

6.5 OBSERVATIONAL RESEARCH

Learning Objectives

1. List the various types of observational research methods and distinguish between each
2. Describe the strengths and weakness of each observational research method.

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 research designs 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. 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 at times acting like complete fools in front of the entire nation.

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 is 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 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 passive deception is being used. The researcher is passively 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)^[1] 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 (published in *Social Psychology Quarterly*) on a university-based religious organization that emphasized how happy its members were (Wilkins, 2008)^[2]. 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 researcher is in a much better position to understand the viewpoint and experiences of the

people they are studying when they are apart 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 when researchers because active members of the social group they are studying, additional concerns arise 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 and participant observation. Often the setting in which the observations are made is not the natural setting, rather 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 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 and 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.

Structured observation is very similar to naturalistic observation and participant observation in that in all 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 and 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.

Researchers Robert Levine and Ara Norenzayan used structured observation to study differences in the “pace of life” across countries (Levine & Norenzayan, 1999)^[3]. One of their measures involved observing pedestrians in a large city to see how long it took them to walk 60 feet. They found that people in some countries walked reliably faster than people in other countries. For example, people in Canada and Sweden covered 60 feet in just under 13 seconds on average, while people in Brazil and Romania took close to 17 seconds. When structured observation takes place in the complex and even chaotic “real world,” the questions of when, where, and under what conditions the observations will be made, and who exactly will be observed are important to consider. Levine and Norenzayan described their sampling process as follows:

“Male and female walking speed over a distance of 60 feet was measured in at least two locations in main downtown areas in each city. Measurements were taken during main business hours on clear summer days. All locations were flat, unobstructed, had broad sidewalks, and were sufficiently uncrowded to allow pedestrians to move at potentially maximum speeds. To control for the effects of socializing, only pedestrians walking alone were used. Children, individuals with obvious physical handicaps, and window-shoppers were not timed. Thirty-five men and 35 women were timed in most cities.” (p. 186). Precise specification of the sampling process in this way makes data collection manageable for the observers, and it also provides some control over important extraneous variables. For example, by making their observations on clear summer days in all countries, Levine and Norenzayan controlled for effects of the weather on people’s walking speeds. In Levine and Norenzayan’s study, measurement was relatively straightforward. They simply measured out a 60-foot distance along a city sidewalk and then used a stopwatch to time participants as they walked over that distance.

As another 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)^[4]. 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.

When the observations require a judgment on the part of the observers—as in Kraut and Johnston’s study—this process is often described as **coding**. 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 different observers code them in the same way. This difficulty with coding is the issue of interrater reliability, as mentioned in Chapter 4. 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.

CASE STUDIES

A **case study** is an in-depth examination of an individual. Sometimes case studies are also completed on social units (e.g., a cult) and events (e.g., a natural disaster). Most commonly in psychology, however, case studies provide a detailed description and analysis of an individual. Often the individual has a rare or unusual condition or disorder or has damage to a specific region of the brain.

Like many observational research methods, case studies tend to be more qualitative in nature. Case study methods involve an in-depth, and often a longitudinal examination of an individual. Depending on the focus of the case study, individuals may or may not be observed in their natural setting. If the natural setting is not what is of interest, then the individual may be brought into a therapist's office or a researcher's lab for study. Also, the bulk of the case study report will focus on in-depth descriptions of the person rather than on statistical analyses. With that said some quantitative data may also be included in the write-up of a case study. For instance, an individual's depression score may be compared to normative scores or their score before and after treatment may be compared. As with other qualitative methods, a variety of different methods and tools can be used to collect information on the case. For instance, interviews, naturalistic observation, structured observation, psychological testing (e.g., IQ test), and/or physiological measurements (e.g., brain scans) may be used to collect information on the individual.

