5 Causation and Research Design

Nomothetic Causal Explanation

Experimental Design and the Criteria for Causal Explanation
Association
Time Order
Nonspuriousness
Mechanism
Context

Nonexperimental Designs and the Criteria for Causal Explanation
Cross-Sectional Designs
Longitudinal Designs
  Repeated Cross-Sectional Designs
  Fixed-Sample Panel Designs
  Event-Based Designs

Units of Analysis and Errors in Causal Reasoning
Individual and Group Units of Analysis
The Ecological Fallacy and Reductionism

Causation in Qualitative Research

Conclusion

KEY TERMS
HIGHLIGHTS
EXERCISES
DEVELOPING A RESEARCH PROPOSAL
WEB EXERCISES
SPSS EXERCISES
Identifying causes—figuring out why things happen—is the goal of most social science research, as well as a critical interest of newspaper reporters, government officials, and ordinary citizens. Unfortunately, valid explanations of the causes of social phenomena do not come easily. Why did the homicide rate drop for 15 years and then start to rise in 1999 (Butterfield, 2000:12)? Changes in the style of policing (Radin, 1997:B7)? Changing attitudes among young people (Butterfield, 1996a)? Variation in patterns of drug use (Krauss, 1996)? Tougner prison sentences (Butterfield, 1996a)? More stringent hand gun regulations (Butterfield, 1996b)? In order to distinguish these possibilities we must design our research strategies carefully. This chapter considers the meaning of causation, the criteria for achieving causally valid explanations, and the ways in which different research designs seek to meet these criteria. I will focus attention on the differences between the experimental and nonexperimental approaches to causation. I will also contrast the way in which many quantitative researchers think about causation with a different approach that is more often used by qualitative researchers. I will focus special attention on the problem of time order—establishing whether the cause truly precedes the effect—and on the identification of units of analysis, which is a prerequisite for properly stating causal conclusions. By the end of the chapter, you should have a good grasp of the meaning of causation and be able to ask the right questions to determine whether causal inferences are likely to be valid. And perhaps you will have another answer or two about the causes of crime and violence.

**Nomothetic Causal Explanation**

A cause is an explanation for some characteristic, attitude, or behavior of groups, individuals, or other entities (such as families, organizations, or cities) or for events. For example, a sociologist may seek to explain adults’ propensities to commit crimes. One explanation, or cause, may be that the experience of abuse during childhood creates a predisposition for crime. You should recognize that there’s a hypothesis here: Adults who experienced abuse as children are more likely to commit crimes than those who did not experience abuse. (Can you identify the dependent and independent variables?)

This is a nomothetic causal explanation; it means that we believe that variation in the independent variable will be followed by variation in the dependent variable, when all other things are equal ("ceteris paribus"). I admit that you can legitimately argue that "all" other things can't literally be equal: We will not be able to compare the same people at the same time in exactly the same circumstances except for the variation in the independent variable (King, Keohane, & Verba, 1994). However, you will see that we can design research to create conditions that are very comparable, so
that we can isolate the impact of the independent variable on the dependent variable.

Quantitative researchers seek nomothetic causal explanations, whether they use experimental or nonexperimental research designs. However, the way in which experimental designs attempt to identify causes differs quite a bit from the way in which nonexperimental designs attempt to identify causes. I will discuss the experimental approach first.

Causal effect (nomothetic perspective) The finding that change in one variable leads to change in another variable, ceteris paribus (other things being equal).

Example of a nomothetic causal explanation: Individuals arrested for domestic assault tend to commit fewer subsequent assaults than similar individuals who are accused in the same circumstances but not arrested.

Experimental Design and the Criteria for Causal Explanation

Feeling angry: Get it out of your system: run on a pillow or yell at the wind and then you won't feel like directing some aggressive act at other people. It's called catharsis, and it's an idea that has almost achieved the status of folk wisdom. Moreover, catharsis theory is supported by the respected Freudian "hydraulic model" of anger; this theory expects frustration-induced anger to build unless it is released (Bushman, Baumeister, & Stack, 1999:367-368). Have you ever experienced what felt like catharsis yourself? Have you seen evidence of it in other people? Does it just make sense to you? Does it come as a surprise to learn that social science research has yielded little evidence in support of catharsis theory? Could you have overgeneralized from your experiences? Selectively observed others? Reasoned illogically?

Five criteria must be considered when deciding whether a causal connection exists. Research designs that allow us to establish these criteria require very careful planning, implementation, and analysis. Many times, researchers have to leave one or more of the criteria unmet and therefore are left with some important doubts about the validity of their causal conclusions; or they may avoid even making any causal assertions. The first three of the criteria are generally considered the most important bases for identifying a nomothetic causal effect: empirical association, appropriate time order, and nonspuriousness. Evidence that meets the other two criteria—identifying a causal mechanism and specifying the context in which the effect occurs—can considerably strengthen causal explanations.
Brad Bushman, Roy Baumeister, and Angela Stack's 1999 study of the effect of catharsis on aggression illustrates how an experimental approach can be used to meet the five criteria for establishing causal relationships. Bushman and his colleagues recruited 707 undergraduate students from introductory psychology courses. The experiment itself had several stages, but I will focus on only a few (see Exhibit 5.1). First, students were told they were in a study of the accuracy of perceptions in social interactions. They gave their consent to participate and were then asked to read either a statement that endorsed the catharsis effect (on reducing aggression), a statement that disputed the catharsis effect, or no statement. Next, each student wrote a short essay on abortion, which was then evaluated by a student in another room. The evaluations, all of them very negative, were returned to the students. At this point, all the students were invited to hit a punching bag, alone, for two minutes.

Now the students were told they were to engage in a competitive reaction-time task. This task consisted of trying to press a button faster than a partner. Each time the student pressed the button faster, they were able to "blast" their competitor with a noise that was as loud and long as they liked (within limits). It turned out that students who had read the pro-catharsis message "blasted" their competitors more than those who had read no message, while those who had read the anti-catharsis message blasted their competitors the least (see Exhibit 5.2). Bushman and his colleagues concluded that reading a pro-catharsis message increased rather than decreased interpersonal aggression.

