s original (non-experimental) obedience study, he was primarily interested in one variable—the extent to which participants obeyed the researcher when he told them to shock the confederate—and he observed all participants performing the same task under the same conditions. However, Milgram subsequently conducted experiments to explore the factors that affect obedience. He manipulated several independent variables, such as the distance between the experimenter and the participant, the participant and the confederate, and the location of the study (Milgram, 1974).

## **Types of Non-Experimental Research**

Non-experimental research falls into two broad categories: correlational research and observational research.

The most common type of non-experimental research conducted in psychology is correlational research. Correlational research is considered non-experimental because it focuses on the statistical relationship between two variables but does not include the manipulation of an independent variable. More specifically, in correlational research, the researcher measures two variables with little or no attempt to control extraneous variables and then assesses the relationship between them. As an example, a researcher interested in the relationship between self-esteem and school achievement could collect data on students’ self-esteem and their GPAs to see if the two variables are statistically related.

Observational research is non-experimental because it focuses on making observations of behavior in a natural or laboratory setting without manipulating anything. Milgram’s original obedience study was non-experimental in this way. He was primarily interested in the extent to which participants obeyed the researcher when he told them to shock the confederate and he observed all participants performing the same task under the same conditions. The study by Loftus and Pickrell (1995) is also a good example of observational research. The dependent variable was whether participants “remembered” having experienced mildly traumatic childhood events (e.g., getting lost in a shopping mall) that they had not actually experienced but that the researchers asked them about repeatedly. In this particular study, nearly a third of the participants “remembered” at least one event. As with Milgram’s original study, this study inspired several later studies using experimental research methods on the factors that affect these phenomena.

## Internal Validity in Non-Experimental Design

Recall that internal validity is the extent to which the design of a study supports the conclusion that changes in the independent variable caused any observed differences in the dependent variable. Experimental research is highest in internal validity, when properly controlled, because the use of manipulation (of the independent variable) and control (of extraneous variables) help to rule out alternative explanations for the observed relationships. If the average score on the dependent variable in an experiment differs across conditions, it is quite likely that the independent variable is responsible for that difference. Non-experimental (correlational) research is lowest in internal validity because these designs fail to use manipulation or control. There are also research approaches that blend both experimental and non-experimental approaches described as **quasi-experimental research** (which will be described in more detail in a subsequent chapter). This type of research contains some, but not all, of the features of a true experiment. For instance, it may fail to use random assignment to assign participants to groups or fail to use counterbalancing to control for potential order effects. Imagine, for example, that a researcher finds two similar schools, starts an anti-bullying program in one, and then finds fewer bullying incidents in that “treatment school” than in the “control school.” While a comparison is being made with a control condition, the inability to randomly assign children to schools could still mean that students in the treatment school differed from students in the control school in some other way that could explain the difference in bullying (e.g., there may be a selection effect).

# What Is Correlational Research?

Correlational research is a type of non-experimental research in which the researcher measures two variables (binary or continuous) and assesses the statistical relationship (i.e., the correlation) between them with little or no effort to control extraneous variables. There are many reasons that researchers interested in statistical relationships between variables would choose to conduct a correlational study rather than an experiment. The first is that they do not believe that the statistical relationship is a causal one or are not interested in causal relationships. Recall two goals of science are to describe and to predict and the correlational research strategy allows researchers to achieve both of these goals. Specifically, this strategy can be used to describe the strength and direction of the relationship between two variables and if there is a relationship between the variables then the researchers can use scores on one variable to predict scores on the other (the specific analytic approach will be discussed in Chapter 16).

Another reason that researchers would choose to use a correlational study rather than an experiment is that the statistical relationship of interest is thought to be causal, but the researcher cannot manipulate the independent variable because it is impossible, impractical, or unethical. For example, while a researcher might be interested in the relationship between the frequency people use cannabis and their memory abilities, they cannot ethically manipulate the frequency that people chronically use cannabis. As such, they must rely on the correlational research strategy; they must simply measure the frequency that people use cannabis and measure their memory abilities using a standardized test of memory and then determine whether the frequency people use cannabis is statistically related to memory test performance.

Another strength of correlational research is that it is often higher in external validity than experimental research. Recall there is typically a trade-off between internal validity and external validity. As greater controls are added to experiments, internal validity is increased but often at the expense of external validity as artificial conditions are introduced that do not exist in reality. In contrast, correlational studies typically have low internal validity because nothing is manipulated or controlled but they often have high external validity. Since nothing is manipulated or controlled by the experimenter the results are more likely to reflect relationships that exist in the real world.