HM is one of the most notorious case studies in psychology. HM suffered from intractable and very severe epilepsy. A surgeon localized HM's epilepsy to his medial temporal lobe and in 1953 he removed large sections of his hippocampus in an attempt to stop the seizures. The treatment was a success, in that it resolved his epilepsy and his IQ and personality were unaffected. However, the doctors soon realized that HM exhibited a strange form of amnesia, called anterograde amnesia. HM was able to carry out a conversation and he could remember short strings of letters, digits, and words. Basically, his short term memory was preserved. However, HM could not commit new events to memory. He lost the ability to transfer information from his short-term memory to his long term memory, something memory researchers call consolidation. So while he could carry on a conversation with someone, he would completely forget the conversation after it ended. This was an extremely important case study for memory researchers because it suggested that there's a dissociation between short-term memory and long-term memory, it suggested that these were two different abilities sub-served by different areas of the brain. It also suggested that the temporal lobes are particularly

important for consolidating new information (i.e., for transferring information from short-term memory to long-term memory).

www.youtube.com/watch?v=KkaXNvzE4pk

The history of psychology is filled with influential cases studies, such as Sigmund Freud's description of "Anna O." (see Note 6.1 "The Case of "Anna O."") and John Watson and Rosalie Rayner's description of Little Albert (Watson & Rayner, 1920)^[5], who learned to fear a white rat—along with other furry objects—when the researchers made a loud noise while he was playing with the rat.

The Case of "Anna O."

Sigmund Freud used the case of a young woman he called "Anna O." to illustrate many principles of his theory of psychoanalysis (Freud, 1961)^[6]. (Her real name was Bertha Pappenheim, and she was an early feminist who went on to make important contributions to the field of social work.) Anna had come to Freud's colleague Josef Breuer around 1880 with a variety of odd physical and psychological symptoms. One of them was that for several weeks she was unable to drink any fluids. According to Freud,

She would take up the glass of water that she longed for, but as soon as it touched her lips she would push it away like someone suffering from hydrophobia....She lived only on fruit, such as melons, etc., so as to lessen her tormenting thirst. (p. 9)

But according to Freud, a breakthrough came one day while Anna was under hypnosis.

[S]he grumbled about her English "lady-companion," whom she did not care for, and went on to describe, with every sign of disgust, how she had once gone into this lady's room and how her little dog—horrid creature!—had drunk out of a glass there. The patient had said nothing, as she had wanted to be polite. After giving further energetic expression to the anger she had held back, she asked for something to drink, drank a large quantity of water without any difficulty, and awoke from her hypnosis with the glass at her lips; and thereupon the disturbance vanished, never to return. (p.9)

Freud's interpretation was that Anna had repressed the memory of this incident along with the emotion that it triggered and that this was what had caused her inability to drink. Furthermore, her recollection of the incident,

along with her expression of the emotion she had repressed, caused the symptom to go away.

As an illustration of Freud's theory, the case study of Anna O. is quite effective. As evidence for the theory, however, it is essentially worthless. The description provides no way of knowing whether Anna had really repressed the memory of the dog drinking from the glass, whether this repression had caused her inability to drink, or whether recalling this "trauma" relieved the symptom. It is also unclear from this case study how typical or atypical Anna's experience was.



Figure 10.1 Anna O. "Anna O." was the subject of a famous case study used by Freud to illustrate the principles of psychoanalysis. Source: http://en.wikipedia.org/wiki/File:Pappenheim_1882.jpg

Case studies are useful because they provide a level of detailed analysis not found in many other research methods and greater insights may be gained from this more detailed analysis. As a result of the case study, the researcher may gain a sharpened understanding of what might become important to look at more extensively in future more controlled research. Case studies are also often the only way to study rare conditions because it may be impossible to find a large enough sample to individuals with the condition to use quantitative methods. Although at first glance a case study of a rare individual might seem to tell us little about ourselves, they often do provide insights into normal behavior. The case of HM provided important insights into the role of the hippocampus in memory consolidation. However, it is important to note that while case studies can provide *insights* into certain areas and variables to study, and can be useful in helping develop theories, they should never be used as evidence for theories. In other words, case studies can be used as

inspiration to formulate theories and hypotheses, but those hypotheses and theories then need to be formally tested using more rigorous quantitative methods.

The reason case studies shouldn't be used to provide support for theories is that they suffer from problems with internal and external validity. Case studies lack the proper controls that true experiments contain. As such they suffer from problems with internal validity, so they cannot be used to determine causation. For instance, during HM's surgery, the surgeon may have accidentally lesioned another area of HM's brain (indeed questioning into the possibility of a separate brain lesion began after HM's death and dissection of his brain) and that lesion may have contributed to his inability to consolidate new information. The fact is, with case studies we cannot rule out these sorts of alternative explanations. So as with all observational methods case studies do not permit determination of causation. In addition, because case studies are often of a single individual, and typically a very abnormal individual, researchers cannot generalize their conclusions to other individuals. Recall that with most research designs there is a trade-off between internal and external validity, with case studies, however, there are problems with both internal validity and external validity. So there are limits both to the ability to determine causation and to generalize the results. A final limitation of case studies is that ample opportunity exists for the theoretical biases of the researcher to color or bias the case description. Indeed, there have been accusations that the woman who studied HM destroyed a lot of her data that were not published and she has been called into question for destroying contradictory data that didn't support her theory about how memories are consolidated. There is a fascinating New York Times article that describes some of the controversies that ensued after HM's death and analysis of his brain that can be found at:

https://www.nytimes.com/2016/08/07/magazine/the-brain-that-couldnt-remember.html?_r=0

ARCHIVAL RESEARCH

Another approach that is often considered observational research is the use of **archival research** which involves analyzing data that have already been collected for some other purpose. An example is a study by Brett Pelham and his colleagues on "implicit egotism"—the tendency for people to prefer people, places, and things that are similar to themselves (Pelham, Carvallo, & Jones, 2005)^[7]. In one study, they examined Social Security records to show that women with the names Virginia, Georgia, Louise, and Florence were especially

likely to have moved to the states of Virginia, Georgia, Louisiana, and Florida, respectively.

As with naturalistic observation, measurement can be more or less straightforward when working with archival data. For example, counting the number of people named Virginia who live in various states based on Social Security records is relatively straightforward. But consider a study by Christopher Peterson and his colleagues on the relationship between optimism and health using data that had been collected many years before for a study on adult development (Peterson, Seligman, & Vaillant, 1988)^[8]. In the 1940s, healthy male college students had completed an open-ended questionnaire about difficult wartime experiences. In the late 1980s, Peterson and his colleagues reviewed the men's questionnaire responses to obtain a measure of explanatory style—their habitual ways of explaining bad events that happen to them. More pessimistic people tend to blame themselves and expect long-term negative consequences that affect many aspects of their lives, while more optimistic people tend to blame outside forces and expect limited negative consequences. To obtain a measure of explanatory style for each participant, the researchers used a procedure in which all negative events mentioned in the questionnaire responses, and any causal explanations for them were identified and written on index cards. These were given to a separate group of raters who rated each explanation in terms of three separate dimensions of optimism-pessimism. These ratings were then averaged to produce an explanatory style score for each participant. The researchers then assessed the statistical relationship between the men's explanatory style as undergraduate students and archival measures of their health at approximately 60 years of age. The primary result was that the more optimistic the men were as undergraduate students, the healthier they were as older men. Pearson's r was $+.25$.

This method is an example of **content analysis**—a family of systematic approaches to measurement using complex archival data. Just as structured observation requires specifying the behaviors of interest and then noting them as they occur, content analysis requires specifying keywords, phrases, or ideas and then finding all occurrences of them in the data. These occurrences can then be counted, timed (e.g., the amount of time devoted to entertainment topics on the nightly news show), or analyzed in a variety of other ways.

Key Takeaways

- 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 analysing existing data.

Exercises

1. 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 systematic single-subject or group study.

-
1. Rosenhan, D. L. (1973). On being sane in insane places. *Science*, 179, 250–258. ↵
 2. Wilkins, A. (2008). “Happier than Non-Christians”: Collective emotions and symbolic boundaries among evangelical Christians. *Social Psychology Quarterly*, 71, 281–301. ↵
 3. Levine, R. V., & Norenzayan, A. (1999). The pace of life in 31 countries. *Journal of Cross-Cultural Psychology*, 30, 178–205. ↵
 4. Kraut, R. E., & Johnston, R. E. (1979). Social and emotional messages of smiling: An ethological approach. *Journal of Personality and Social Psychology*, 37, 1539–1553. ↵
 5. Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology*, 3, 1–14. ↵
 6. Freud, S. (1961). *Five lectures on psycho-analysis*. New York, NY: Norton. ↵
 7. Pelham, B. W., Carvallo, M., & Jones, J. T. (2005). Implicit egotism. *Current Directions in Psychological Science*, 14, 106–110. ↵
 8. Peterson, C., Seligman, M. E. P., & Vaillant, G. E. (1988). Pessimistic explanatory style is a risk factor for physical illness: A thirty-five year longitudinal study. *Journal of Personality and Social Psychology*, 55, 23–27. ↵