Was this causal conclusion justified? How confident can we be in its internal validity? I will answer this question by reviewing how the experiment attempted to meet each of the causal criteria; this review will also clarify why the study had so many different components and what the key features of a "true experiment" are. You will then consider how nonexperimental designs attempt to meet the causal criteria. We will study variations
Brad Bushman, Roy Baumeister, and Angela Stack's 1999 study of the effect of catharsis on aggression illustrates how an experimental approach can be used to meet the five criteria for establishing causal relationships. Bushman and his colleagues recruited 707 undergraduate students from introductory psychology courses. The experiment itself had several stages, but I will focus on only a few (see Exhibit 5.1). First, students were told they were in a study of the accuracy of perceptions in social interactions. They gave their consent to participate and were then asked to read either a statement that endorsed the catharsis effect (on reducing aggression), a statement that disputed the catharsis effect, or no statement. Next, each student wrote a short essay on abortion, which was then evaluated by a student in another room. The evaluations, all of them very negative, were returned to the students. At this point, all the students were invited to hit a punching bag, alone, for two minutes.

Now the students were told they were to engage in a competitive reaction-time task. This task consisted of trying to press a button faster than a partner. Each time the student pressed the button faster, they were able to "blast" their competitor with a noise that was as loud and long as they liked (within limits). It turned out that students who had read the pro-catharsis message "blasted" their competitors more than did those who had read no message, while those who had read the anti-catharsis message blasted their competitors the least (see Exhibit 5.2). Bushman and his colleagues concluded that reading a pro-catharsis message increased rather than decreased interpersonal aggression.

Was this causal conclusion justified? How confident can we be in its internal validity? I will answer this question by reviewing how the experimental attempt to meet each of the causal criteria; this review will also clarify why the study had so many different components and what the key features of a "true experiment" are. You will then consider how non-experimental designs attempt to meet the causal criteria.
Exhibit 5.2  Interpersonal Aggression as a Function of Message Content

Source: Bushman et al., 1999:372.

on experimental design in Chapter 6 and you will explore nonexperimental designs in Chapters 7, 8, and 9.

Association

We say that there was an association between interpersonal aggression and type of message because the level of interpersonal aggression varied according to the type of message. An empirical (or observed) association between the independent and dependent variables is the first criterion for identifying a nomothetic causal effect.

We can determine whether an association exists between the independent and dependent variables in a true experiment because there are two or more groups that differ in terms of their value on the independent variable. One group receives some “treatment,” such as reading a cathartic message, that is a manipulation of the value of the independent variable. This group is termed the “experimental group.” In a simple experiment, there may be one other group that does not receive the treatment; it is termed the “control group.” The Bushman study, as I have described it, compared three groups; other experiments may compare more groups that represent multiple values of the independent variable or even combinations of the values of two or more independent variables. In fact, Bushman actually compared groups that differed in having direct or displaced targets for their aggression (see Exhibit 5.3).
Time Order

Association is a necessary criterion for establishing a causal effect, but it is not sufficient. We must also ensure that the variation in the dependent variable occurred after the variation in the independent variable. This is the criterion of time order. In a true experiment, the time order is determined by the researcher. Bushman and his colleagues first exposed the students to the different messages and then measured their level of interpersonal aggression. If we find an association between the types of messages people have read and their aggressiveness outside of an experimental situation, the criterion of time order may not be met. People who are more inclined to interpersonal aggression may be more likely than others to read messages that encourage displays of aggressiveness. This would result in an association between type of message and interpersonal aggression, but the association would reflect the influence of aggression on type of message rather than the other way around.

Nonspuriousness

Another essential criterion for establishing the existence of a causal effect of an independent variable on a dependent variable is nonspuriousness. We say that a relationship between two variables is not spurious when it is not due to variation in a third variable. Have you heard the old adage “Correlation does not prove causation”? It is meant to remind us that an association between two variables might be caused by something else. If we measure children’s shoe sizes and their academic knowledge, for example, we will find a positive association. However, the association may result from the fact that older children have larger feet as well as more academic knowledge. Shoe size does not cause knowledge or vice versa.

Do storks bring babies? If you believe that correlation proves causation, then you might think so. The more storks that appear in certain districts in Holland, the more babies are born. But the association in Holland between number of storks and number of babies is spurious. In fact, both the number of storks and the birth rate are higher in rural districts than in urban

---

**Exhibit 5.3 Experimental Conditions: A “3 × 2” Design**

<table>
<thead>
<tr>
<th>Message Content (3):</th>
<th>Procatharsis</th>
<th>Procatharsis</th>
<th>Procatharsis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggressiveness Target (2):</td>
<td>Direct</td>
<td>Displaced</td>
<td>Direct</td>
</tr>
<tr>
<td></td>
<td>$X_1$</td>
<td>$X_2$</td>
<td>$X_3$</td>
</tr>
</tbody>
</table>

Source: Bushman et al., 1999:371.
The rural or urban character of the districts (the extraneous variable) causes variation in the other two variables.

If you think this point is obvious, consider a social science example. Do schools with better resources produce better students? Before you answer the question, consider the fact that parents with more education and higher income tend to live in neighborhoods that spend more on their schools. These parents are also more likely to have books in the home and provide other advantages for their children. Do the parents cause variation in both school resources and student performance? If so, there would be an association between school resources and student performance that was at least partially spurious.

A true experiment like Bushman's study of catharsis uses a technique called randomization to reduce the risk of spuriousness. Students in Bushman's experiment were asked to select a message to read by drawing a random number out of a bag. That is, the students were assigned randomly to a treatment condition. If students were assigned to only two groups, a coin toss could have been used (see Exhibit 5.4). Random assignment ensures that neither students' aggressiveness nor any of their other characteristics or attitudes could influence which of the messages they read. As a result, the different groups are likely to be equivalent in all respects at the outset of the experiment. The greater the number of cases assigned randomly to the groups, the more likely that the groups will be equivalent in all respects. Whatever the preexisting sources of variation among the students, these could not explain why the group that read the pro-catharsis message became more aggressive, while the others didn't.

These defining features of true experimental designs give us a great deal of confidence that we can meet the three basic criteria for identifying nomothetic causes: association, time order, and nonspuriousness. However, we can strengthen our understanding of causal connections, and increase the likelihood of drawing causally valid conclusions, by also investigating causal mechanism and causal context.
A causal mechanism is some process that creates the connection between variation in an independent variable and the variation in the dependent variable it is hypothesized to cause (Cook & Campbell, 1979:35; Marini & Singer, 1988). Many social scientists (and scientists in other fields) argue that no nomothetic causal explanation is adequate until a causal mechanism is identified.