Finally, extending upon this trade-off between internal and external validity, correlational research can help to provide converging evidence for a theory. If a theory is supported by a true experiment that is high in internal validity as well as by a correlational study that is high in external validity, then the researchers can have more confidence in the validity of their theory. As a concrete example, correlational studies establishing that there is a relationship between watching violent television and aggressive behavior have been complemented by experimental studies confirming that the relationship is a causal one (Bushman & Huesmann, 2001).

## Does Correlational Research Always Involve Quantitative Variables?

A common misconception among beginning researchers is that correlational research must involve two quantitative variables, such as scores on two extraversion tests or the number of daily hassles and number of symptoms people have experienced. However, the defining feature of correlational research is that the two variables are measured—neither one is manipulated—and this is true regardless of whether the variables are quantitative or categorical. Imagine, for example, that a researcher administers the Rosenberg Self-Esteem Scale to 50 American college students and 50 Japanese college students. Although this “feels” like a between-subjects experiment, it is a correlational study because the researcher did not manipulate the students’ nationalities. The same is true of the study by Cacioppo and Petty (1982) comparing college faculty and factory workers in terms of their need for cognition. It is a correlational study because the researchers did not manipulate the participants’ occupations.

The statistical tool based on calculating a correlation coefficient (Chapter 16) does require having two variables that are quantitative and not categorical. While it is commonly the case that correlational research depends on calculating correlation coefficients, be aware that the terms are not actually synonymous. While it is technically possible to have an experimental design that is evaluated statistically with a correlation analysis (calculating a correlation coefficient), in practice this is extremely rare. It would require having a large number of possibilities across the manipulated independent variable, which is a very rare case. In all other cases, the term ‘correlation’ should be avoided in describing experimental research as the neither the statistical tool nor the general method applies (if necessary, the related word ‘associated’ should be preferred over ‘correlated’).

## **Data Collection in Correlational Research**

Again, the defining feature of correlational research is that neither variable is manipulated. It does not matter how or where the variables are measured. A researcher could have participants come to a laboratory to complete a computerized backward digit span task and a computerized risky decision-making task and then assess the relationship between participants’ scores on the two tasks. Or a researcher could go to a shopping mall to ask people about their attitudes toward the environment and their shopping habits and then assess the relationship between these two variables. Both of these studies would be correlational because no independent variable is manipulated.

## Correlation Does Not Imply Causation

You have probably heard repeatedly that “Correlation does not imply causation.” An amusing example of this comes from a 2012 study that showed a positive correlation (Pearson’s r = 0.79) between the per capita chocolate consumption of a nation and the number of Nobel prizes awarded to citizens of that nation. It seems clear, however, that this does not mean that eating chocolate causes people to win Nobel prizes, and it would not make sense to try to increase the number of Nobel prizes won by recommending that parents feed their children more chocolate.

There are two reasons that correlation does not imply causation. The first is called the directionality problem. Two variables, X and Y, can be statistically related because X causes Y or because Y causes X. Consider, for example, a study showing that whether or not people exercise is statistically related to how happy they are—such that people who exercise are happier on average than people who do not. This statistical relationship is consistent with the idea that exercising causes happiness, but it is also consistent with the idea that happiness causes exercise. Perhaps being happy gives people more energy or leads them to seek opportunities to socialize with others by going to the gym. The second reason that correlation does not imply causation is called the third-variable problem. Two variables, X and Y, can be statistically related not because X causes Y, or because Y causes X, but because some third variable, Z, causes both X and Y. For example, the fact that nations that have won more Nobel prizes tend to have higher chocolate consumption probably reflects geography in that European countries tend to have higher rates of per capita chocolate consumption and invest more in education and technology (once again, per capita) than many other countries in the world. Similarly, the statistical relationship between exercise and happiness could mean that some third variable, such as physical health, causes both of the others. Being physically healthy could cause people to exercise and cause them to be happier. Correlations that are a result of a third-variable are often referred to as spurious correlations.

Some other excellent and amusing examples of spurious correlations can be found at <http://www.tylervigen.com> (Above is one such example). Chart, line chart

Description automatically generated

“Lots of Candy Could Lead to Violence”

Although researchers in psychology know that correlation does not imply causation, many journalists do not. One website about correlation and causation, <http://jonathan.mueller.faculty.noctrl.edu/100/correlation_or_causation.htm>, links to dozens of media reports about real biomedical and psychological research. Many of the headlines suggest that a causal relationship has been demonstrated when a careful reading of the articles shows that it has not because of the directionality and third-variable problems.

One such article is about a study showing that children who ate candy every day were more likely than other children to be arrested for a violent offense later in life. But could candy really “lead to” violence, as the headline suggests? What alternative explanations can you think of for this statistical relationship? How could the headline be rewritten so that it is not misleading?