Up to this point, I have actually described only a portion of Bushman's experiment about catharsis. He and his colleagues actually conducted two experiments and measured some other variables that I haven't described. They tested the effect of reading a procatharsis message on the desire to hit a punching bag; they asked students how much they enjoyed hitting the punching bag; and they compared the aggressiveness of students who had hit the punching bag with those who hadn't. The findings that emerged from these additional study components provided a substantial bit of information about the causal mechanism that linked reading the procatharsis message to aggressive behavior. They concluded that the process went like this: Participants who read a procatharsis message were more likely to want to hit a punching bag; these students were then more likely to enjoy hitting the punching bag, and they were more likely to be aggressive in their actions toward others. When Bushman and his colleagues found that aggressiveness remained high throughout subsequent competitions, they also speculated that there was a "self-defeating prophecy" at work: When participants did not experience a reduction in their anger after expressing it, they became even more frustrated and angry.

Figuring out some aspects of the process by which the independent variable influenced the variation in the dependent variable should increase confidence in our conclusion that there was a causal effect (Costner, 1989). However, there may be many components to the causal mechanism and we cannot hope to identify them all in one study. For example, Bushman and his colleagues (1999:374) acknowledged that they had not identified empirically "the intrapsychic process, or mechanism that mediated effects of the persuasive messages." They did speculate that both anger and diminished self-esteem might be important, but an empirical test is left for another project.

In their study of deterrence of spouse abuse (introduced in Chapter 2), Lawrence Sherman and Richard Berk (1984) designed follow-up experiments to test or control for several causal mechanisms that they wondered about after their first experiment: Did recidivism decrease for those who were arrested for spouse abuse because of the exemplary work of the arresting officers? Did recidivism increase for arrestees as time passed and they experienced more stressors with their spouses? Investigating these and other possible causal mechanisms enriched Sherman and Berk's eventual explanation of how arrest influences recidivism.
Context

No cause has its effect apart from some larger context involving other variables. For whom and when and in what conditions does this effect occur? A cause is really one among a set of interrelated factors required for the effect (Hage & Meeker, 1988; Papineau, 1978). Identification of the context in which a causal effect occurs is not itself a criterion for a valid causal conclusion and it is not always attempted, but it does help us to understand the causal relationship.

Bushman and colleagues tested the effect of several contextual factors having to do with the types of persons reading the messages. They found that men and women had the same response to the cathartic message as did those who competed against the person who was the source of their angry feelings (the person who wrote the negative evaluation) and those who did not. However, being angry was a precondition for the effect of the procatharsis and anticatharsis articles.

Context was also important in Sherman and Berk’s research on domestic violence. Arrest was less effective in reducing subsequent domestic violence in cities with high levels of unemployment than in cities with low levels of unemployment. This seemed to be more evidence of the importance of individuals having a “stake in conformity” (Berk et al., 1992).

Nonexperimental Designs and the Criteria for Causal Explanation

The nonexperimental approach to establishing causality (sometimes called the descriptive or observational approach) involves studying naturally occurring variation in the dependent and independent variables, without any intervention by the researchers. Nonexperimental research designs can be either cross-sectional or longitudinal. In a cross-sectional research design, all data are collected at one point in time. Identifying the time order of effects can be an insurmountable problem with such a design. In longitudinal research designs, data are collected at two or more points in time, and so identification of the time order of effects can be quite straightforward.

Cross-Sectional Designs

Robert Sampson and Stephen Raudenbush (1999) used a very ambitious cross-sectional design to study the effect of visible public social and physical disorder on the crime rate in Chicago neighborhoods. Their theoretical framework focused on the concept of informal social control: the ability of residents to regulate social activity in their neighborhoods through their collective efforts according to desired principles. They believed that infor-
A Spurious Relationship

Spurious relationship: "Broken windows" theory

Social Disorder → Crime

The extraneous variable creates the spurious relationship:
Informal social control theory

Social Disorder

Collective Efficacy

Crime


Social control would vary between neighborhoods, and they hypothesized that it was the strength of informal social control that would explain variation in crime rates rather than just the visible signs of disorder. They contrasted this prediction to the "broken windows" theory: the belief that signs of disorder themselves cause crime. In the theory proposed by Sampson and Raudenbush, both visible disorder and crime were consequences of low levels of informal social control; one did not cause the other (Exhibit 5.5).

Sampson and Raudenbush measured visible disorder through direct observation: Trained observers rode slowly around every street in 196 Chicago census tracts. They also conducted a survey of residents and examined police records. Both survey responses and police records were used to measure crime levels. The level of neighborhood informal social control and other variables were measured with the average resident responses to several survey questions. Both the crime rate and the level of social and physical disorder varied between neighborhoods in relation to the level of informal social control. Informal social control (collective efficacy) was a much more important factor in the neighborhood crime rate than visible social and physical disorder (Exhibit 5.6).

How well does the study meet the criteria for establishing a causal connection? Sampson and Raudenbush showed that there was an association between variation in the independent variable, the level of informal social
control, and the dependent variable, the crime rate. However, their design could not establish directly that the variation in the crime rate occurred after variation in informal social control. Maybe it was a high crime rate that led residents to stop trying to exert much control over deviant activities in the neighborhood, perhaps because of fear of crime. It is difficult to discount such a possibility when only cross-sectional data are available.

A nonexperimental study like Sampson and Raudenbush's cannot use random assignment to comparison groups in order to minimize the risk of spurious effects. Even if we wanted to, we couldn't randomly assign people to live in neighborhoods with different levels of informal social control. Instead, nonexperimental researchers commonly use an alternative approach to try to achieve the criterion of nonspuriousness. The technique of statistical control allows researchers to determine whether the relationship between the independent and dependent variables still occurs while we hold constant the values of other variables. If it does, the relationship could not be caused by variation in that other variable.

Sampson and Raudenbush designed their study in part to determine whether the apparent effect of visible disorder on crime—the "broken windows" thesis—was spurious due to the effect of informal social control (see Exhibit 5.7). Exhibit 5.7 shows how statistical control was used to test this possibility. The data for all neighborhoods show that neighborhoods with much visible disorder had higher crime rates than those with less visible disorder. However, when we examine the relationship between visible disorder and neighborhood crime rate separately for neighborhoods with high and low levels of informal social control (when we "statistically control for social control level"), we see that the crime rate no longer varies with vis-

---

**Exhibit 5.6**  
Effect of Social Disorder and Collective Efficacy on Personal Violent Crimes

![Exhibit 5.6](image-url)
their design occurred at a rate that activities in crime to dis-
control. In the risk of assign people to test this hile we hold ip could not determine broken windows’ was spurious due to level of informal social control. Neigh-
levels of social control and physical disorder, and they were also more likely to have a high crime rate, but the visible disorder itself did not alter the crime rate.

**Statistical control** A technique used in nonexperimental research to reduce the risk of spuriousness. The effect of one or more variables are removed, for example, by holding them constant, so that the relationship between the independent and dependent variables can be assessed without the influence of variation in the control variables.

**Example:** In a different study, Sampson (1987) found a relationship between rates of family disruption and violent crime. He then classified cities by their level of joblessness (the control variable) and found that same relationship between the rates of family disruption and violent crime among cities with different levels of joblessness. So the rate of joblessness could not have caused the association between family disruption and violent crime.

Our confidence in causal conclusions based on nonexperimental research also increases with identification of a causal mechanism. Such mechanisms, which are termed **intervening variables** in nonexperimental
Exhibit 5.8  Intervening Variables in Nonexperimental Research: Structural Disadvantage and Juvenile Delinquency (Sampson & Laub, 1994)

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Intervening Variable (causal mechanism)</th>
<th>Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural disadvantage: Family poverty Geographic mobility</td>
<td>Informal social control: Low parent-child, attachment Low maternal supervision More erratic or harsh discipline</td>
<td>Juvenile delinquency</td>
</tr>
</tbody>
</table>

In a study that reanalyzed data from Sheldon Glueck and Elenor Glueck’s (1950) path-breaking study of juvenile delinquency, Robert Sampson and John Laub (1994) found that children who grew up with such structural disadvantages as family poverty and geographic mobility were more likely to become juvenile delinquents. Why did this occur? Their analysis indicated that these structural disadvantages led to lower levels of informal social control in the family (less parent-child attachment, less maternal supervision, and more erratic or harsh discipline). Lower levels of informal social control resulted in a higher probability of delinquency (Exhibit 5.8). Informal social control intervened in the relationship between structural context and delinquency.

Of course, identification of one (or two or three) intervening variables does not end the possibilities for clarifying the causal mechanisms. You might ask why structural disadvantage tends to result in lower levels of family social control or how family social control influences delinquency. You could then conduct research to identify the mechanisms that link, for example, family social control and juvenile delinquency. (Perhaps the children feel they’re not cared for, so they become less concerned with conforming to social expectations.) This process could go on and on. The point is that identification of a mechanism through which the independent variable influences the dependent variable increases our confidence in the conclusion that a causal connection does indeed exist.

When you think about the role of variables in causal relationships, don’t confuse variables that cause spurious relationships with those that intervene in causal relationships—even though both are “third variables” that do not appear in the initial hypothesis. In Exhibit 5.9 the extraneous variable, joblessness, creates a spurious relationship. By contrast, in Exhibit 5.8, the intervening variable is part of the process that links the independent
Hypothesized causal relationship

Spurious causal relationship: Independent and dependent variables are merely correlated.

Variable and the dependent variable; intervening variables help to explain the relationship between the independent variable (structural disadvantage) and the dependent variable (juvenile delinquency) (Davis, 1985).

Specifying the context in which causal effects occur is no less important in nonexperimental than in experimental research. Nonexperimental research is, in fact, well suited to exploring the context in which causal effects occur. Administering surveys in many different settings and to different types of individuals is usually much easier than administering experiments in different ways.

Longitudinal Designs

It is risky to draw conclusions about time order on the basis of cross-sectional data, except in four special cases (see below). In longitudinal research, in contrast, data are collected that can be ordered in time. By measuring the value of cases on an independent variable and a dependent variable at each of these different times, the researcher can determine whether variation in the independent variable precedes variation in the dependent variable.

The four special circumstances in which cross-sectional data can reasonably be used to infer the time order of effects can actually be thought of as longitudinal designs, in the sense that the data can be ordered in time (Campbell, 1992):

- The independent variable is fixed at some point prior to the variation in the dependent variable. So-called demographic variables that are determined at birth—such as sex, race, and age—are fixed in this way. So are variables
like education and marital status, if we know when the value of cases on these variables was established and if we know that the value of cases on the dependent variable was set some time later. For example, say we hypothesize that education influences the type of job individuals have. If we know that respondents completed their education before taking their current jobs, we would satisfy the time order requirement even if we were to measure education at the same time we measure type of job. However, if some respondents possibly went back to school as a benefit of their current job, the time order requirement would not be satisfied.

We believe that respondents can give us reliable reports of what happened to them or what they thought at some earlier point in time. Julie Horney, D. Wayne Osgood, and Ineke Haen Marshall (1995) provide an interesting example of the use of such retrospective data. The researchers wanted to identify how criminal activity varies in response to changes in life circumstances. They interviewed 658 newly convicted male offenders sentenced to a Nebraska state prison. In a 45- to 90-minute interview, they recorded each inmate’s report of his life circumstances and of his criminal activities for the preceding two to three years. They then found that criminal involvement was related strongly to adverse changes in life circumstances, such as marital separation or drug use. Retrospective data are often inadequate for measuring variation in past psychological states or behaviors, however, because what we recall about our feelings or actions in the past is likely to be influenced by what we feel in the present. For example, retrospective reports by both adult alcoholics and their parents appear to greatly overestimate the frequency of childhood problems (Vaillant, 1995). People cannot report reliably the frequency and timing of many past events, from hospitalization to hours worked. However, retrospective data tends to be reliable when it concerns major, persistent experiences in the past, such as what type of school someone went to or how a person’s family was structured (Campbell, 1992).

Our measures are based on records that contain information on cases in earlier periods. Government, agency, and organizational records are an excellent source of time-ordered data after the fact. However, sloppy record keeping and changes in data-collection policies can lead to inconsistencies, which must be taken into account. Another weakness of such archival data is that they usually contain measures of only a fraction of the variables that we think are important.