As you have learned by reading this book, there are various ways that researchers address the directionality and third-variable problems. The most effective is to conduct an experiment. For example, instead of simply measuring how much people exercise, a researcher could bring people into a laboratory and randomly assign half of them to run on a treadmill for 15 minutes and the rest to sit on a couch for 15 minutes. Although this seems like a minor change to the research design, it is extremely important. Now if the exercisers end up in more positive moods than those who did not exercise, it cannot be because their moods affected how much they exercised (because it was the researcher who used random assignment to determine how much they exercised). Likewise, it cannot be because some third variable (e.g., physical health) affected both how much they exercised and what mood they were in. Thus, experiments eliminate the directionality and third-variable problems and allow researchers to draw firm conclusions about causal relationships.

# **What Is Qualitative Research?**

This textbook is primarily about quantitative research, in part because most studies conducted in psychology are quantitative in nature. Quantitative researchers typically start with a focused research question or hypothesis, collect a small amount of numerical data from a large number of individuals, describe the resulting data using statistical techniques, and draw general conclusions about some large population. Although this method is by far the most common approach to conducting empirical research in psychology, there is an important alternative called qualitative research. Qualitative research originated in the disciplines of anthropology and sociology but is also used to study psychological topics. Qualitative researchers generally begin with a less focused research question, collect large amounts of relatively “unfiltered” data from a relatively small number of individuals, and describe their data using nonstatistical techniques, such as grounded theory, thematic analysis, critical discourse analysis, or interpretative phenomenological analysis. They are usually less concerned with drawing general conclusions about human behavior than with understanding in detail the experience of their research participants.

Consider, for example, a study by researcher Per Lindqvist and his colleagues, who wanted to learn how the families of teenage suicide victims cope with their loss (Lindqvist, Johansson, Karlsson, 2008). They did not have a specific research question or hypothesis, such as, What percentage of family members join suicide support groups? Instead, they wanted to understand the variety of reactions that families had, with a focus on what it is like from their perspectives. To address this question, they interviewed the families of 10 teenage suicide victims in their homes in rural Sweden. The interviews were relatively unstructured, beginning with a general request for the families to talk about the victim and ending with an invitation to talk about anything else that they wanted to tell the interviewer. One of the most important themes that emerged from these interviews was that even as life returned to “normal,” the families continued to struggle with the question of why their loved one committed suicide. This struggle appeared to be especially difficult for families in which the suicide was most unexpected.

## **The Purpose of Qualitative Research**

Again, this textbook is primarily about quantitative research in psychology. The strength of quantitative research is its ability to provide precise answers to specific research questions and to draw general conclusions about human behavior. This method is how we know that people have a strong tendency to obey authority figures, for example, and that female undergraduate students are not substantially more talkative than male undergraduate students. But while quantitative research is good at providing precise answers to specific research questions, it is not nearly as good at generating novel and interesting research questions. Likewise, while quantitative research is good at drawing general conclusions about human behavior, it is not nearly as good at providing detailed descriptions of the behavior of particular groups in particular situations. And quantitative research is not very good at communicating what it is actually like to be a member of a particular group in a particular situation.

But the relative weaknesses of quantitative research are the relative strengths of qualitative research. Qualitative research can help researchers to generate new and interesting research questions and hypotheses. The research of Lindqvist et al (2008), for example, suggests that there may be a general relationship between how unexpected a suicide is and how consumed the family is with trying to understand why the teen committed suicide. This relationship can now be explored using quantitative research. But it is unclear whether this question would have arisen at all without the researchers sitting down with the families and listening to what they themselves wanted to say about their experience. Qualitative research can also provide rich and detailed descriptions of human behavior in the real-world contexts in which it occurs. Among qualitative researchers, this depth is often referred to as “thick description” (Geertz, 1973). Similarly, qualitative research can convey a sense of what it is actually like to be a member of a particular group or in a particular situation—what qualitative researchers often refer to as the “lived experience” of the research participants. Lindqvist and colleagues, for example, describe how all the families spontaneously offered to show the interviewer the victim’s bedroom or the place where the suicide occurred—revealing the importance of these physical locations to the families. It seems unlikely that a quantitative study would have discovered this detail.