We know that cases were equivalent on the dependent variable prior to the treatment. For example, we may hypothesize that a training program (independent variable) improves the English-speaking abilities (dependent variable) of a group of recent immigrants. If we know that none of the immigrants could speak English prior to enrolling in the training
In cases on value of cases sample, say we individuals have. before taken even if type of job, as a benefit satisfied. It happened to Julie Horney, an interesting researcher to changes male 0-minute circumstances years. They y to advise or drug use. at we recall influenced by both estimate the cannot report hospitals to be reliable, such as family was uses in earlier an excellent record inconsistencies, such archival of the variability prior to the program (depend that none of the training

program, we can be confident that any subsequent variation in their ability to speak English did not precede exposure to the training program. This is one way that traditional experiments establish time order: Two or more equivalent groups are formed prior to exposing one of them to some treatment.

When these special circumstances do not exist in nonexperimental research, we must actually collect data at two or more points in time in order to establish empirically the time order of effects (Campbell, 1992). In some longitudinal designs, the same sample (or panel) is followed over time; in other designs, sample members are rotated or completely replaced. The population from which the sample is selected may be defined broadly, as when a longitudinal survey of the general population is conducted. Or the population may be defined narrowly, as when members of a specific age group are sampled at multiple points in time. The frequency of follow-up measurement can vary, ranging from a before-and-after design with just one follow-up to studies in which various indicators are measured every month for many years.

Certainly it is more difficult to collect data at two or more points in time than at one time. Quite frequently researchers simply cannot, or are unwilling to, delay completion of a study for even one year in order to collect follow-up data. But think of the many research questions that really should involve a much longer follow-up period: What is the impact of job training on subsequent employment? How effective is a school-based program in improving parenting skills? Under what conditions do traumatic experiences in childhood result in mental illness? It is safe to say that we will never have enough longitudinal data to answer many important research questions. The value of longitudinal data is so great that every effort should be made to develop longitudinal research designs when they are appropriate for the research question asked. The following discussion of the three major types of longitudinal design will give you a sense of the possibilities (see Exhibit 5.10).

**Repeated cross-sectional design** A type of longitudinal study in which data are collected at two or more points in time from different samples of the same population.

**Fixed-sample panel design** A type of longitudinal study in which data are collected from the same individuals—the panel—at two or more points in time. In another type of panel design, panel members who leave are replaced with new members.

**Event-based design** A type of longitudinal study in which data are collected at two or more points in time from individuals in a cohort.
Exhibit 5.10  Three Types of Longitudinal Design

Repeated Cross-Sectional Design

Fixed-Sample Panel Design

Event-Based Design

Cohort  Individuals or groups with a common starting point. Examples include college class of 1997, people who graduated from high school in the 1980s, General Motors employees who started work between 1990 and the year 2000, and people who were born in the late 1940s or the 1950s (the "baby boom generation").

Repeated Cross-Sectional Designs
Repeated cross-sectional studies, also known as trend studies, have become fixtures of the political arena around election time. Particularly in presidential election years, we have all become accustomed to reading weekly, even daily, reports on the percentage of the population that supports each candidate. Similar polls are conducted to track sentiment on many other social issues. For example, a 1993 poll reported that 52% of adult Americans supported a ban on the possession of handguns, compared to 41% in a similar poll conducted in 1991. According to pollster Louis Harris, this increase indicated a "sea change" in public attitudes (Barringer, 1993). Another researcher said, "It shows that people are responding to their experience [of an increase in handgun-related killings]" (Barringer, 1993:1).
Repeated cross-sectional surveys are conducted as follows:

1. A sample is drawn from a population at time 1, and data are collected from the sample.
2. As time passes, some people leave the population and others enter it.
3. At time 2 a different sample is drawn from this population.

These features make the repeated cross-sectional design appropriate when the goal is to determine whether a population has changed over time. Has racial tolerance increased among Americans in the past 20 years? Are employers more likely to pay maternity benefits today than they were in the 1950s? These questions concern changes in the population as a whole, not changes in individuals within the population. We want to know whether racial tolerance increased in society, not whether this change was due to migration that brought more racially tolerant people into the country or to individual U.S. citizens becoming more tolerant. We are asking whether employers overall are more likely to pay maternity benefits today than they were yesterday, not whether any such increase was due to recalcitrant employers going out of business or to individual employers changing their maternity benefits. When we do need to know whether individuals in the population changed, we must turn to a panel design.

**Fixed-Sample Panel Designs**

Panel designs allow identification of changes in individuals, groups, or whatever we are studying. This is the process for conducting fixed-sample panel studies:

1. A sample (called a panel) is drawn from a population at time 1, and data are collected from the sample.
2. As time passes, some panel members become unavailable for follow-up, and the population changes.
3. At time 2, data are collected from the same people as at time 1 (the panel)—except for those people who cannot be located.

Because a panel design follows the same individuals, it is better than a repeated cross-sectional design for testing causal hypotheses. For example, Sampson and Laub (1990) used a fixed-sample panel design to investigate the effect of childhood deviance on adult crime. They studied a sample of white males in Boston when the subjects were between 10 and 17 years old and then followed up when the subjects were in their adult years. Data were collected from multiple sources, including the subjects themselves and criminal justice records. Sampson and Laub (1990:614) found that children who had been committed to a correctional school for persistent delinquency were much more likely to commit crimes as adults: 61% were arrested between the ages of 25 and 32, compared to 14% of those who had not been in
Exhibit 5.11  Causality in Panel Studies

Although delinquency in the 11th grade and grades in the 12th grade are clearly associated and the time order is clear, causality cannot be determined. In reality, grades in the 7th grade also play a role.

correctional schools as juveniles. In this study, juvenile delinquency unquestionably occurred before adult criminality. If the researchers had used a cross-sectional design to study the past of adults, the juvenile delinquency measure might have been biased by memory lapses, by self-servings recollections about behavior as juveniles, or by loss of agency records.

If you now wonder why every longitudinal study isn’t designed as a panel study, you’ve understood the advantages of panel designs. However, remember that this design does not in itself establish causality. Variation in both the independent variable and the dependent variables may be due to some other variable, even to earlier variation in what is considered the dependent variable. In the example in Exhibit 5.11 there is a hypothesized association between delinquency in the 11th grade and grades obtained in the 12th grade (the dependent variable). The time order is clear. However, both variables are consequences of grades obtained in the 7th grade. The apparent effect of 11th-grade delinquency on 12th-grade grades is spurious because of variation in the “dependent” variable (grades) at an earlier time.