 Some contrasts between qualitative and quantitative research

|  |  |
| --- | --- |
| Qualitative | Quantitative |
| 1. In-depth information about relatively few people | 1. Less depth information with larger samples |
| 2. Conclusions are based on interpretations drawn by the investigator | 2. Conclusions are based on statistical analyses |
| 3. Global and exploratory | 3. Specific and focused |
|  |  |

## **Data Collection and Analysis in Qualitative Research**

Data collection approaches in qualitative research are quite varied and can involve naturalistic observation, participant observation, archival data, artwork, and many other things. But one of the most common approaches, especially for psychological research, is to conduct interviews. Interviews in qualitative research can be unstructured—consisting of a small number of general questions or prompts that allow participants to talk about what is of interest to them—or structured, where there is a strict script that the interviewer does not deviate from. Most interviews are in between the two and are called semi-structured interviews, where the researcher has a few consistent questions and can follow up by asking more detailed questions about the topics that come up. Such interviews can be lengthy and detailed, but they are usually conducted with a relatively small sample. The unstructured interview was the approach used by Lindqvist and colleagues in their research on the families of suicide victims because the researchers were aware that how much was disclosed about such a sensitive topic should be led by the families, not by the researchers.

Another approach used in qualitative research involves small groups of people who participate together in interviews focused on a particular topic or issue, known as focus groups. The interaction among participants in a focus group can sometimes bring out more information than can be learned in a one-on-one interview. The use of focus groups has become a standard technique in business and industry among those who want to understand consumer tastes and preferences. The content of all focus group interviews is usually recorded and transcribed to facilitate later analyses. However, we know from social psychology that group dynamics are often at play in any group, including focus groups, and it is useful to be aware of those possibilities. For example, the desire to be liked by others can lead participants to provide inaccurate answers that they believe will be perceived favorably by the other participants. The same may be said for personality characteristics. For example, highly extraverted participants can sometimes dominate discussions within focus groups.

## **Data Analysis in Qualitative Research**

Although quantitative and qualitative research generally differ along several important dimensions (e.g., the specificity of the research question, the type of data collected), it is the method of data analysis that distinguishes them more clearly than anything else. To illustrate this idea, imagine a team of researchers that conducts a series of unstructured interviews with people recovering from alcohol use disorder to learn about the role of their religious faith in their recovery. Although this project sounds like qualitative research, imagine further that once they collect the data, they code the data in terms of how often each participant mentions God (or a “higher power”), and they then use descriptive and inferential statistics to find out whether those who mention God more often are more successful in abstaining from alcohol. Now it sounds like quantitative research. In other words, the quantitative-qualitative distinction depends more on what researchers do with the data they have collected than with why or how they collected the data.

But what does qualitative data analysis look like? Just as there are many ways to collect data in qualitative research, there are many ways to analyze data. Here we focus on one general approach called grounded theory (Glaser Strauss, 1967). This approach was developed within the field of sociology in the 1960s and has gradually gained popularity in psychology. Remember that in quantitative research, it is typical for the researcher to start with a theory, derive a hypothesis from that theory, and then collect data to test that specific hypothesis. In qualitative research using grounded theory, researchers start with the data and develop a theory or an interpretation that is “grounded in” those data. They do this analysis in stages. First, they identify ideas that are repeated throughout the data. Then they organize these ideas into a smaller number of broader themes. Finally, they write a theoretical narrative—an interpretation of the data in terms of the themes that they have identified. This theoretical narrative focuses on the subjective experience of the participants and is usually supported by many direct quotations from the participants themselves.

As an example, consider a study Abrams & Curran (2009), who used the grounded theory approach to study the experience of postpartum depression symptoms among low-income mothers. Their data were the result of unstructured interviews with 19 participants with the observed broad themes below in a table. In their research report, they provide numerous quotations from their participants, such as this one from “Destiny:”

Well, just recently my apartment was broken into and the fact that his Medicaid for some reason was cancelled so a lot of things was happening within the last two weeks all at one time. So that in itself I don’t want to say almost drove me mad but it put me in a funk.…Like I really was depressed. (p. 357)

Their theoretical narrative focused on the participants’ experience of their symptoms, not as an abstract “affective disorder” but as closely tied to the daily struggle of raising children alone under often difficult circumstances.

Table of Themes and Repeating Ideas in a Study of Postpartum Depression Among Low-Income Mothers. Based on Research by Abrams and Curran (2009).

|  |  |
| --- | --- |
| **Theme** | **Repeating ideas** |
| Ambivalence | “I wasn’t prepared for this baby,” “I didn’t want to have any more children.” |
| Caregiving overload | “Please stop crying,” “I need a break,” “I can’t do this anymore.” |
| Juggling | “No time to breathe,” “Everyone depends on me,” “Navigating the maze.” |
| Mothering alone | “I really don’t have any help,” “My baby has no father.” |
| Real-life worry | “I don’t have any money,” “Will my baby be OK?” “It’s not safe here.” |
|  |  |

## **The Quantitative-Qualitative “Debate”**

Given their differences, it may come as no surprise that quantitative and qualitative research in psychology and related fields do not coexist in complete harmony. Some quantitative researchers criticize qualitative methods on the grounds that they lack objectivity, are difficult to evaluate in terms of reliability and validity, and do not allow generalization to people or situations other than those actually studied. At the same time, some qualitative researchers criticize quantitative methods on the grounds that they overlook the richness of human behavior and experience and instead answer simple questions about easily quantifiable variables.