Panel designs are also a challenge to implement successfully, and often are not even attempted, because of two major difficulties:

- **Expense and attrition.** It can be difficult, and very expensive, to keep track of individuals over a long period, and inevitably the proportion of panel members who can be located for follow-up will decline over time. Panel studies often lose more than one-quarter of their members through attrition (Miller, 1991:170). However, subject attrition can be reduced substantially if sufficient staff can be used to keep track of panel members. In their panel study, Sampson and Laub (1990) lost only 12% of the juveniles in the original sample (8% if you do not count those who had died).
The consequences of a high rate of subject attrition are that the follow-up sample may no longer be representative of the population from which it was drawn and may no longer provide a sound basis for estimating change. Subjects who were lost to follow-up may have been those who changed the most, or the least, over time. It does help to compare the baseline characteristics of those who are interviewed at follow-up with characteristics of those lost to follow-up. If these two groups of panel members were not very different at baseline, it is less likely that changes had anything to do with characteristics of the missing panel members.

- Subject fatigue. Panel members may grow weary of repeated interviews and drop out of the study, or they may become so used to answering the standard questions in the survey that they start giving stock answers rather than actually thinking about their current feelings or actions. (Campbell, 1992). This is called the problem of subject fatigue. Fortunately, subjects do not often seem to become fatigued in this way, particularly if the research staff have maintained positive relations with the subjects. For example, at the end of an 18-month-long experimental study of housing alternatives for persons with mental illness who had been homeless, only 3 or 4 individuals (out of 93 who could still be located) refused to participate in the fourth and final round of interviews. The interviews took a total of about five hours to complete, although participants did receive about $50 (Schutt, Goldfinger, & Penk, 1997).

Because panel studies are so useful, social researchers have developed increasingly effective techniques for keeping track of individuals and overcoming subject fatigue. But when resources do not permit use of these techniques to maintain an adequate panel, repeated cross-sectional designs usually can be employed at a cost that is not a great deal higher than that of a one-time-only cross-sectional study. The payoff in explanatory power should be well worth the cost.

**Event-Based Designs**

In an event-based design, often called a cohort study, the follow-up samples (at one or more times) are selected from the same cohort—people who all have experienced a similar event or a common starting point. Examples include:

- Birth cohorts—those who share a common period of birth (those born in the 1940s, 1950s, 1960s, and so on)
- Seniority cohorts—those who have worked at the same place for about 5 years, about 10 years, and so on
- School cohorts—freshmen, sophomores, juniors, seniors
An event-based design can be a type of repeated cross-sectional design or a type of panel design. In an event-based repeated cross-sectional design, separate samples are drawn from the same cohort at two or more different times. In an event-based panel design, the same individuals from the same cohort are studied at two or more different times.

We can see the value of event-based research in a comparison of two studies that estimated the impact of public and private schooling on high school students’ achievement test scores. In a cross-sectional study, James Coleman, Thomas Hoffer, and Sally Kilgore (1982) compared standardized achievement test scores of high school sophomores and seniors in public, Catholic, and other private schools. They found that test scores were higher in the private high schools (both Catholic and other) than in the public high schools. But was this difference a causal effect of private schooling? Perhaps the parents of higher-performing children were choosing to send them to private rather than to public schools. In other words, the higher achievement levels of private-sector students might have been in place before they started high school and not have developed as a consequence of their high school education.

The researchers tried to reduce the impact of this problem by statistically controlling for a range of family background variables: family income, parents’ education, race, number of siblings, number of rooms in the home, number of parents present, mother working, and other indicators of a family orientation to education. But some critics pointed out that even with all these controls for family background, the cross-sectional study did not ensure that the students had been comparable in achievement when they started high school.

So James Coleman and Thomas Hoffer (1987) went back to the high schools and studied the test scores of the former sophomores two years later, when they were seniors; in other words, the researchers used an event-based panel design. This time they found that the verbal and math achievement test scores of the Catholic school students had increased more over the two years than was the case for the public school students; it was not clear whether the scores of the other private school students had increased. Irrespective of students’ initial achievement test scores, the Catholic schools seemed to “do more” for their students than did the public schools. This finding continued to be true even when dropouts were studied, too. The researchers’ causal conclusion rested on much stronger ground because they used an event-based panel design.

Units of Analysis and Errors in Causal Reasoning

Regardless of the research design, we can easily come to invalid conclusions about causal influences if we do not know what units of analysis the
measures in our study refer to—that is, the level of social life on which the research question is focused, such as individuals, groups, towns, or nations.

Individual and Group Units of Analysis

In most sociological and psychological studies, the units of analysis are individuals. The researcher may collect survey data from individuals, analyze the data, and then report on, say, how many individuals felt socially isolated and whether substance abuse by individuals was related to their feelings of social isolation.

The units of analysis may instead be groups of some sort, such as families, schools, work organizations, towns, states, or countries. For example, a researcher may collect data from town and police records on the number of accidents in which a driver was intoxicated and the presence or absence of a server liability law in the town. (These laws make those who serve liquor liable for accidents caused by those to whom they served liquor.) The researcher can then analyze the relationship between server liability laws and the frequency of accidents due to drunk driving (perhaps also taking into account town population). Because the data describe the town, towns are the units of analysis.

In some studies, groups are the units of analysis but data are collected from individuals. For example, in an earlier article from their study of influences on violent crime in Chicago neighborhoods, Robert Sampson, Stephen Raudenbush, and Felton Earls (1997) hypothesized that “collective efficacy” would influence neighborhood crime rates. Collective efficacy was defined conceptually as a characteristic of the neighborhood: the extent to which residents were likely to help other residents and were trusted by other residents. However, they measured this variable in a survey of individuals. The responses of individual residents about their perceptions of their neighbors’ helpfulness and trustworthiness were averaged together to create a collective efficacy score for each neighborhood. It was this neighborhood measure of collective efficacy that was used to explain variation in the rate of violent crime between neighborhoods. The data were collected from individuals and were about individuals, but they were combined (aggregated) so as to describe neighborhoods. The units of analysis were thus groups (neighborhoods).

In a study like Sampson’s, we can distinguish the concept of units of analysis from the units of observation. Data were collected from individuals, the units of observation, and then the data were aggregated and analyzed at the group level. In some studies, the units of observation and the units of analysis are the same. The important point is to know. A conclusion that “crime increases with joblessness” could imply either that individuals who lose their jobs are more likely to commit a crime or that a community with a high unemployment rate is likely to have a high crime
rate—or both. Whether we are drawing conclusions from data or interpreting others’ conclusions, it is important to be clear about which relationship is being referred to.