However many researchers agree that the two approaches can and should be combined into what has come to be called mixed-methods research (Todd, Nerlich, McKeown, Clarke, 2004). One approach to combining quantitative and qualitative research is to use qualitative research for hypothesis generation and quantitative research for hypothesis testing. Again, while a qualitative study might suggest that families who experience an unexpected suicide have more difficulty resolving the question of why, a well-designed quantitative study could test a hypothesis by measuring these specific variables in a large sample. A second approach to combining quantitative and qualitative research is referred to as triangulation. The idea is to use both quantitative and qualitative methods simultaneously to study the same general questions and to compare the results. If the results of the quantitative and qualitative methods converge on the same general conclusion, they reinforce and enrich each other. If the results diverge, then they suggest an interesting new question: Why do the results diverge and how can they be reconciled?

Using qualitative research can often help clarify quantitative results via triangulation. Trenor, Yu, Waight, Zerda, and Sha (2008) investigated the experience of female engineering students at a university. In the first phase, female engineering students were asked to complete a survey, where they rated a number of their perceptions, including their sense of belonging.  Their results were compared across the student ethnicities, and statistically, the various ethnic groups showed no differences in their ratings of their sense of belonging.  One might look at that result and conclude that ethnicity does not have anything to do with one’s sense of belonging.  However, in the second phase, the authors also conducted interviews with the students, and in those interviews, many minority students reported how the diversity of cultures at the university enhanced their sense of belonging. Without the qualitative component, we might have drawn the wrong conclusion about the quantitative results. This example shows how qualitative and quantitative research work together to help us understand human behavior.

# What Is Observational Research?

The term observational research is used to refer to several different types of non-experimental studies in which behavior is systematically observed and recorded. The goal of observational research is to describe a variable or set of variables. More generally, the goal is to obtain a snapshot of specific characteristics of an individual, group, or setting. As described previously, observational research is non-experimental because nothing is manipulated or controlled, and as such we cannot arrive at causal conclusions using this approach. The data that are collected in observational research studies are often qualitative in nature but they may also be quantitative or both (mixed-methods). There are several different types of observational methods that will be described below.

## Naturalistic Observation

Naturalistic observation is an observational method that involves observing people’s behavior in the environment in which it typically occurs. Thus, naturalistic observation is a type of field research (as opposed to a type of laboratory research). Jane Goodall’s famous research on chimpanzees is a classic example of naturalistic observation. Dr. Goodall spent three decades observing chimpanzees in their natural environment in East Africa. She examined such things as chimpanzee’s social structure, mating patterns, gender roles, family structure, and care of offspring by observing them in the wild. However, naturalistic observation could more simply involve observing shoppers in a grocery store, children on a school playground, or psychiatric inpatients in their wards. Researchers engaged in naturalistic observation usually make their observations as unobtrusively as possible so that participants are not aware that they are being studied. Such an approach is called disguised naturalistic observation. Ethically, this method is considered to be acceptable if the participants remain anonymous and the behavior occurs in a public setting where people would not normally have an expectation of privacy. Grocery shoppers putting items into their shopping carts, for example, are engaged in public behavior that is easily observable by store employees and other shoppers. For this reason, most researchers would consider it ethically acceptable to observe them for a study. On the other hand, one of the arguments against the ethicality of the naturalistic observation of “bathroom behavior” discussed earlier in the book is that people have a reasonable expectation of privacy even in a public restroom and that this expectation was violated.

In cases where it is not ethical or practical to conduct disguised naturalistic observation, researchers can conduct undisguised naturalistic observation where the participants are made aware of the researcher presence and monitoring of their behavior. However, one concern with undisguised naturalistic observation is reactivity. Reactivity refers to when a measure changes participants’ behavior. In the case of undisguised naturalistic observation, the concern with reactivity is that when people know they are being observed and studied, they may act differently than they normally would. This type of reactivity is known as the Hawthorne effect. For instance, you may act much differently in a bar if you know that someone is observing you and recording your behaviors and this would invalidate the study. So disguised observation is less reactive and therefore can have higher validity because people are not aware that their behaviors are being observed and recorded. However, we now know that people often become used to being observed and with time they begin to behave naturally in the researcher’s presence. In other words, over time people habituate to being observed. Think about reality shows like Big Brother or Survivor where people are constantly being observed and recorded. While they may be on their best behavior at first, in a fairly short amount of time they are flirting, having sex, wearing next to nothing, screaming at each other, and occasionally behaving in ways that are embarrassing.

## **Participant Observation**

Another approach to data collection in observational research is participant observation. In participant observation, researchers become active participants in the group or situation they are studying. Participant observation is very similar to naturalistic observation in that it involves observing people’s behavior in the environment in which it typically occurs. As with naturalistic observation, the data that are collected can include interviews (usually unstructured), notes based on their observations and interactions, documents, photographs, and other artifacts. The only difference between naturalistic observation and participant observation is that researchers engaged in participant observation become active members of the group or situations they are studying. The basic rationale for participant observation is that there may be important information that is only accessible to, or can be interpreted only by, someone who is an active participant in the group or situation. Like naturalistic observation, participant observation can be either disguised or undisguised. In disguised participant observation, the researchers pretend to be members of the social group they are observing and conceal their true identity as researchers.

In a famous example of disguised participant observation, Leon Festinger and his colleagues infiltrated a doomsday cult known as the Seekers, whose members believed that the apocalypse would occur on December 21, 1954. Interested in studying how members of the group would cope psychologically when the prophecy inevitably failed, they carefully recorded the events and reactions of the cult members in the days before and after the supposed end of the world. Unsurprisingly, the cult members did not give up their belief but instead convinced themselves that it was their faith and efforts that saved the world from destruction. Festinger and his colleagues later published a book about this experience, which they used to illustrate the theory of cognitive dissonance (Festinger, Riecken, Schachter, 1956).

In contrast with undisguised participant observation, the researchers become a part of the group they are studying and they disclose their true identity as researchers to the group under investigation. Once again there are important ethical issues to consider with disguised participant observation. First no informed consent can be obtained and second deception is being used. The researcher is deceiving the participants by intentionally withholding information about their motivations for being a part of the social group they are studying. But sometimes disguised participation is the only way to access a protective group (like a cult). Further, disguised participant observation is less prone to reactivity than undisguised participant observation.

Rosenhan’s study (1973) of the experience of people in a psychiatric ward would be considered disguised participant observation because Rosenhan and his pseudopatients were admitted into psychiatric hospitals on the pretense of being patients so that they could observe the way that psychiatric patients are treated by staff. The staff and other patients were unaware of their true identities as researchers.

Another example of participant observation comes from a study by sociologist Amy Wilkins on a university-based religious organization that emphasized how happy its members were (Wilkins, 2008). Wilkins spent 12 months attending and participating in the group’s meetings and social events, and she interviewed several group members. In her study, Wilkins identified several ways in which the group “enforced” happiness—for example, by continually talking about happiness, discouraging the expression of negative emotions, and using happiness as a way to distinguish themselves from other groups.

One of the primary benefits of participant observation is that the researchers are in a much better position to understand the viewpoint and experiences of the people they are studying when they are a part of the social group. The primary limitation with this approach is that the mere presence of the observer could affect the behavior of the people being observed. While this is also a concern with naturalistic observation, additional concerns arise when researchers become active members of the social group they are studying because that they may change the social dynamics and/or influence the behavior of the people they are studying. Similarly, if the researcher acts as a participant observer there can be concerns with biases resulting from developing relationships with the participants. Concretely, the researcher may become less objective resulting in more experimenter bias.

## **Structured Observation**

Another observational method is structured observation. Here the investigator makes careful observations of one or more specific behaviors in a particular setting that is more structured than the settings used in naturalistic or participant observation. Often the setting in which the observations are made is not the natural setting. Instead, the researcher may observe people in the laboratory environment. Alternatively, the researcher may observe people in a natural setting (like a classroom setting) that they have structured some way, for instance by introducing some specific task participants are to engage in or by introducing a specific social situation or manipulation.

Structured observation is very similar to naturalistic observation and participant observation in that in all three cases researchers are observing naturally occurring behavior; however, the emphasis in structured observation is on gathering quantitative rather than qualitative data. Researchers using this approach are interested in a limited set of behaviors. This allows them to quantify the behaviors they are observing. In other words, structured observation is less global than naturalistic or participant observation because the researcher engaged in structured observations is interested in a small number of specific behaviors. Therefore, rather than recording everything that happens, the researcher only focuses on very specific behaviors of interest.