We also have to know the units of analysis to interpret statistics appropriately. Measures of association tend to be stronger for group-level than for individual-level data because measurement errors at the individual level tend to cancel out at the group level (Bridges & Weis, 1989:29–31).

The Ecological Fallacy and Reductionism

Researchers should make sure that their causal conclusions reflect the units of analysis in their study. Conclusions about processes at the individual level should be based on individual-level data; conclusions about group-level processes should be based on data collected about groups. In most cases, violation of this rule creates one more reason to suspect the validity of the causal conclusions.

A researcher who draws conclusions about individual-level processes from group-level data is making what is termed an ecological fallacy (see Exhibit 5.12). The conclusions may or may not be correct, but we must recognize that group-level data do not describe individual-level processes.

Exhibit 5.12  Errors in Causal Conclusions

<table>
<thead>
<tr>
<th>You collect data from</th>
<th>Groups</th>
<th>Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>OK!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Groups</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STOP! Reductionist fallacy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STOP! Ecological fallacy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OK!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For example, a researcher may examine factory records and find that the higher the percentage of unskilled workers in factories, the higher the rate of employee sabotage in those factories. But the researcher would commit an ecological fallacy if she then concluded that individual unskilled factory workers are more likely to engage in sabotage. This conclusion is about an individual-level causal process (the relationship between the occupation and criminal propensities of individuals), even though the data describe groups (factories). It could actually be that white-collar workers are the ones more likely to commit sabotage, perhaps because in factories with more unskilled workers the white-collar workers feel they won’t be suspected.

Bear in mind that conclusions about individual processes based on group-level data are not necessarily wrong. We just don’t know for sure. Say that we find communities with higher average incomes have lower crime rates. The only thing special about these communities may be that they have more individuals with higher incomes, who tend to commit fewer crimes. Even though we collected data at the group level and analyzed them at the group level, they reflect a causal process at the individual level (Sampson & Lauritsen, 1994:80–83).

When data about individuals are used to make inferences about group-level processes, a problem occurs that can be thought of as the mirror image of the ecological fallacy: the reductionist fallacy, or reductionism (see Exhibit 5.12). For example, William Wilson (1987:58) notes that we can be misled into concluding from individual-level data that race has a causal effect on violence because there is an association at the individual level between race and the likelihood of arrest for violent crime. However, community-level data reveal that almost 40% of poor blacks lived in extremely poor areas in 1980, compared to only 7% of poor whites. The concentration of African-Americans in poverty areas, not the race or other characteristics of the individuals in these areas, may be the cause of higher rates of violence. Explaining violence in this case requires community-level data.

The fact that errors in causal reasoning can be made should not deter you from conducting research with aggregate data nor make you unduly critical of researchers who make inferences about individuals on the basis of aggregate data. When considered broadly, many research questions point to relationships that could be manifested in many ways and on many levels. Sampson’s (1987) study of urban violence is a case in point. His analysis involved only aggregate data about cities, and he explained his research approach as in part a response to the failure of other researchers to examine this problem at the structural, aggregate level. Moreover, Sampson argued that the rates of joblessness and family disruption in communities influence community social processes, not just the behavior of the specific individuals who are unemployed or who grew up without two parents. Yet Sampson
suggestions that the experience of joblessness and poverty is what tends to reduce the propensity of individual men to marry and that the experience of growing up in a home without two parents in turn increases the propensity of individual juveniles to commit crimes. These conclusions about the behavior of individuals seem consistent with the patterns Sampson found in his aggregate, city-level data, so it seems unlikely that he is committing an ecological fallacy when he proposes them.

The solution is to know what the units of analysis and units of observation were in a study and to take these into account in weighing the credibility of the researcher's conclusions. The goal is not to reject out of hand conclusions that refer to a level of analysis different from what was actually studied. Instead, the goal is to consider the likelihood that an ecological fallacy or a reductionist fallacy has been made when estimating the causal validity of the conclusions.

**Causation in Qualitative Research**

When qualitative researchers attempt to develop causal explanations, they are likely to use an "idiographic" conception of cause rather than the "nomothetic" conception used by quantitative researchers. An idiographic causal explanation is one that identifies the concrete, individual sequence of events, thoughts, or actions that resulted in a particular outcome for a particular individual or that led to a particular event (Hage & Meeker, 1988). An idiographic explanation also may be termed a narrative, individualist, historicist, or case-oriented explanation.

A causal explanation that is idiographic includes statements of initial conditions and then relates a series of events at different times that led to the outcome, or causal effect. This narrative, or story, is the critical element in an idiographic explanation, which may therefore be classified as narrative reasoning (Richardson, 1995:200–201). Idiographic explanations focus on particular social actors, in particular social places, at particular social times (Abbott, 1992). Idiographic explanations are also "holistic": they typically are very concerned with context, with understanding the particular outcome as part of a larger set of interrelated circumstances.

Elijah Anderson's (1990) field research in a poor urban community produced a narrative account of how drug addiction often resulted in a downward slide into residential instability and crime:

When addicts deplete their resources, they may go to those closest to them, drawing them into their schemes... [T]he family may put up with the person for a while. They provide money if they can. ... They come to realize that the person is on drugs. ... Slowly the reality sets in more and more completely, and the family becomes drained of both
financial and emotional resources. Close relatives lose faith and begin to see the person as untrustworthy and weak. Eventually the addict begins to “mess up” in a variety of ways, taking furniture from the house [and] anything of value. Relatives and friends begin to see the person . . . as “out there” in the streets. . . . One deviant act leads to another. (Anderson, 1990:86–87)

**Causal effect (idiographic perspective)** The finding that a series of events following an initial set of conditions leads in a progressive manner to a particular event or outcome.

*Example of an idiographic causal explanation:* An individual is neglected by her parents but has a supportive grandparent. She comes to distrust others, has trouble in school, is unable to keep a job, and eventually becomes homeless. She subsequently develops a supportive relationship with a shelter case manager, who helps her find a job and regain her housing (based on K. Hirsch, 1989).

An idiographic explanation like Anderson’s pays close attention to time order and causal mechanisms. Nonetheless, it is difficult to make a convincing case that one particular causal narrative should be chosen over an alternative narrative (Abbott, 1992). Does low self-esteem result in vulnerability to the appeals of drug dealers, or does a chance drug encounter precipitate a slide in self-esteem? The prudent causal analyst remains open to alternative explanations.