As an example, researchers Robert Kraut and Robert Johnston wanted to study bowlers’ reactions to their shots, both when they were facing the pins and then when they turned toward their companions (Kraut & Johnston, 1979). But what “reactions” should they observe? Based on previous research and their own pilot testing, Kraut and Johnston created a list of reactions that included “closed smile,” “open smile,” “laugh,” “neutral face,” “look down,” “look away,” and “face cover” (covering one’s face with one’s hands). The observers committed this list to memory and then practiced by coding the reactions of bowlers who had been videotaped. During the actual study, the observers spoke into an audio recorder, describing the reactions they observed. Among the most interesting results of this study was that bowlers rarely smiled while they still faced the pins. They were much more likely to smile after they turned toward their companions, suggesting that smiling is not purely an expression of happiness but also a form of social communication.

In another example (this one in a laboratory environment), Dov Cohen and his colleagues had observers rate the emotional reactions of participants who had just been deliberately bumped and insulted by a confederate after they dropped off a completed questionnaire at the end of a hallway. The confederate was posing as someone who worked in the same building and who was frustrated by having to close a file drawer twice in order to permit the participants to walk past them (first to drop off the questionnaire at the end of the hallway and once again on their way back to the room where they believed the study they signed up for was taking place). The two observers were positioned at different ends of the hallway so that they could read the participants’ body language and hear anything they might say. Interestingly, the researchers hypothesized that participants from the southern United States, which is one of several places in the world that has a “culture of honor,” would react with more aggression than participants from the northern United States, a prediction that was in fact supported by the observational data (Cohen, Nisbett, Bowdle, Schwarz, 1996).

When the observations require a judgment on the part of the observers—as in the studies by Kraut and Johnston and Cohen and his colleagues—a process referred to as coding is typically required. Coding generally requires clearly defining a set of target behaviors. The observers then categorize participants individually in terms of which behavior they have engaged in and the number of times they engaged in each behavior. The observers might even record the duration of each behavior. The target behaviors must be defined in such a way that guides different observers to code them in the same way. Researchers are expected to demonstrate the interrater reliability of their coding procedure by having multiple raters code the same behaviors independently and then showing that the different observers are in close agreement. Kraut and Johnston, for example, video recorded a subset of their participants’ reactions and had two observers independently code them. The two observers showed that they agreed on the reactions that were exhibited 97% of the time, indicating good interrater reliability.

One of the primary benefits of structured observation is that it is far more efficient than naturalistic and participant observation. Since the researchers are focused on specific behaviors this reduces time and expense. Also, often times the environment is structured to encourage the behaviors of interest which again means that researchers do not have to invest as much time in waiting for the behaviors of interest to naturally occur. Finally, researchers using this approach can clearly exert greater control over the environment. However, when researchers exert more control over the environment it may make the environment less natural which decreases external validity. It is less clear for instance whether structured observations made in a laboratory environment will generalize to a real-world environment. Furthermore, since researchers engaged in structured observation are often not disguised there may be more concerns with reactivity.

## Key Takeaways and Exercises

* Non-experimental research is research that lacks the manipulation of an independent variable.
* There are two broad types of non-experimental research. Correlational research that focuses on statistical relationships between variables that are measured but not manipulated; and observational research in which participants are observed and their behavior is recorded without the researcher interfering or manipulating any variables.
* In general, experimental research is high in internal validity, correlational research is low in internal validity, and quasi-experimental research is in between.
* Correlational research involves measuring two variables and assessing the relationship between them, with no manipulation of an independent variable.
* Correlation does not imply causation. A statistical relationship between two variables, X and Y, does not necessarily mean that X causes Y. It is also possible that Y causes X, or that a third variable, Z, causes both X and Y.
* While correlational research cannot be used to establish causal relationships between variables, correlational research does allow researchers to achieve many other important objectives (establishing reliability and validity, providing converging evidence, describing relationships, and making predictions)
* Correlation coefficients can range from -1 to +1. The sign indicates the direction of the relationship between the variables and the numerical value indicates the strength of the relationship.
* Researchers often use complex correlational research to explore relationships among several variables in the same study.
* Complex correlational research can be used to explore possible causal relationships among variables using techniques such as partial correlation and multiple regression. Such designs can show patterns of relationships that are consistent with some causal interpretations and inconsistent with others, but they cannot unambiguously establish that one variable causes another.
* Qualitative research is an important alternative to quantitative research in psychology. It generally involves asking broader research questions, collecting more detailed data (e.g., interviews), and using non-statistical analyses.
* Many researchers conceptualize quantitative and qualitative research as complementary and advocate combining them. For example, qualitative research can be used to generate hypotheses and quantitative research to test them.
* There are several different approaches to observational research including naturalistic observation, participant observation, structured observation, case studies, and archival research.
* Naturalistic observation is used to observe people in their natural setting; participant observation involves becoming an active member of the group being observed; structured observation involves coding a small number of behaviors in a quantitative manner; case studies are typically used to collect in-depth information on a single individual; and archival research involves analyzing existing data.

## Exercises

* Discussion: For each of the following studies, decide which type of research design it is and explain why.
  + A researcher conducts detailed interviews with unmarried teenage fathers to learn about how they feel and what they think about their role as fathers and summarizes their feelings in a written narrative.
  + A researcher measures the impulsivity of a large sample of drivers and looks at the statistical relationship between this variable and the number of traffic tickets the drivers have received.
  + A researcher randomly assigns patients with low back pain either to a treatment involving hypnosis or to a treatment involving exercise. She then measures their level of low back pain after 3 months.
* Discussion: For each of the following, decide whether it is most likely that the study described is experimental or non-experimental and explain why.
  + A cognitive psychologist compares the ability of people to recall words that they were instructed to “read” with their ability to recall words that they were instructed to “imagine.”
  + A manager studies the correlation between new employees’ college grade point averages and their first-year performance reports.
  + An automotive engineer installs different stick shifts in a new car prototype, each time asking several people to rate how comfortable the stick shift feels.
  + A food scientist studies the relationship between the temperature inside people’s refrigerators and the amount of bacteria on their food.
  + A social psychologist tells some research participants that they need to hurry over to the next building to complete a study. She tells others that they can take their time. Then she observes whether they stop to help a research assistant who is pretending to be hurt.
* Practice: For each of the following statistical relationships, decide whether the directionality problem is present and think of at least one plausible third variable.
  + People who eat more lobster tend to live longer.
  + People who exercise more tend to weigh less.
  + College students who drink more alcohol tend to have poorer grades.
* Practice: Construct a correlation matrix for a hypothetical study including the variables of depression, anxiety, self-esteem, and happiness. Include the Pearson’s r values that you would expect.
* Discussion: Imagine a correlational study that looks at intelligence, the need for cognition, and high school students’ performance in a critical thinking course. A multiple regression analysis shows that intelligence is not related to performance in the class but that the need for cognition is. Explain what this study has shown in terms of what is related to good performance in the critical thinking course.
* Discussion: What are some ways in which a qualitative study of girls who play youth baseball would likely differ from a quantitative study on the same topic? How would the data differ by interviewing girls one-on-one rather than conducting focus groups or surveys?
* Practice: Find and read a published case study in psychology. (Use case study as a key term in a PsycINFO search.) Then do the following:
  + Describe one problem related to internal validity.
  + Describe one problem related to external validity.
  + Generate one hypothesis suggested by the case study that might be interesting to test in a subsequent study.

## References

Abrams, L. S., Curran, L. (2009). “And you’re telling me not to stress?” A grounded theory study of postpartum depression symptoms among low-income mothers, 351–362.

Bryman, A. (2012), 4th ed. Oxford: OUP.

Bushman, B. J., Huesmann, L. R. (2001). Effects of televised violence on aggression. In D. Singer J. Singer (Eds.), (pp. 223–254). Thousand Oaks, CA: Sage.

Cacioppo, J. T., Petty, R. E. (1982). The need for cognition, 116–131.

Diener, E. (2000). Subjective well-being: The science of happiness, and a proposal for a national index, 34–43.

Geertz, C. (1973). New York, NY: Basic Books.

Glaser, B. G., Strauss, A. L. (1967). Chicago, IL: Aldine.

Jouriles, E. N., Garrido, E., Rosenfield, D., McDonald, R. (2009). Experiences of psychological and physical aggression in adolescent romantic relationships: Links to psychological distress(7), 451–460.

Lindqvist, P., Johansson, L., Karlsson, U. (2008). In the aftermath of teenage suicide: A qualitative study of the psychosocial consequences for the surviving family members, 26. Retrieved from <http://www.biomedcentral.com/1471-244X/8/26>

Messerli, F. H. (2012). Chocolate consumption, cognitive function, and Nobel laureates, 1562-1564.

Milgram, S. (1963). Behavioral study of obedience, 371–378.

Loftus, E. F., Pickrell, J. E. (1995). The formation of false memories, 720–725.

Plomin, R., DeFries, J. C., McClearn, G. E., McGuffin, P. (2008). (5th ed.). New York, NY: Worth.

Rentfrow, P. J., Gosling, S. D. (2008). The do re mi’s of everyday life: The structure and personality correlates of music preferences, 1236–1256.

Rosenhan, D. L. (1973). On being sane in insane places, 250–258.

Todd, Z., Nerlich, B., McKeown, S., Clarke, D. D. (2004). London, UK: Psychology Press.

Trenor, J.M., Yu, S.L., Waight, C.L., Zerda. K.S Sha T.-L. (2008). The relations of ethnicity to female engineering students’ educational experiences and college and career plans in an ethnically diverse learning environment, (4), 449-465.