**Conclusion**

Causation and the means for achieving causally valid conclusions in research is the last of the three legs on which the validity of research rests. In this chapter, you have learned about the two main meanings of causation (nomothetic and idiographic) and about the five criteria used to evaluate the extent to which particular research designs may achieve causally valid findings. You have been exposed to the problem of spuriousness and the ways that randomization and statistical control deal with it. You also have learned how to establish the time order of effects in nonexperimental research and how to come to causal conclusions that are appropriate to the research design.

I should reemphasize that the results of any particular study are part of an always changing body of empirical knowledge about social reality. Thus our understandings of causal relationships are always partial. Researchers always wonder whether they have omitted some relevant variables from
their controls or whether their experimental results would differ if the experiment were conducted in another setting or whether they have overlooked a critical historical event. But by using consistent definitions of terms and maintaining clear standards for establishing the validity of research results—and by expecting the same of others who do research—social researchers can contribute to a growing body of knowledge that can reliably guide social policy and social understanding.

When you read the results of a social scientific study, you should now be able to evaluate critically the validity of the study’s findings. If you plan to engage in social research, you should now be able to plan an approach that will lead to valid findings. And with a good understanding of the three dimensions of validity (measurement validity, generalizability, and causal validity) under your belt, you are ready to focus on the four major methods of data collection used by social scientists. Each of these methods tends to use a somewhat different approach to achieving validity.

**KEY TERMS**

Association  
Causal effect (idiographic perspective)  
Causal effect (nomothetic perspective)  
Ceteris paribus  
Cohort  
Cohort study  
Context  
Counterfactual  
Cross-sectional research design  
Ecological fallacy  
Event-based design  
Experimental approach  
Extraneous variable  
Fixed-sample panel design  
Idiographic causal explanation  
Intervening variable  
Longitudinal research design  
Mechanism  
Nomothetic causal explanation  
Nonexperimental approach  
Nonspuriousness  
Random assignment  
Randomization  
Reductionist fallacy (reductionism)  
Repeated cross-sectional design  
Spurious relationship  
Statistical control  
Subject fatigue  
Time order  
Trend study  
Units of analysis  
Unit of observation

**HIGHLIGHTS**

- Causation can be defined in either nomothetic or idiographic terms. Nomothetic causal explanations deal with effects on average. Idiographic causal explanations deal with the sequence of events that led to a particular outcome.
- The concept of nomothetic causal explanation relies on a comparison. The value of cases on the dependent variable is measured after they have been exposed to variation in an independent variable. This mea-
measurement is compared to what the value of cases on the dependent variable would have been if they had not been exposed to the variation in the independent variable (the counterfactual). The validity of nomothetic causal conclusions rests on how closely the comparison group comes to the ideal counterfactual.

- From a nomothetic perspective three criteria are generally viewed as necessary for identifying a causal relationship: association between the variables, proper time order, and nonspuriousness of the association. In addition, the basis for concluding that a causal relationship exists is strengthened by identification of a causal mechanism and the context.

- Association between two variables is in itself insufficient evidence of a causal relationship. This point is commonly made with the expression “Correlation does not prove causation.”

- Experiments use random assignment to make comparison groups as similar as possible at the outset of an experiment in order to reduce the risk of spurious effects due to extraneous variables.

- Nonexperimental designs use statistical controls to reduce the risk of spuriousness. A variable is controlled when it is held constant so that the association between the independent and dependent variables can be assessed without being influenced by the control variable.

- Ethical and practical constraints often preclude the use of experimental designs.

- Idiographic causal explanations can be difficult to identify, because the starting and ending points of particular events and the determination of which events act as causes in particular sequences may be ambiguous.

- Longitudinal designs are usually preferable to cross-sectional designs for establishing the time order of effects. Longitudinal designs vary in terms of whether the same people are measured at different times, how the population of interest is defined, and how frequently follow-up measurements are taken. Fixed-sample panel designs provide the strongest test for the time order of effects, but they can be difficult to carry out successfully, because of their expense and subject attrition and fatigue.

- We do not fully understand the variables in a study until we know what units of analysis—what level of social life—they refer to.

- Invalid conclusions about causality may occur when relationships between variables measured at the group level are assumed to apply at the individual level (the ecological fallacy) and when relationships between variables measured at the level of individuals are assumed to apply at the group level (the reductionist fallacy). Nonetheless, many research questions point to relationships at multiple levels and may profitably be answered by studying different units of analysis.
1. Review articles in several newspapers, copying down all causal assertions. These might range from assertions that the stock market declined because of uncertainty in the Middle East to explanations about why a murder was committed or why test scores are declining in U.S. schools. Inspect the articles carefully, noting all evidence used to support the causal assertions. Are the explanations nomothetic, idiographic, or a combination of both? Which criteria for establishing causality in a nomothetic framework are met? How satisfactory are the idiographic explanations? What other potentially important influences on the reported outcome have been overlooked?

2. Select several research articles in professional journals that assert, or imply, that they have identified a causal relationship between two or more variables. Are each of the criteria for establishing the existence of a causal relationship met? Find a study in which subjects were assigned randomly to experimental and comparison groups to reduce the risk of spurious influences on the supposedly causal relationship. How convinced are you by the study?

Find a survey study that makes causal assertions based on the relationships, or correlations, among variables. What variables have been statistically controlled? List other variables that might be influencing the relationship but that have not been controlled. How convinced are you by the study?

3. Search Sociological Abstracts or another index to the social science literature for several articles on studies using any type of longitudinal design. You will be searching for article titles that use words like longitudinal, panel, trend, or over time. How successful were the researchers in carrying out the design? What steps did the researchers who used a panel design take to minimize panel attrition? How convinced are you by those using repeated cross-sectional designs that they have identified a process of change in individuals? Did any researchers use retrospective questions? How did they defend the validity of these measures?

4. The practice diskette contains lessons on units of analysis and the related problems of ecological fallacy and reductionism. Choose the Units of Analysis lesson from the main menu. It describes several research projects and asks you to identify the units of analysis in each. Then it presents several conclusions for particular studies and asks you to determine whether an error has been made.

5. Propose a hypothesis involving variables that could be measured with individuals as the units of analysis. How might this hypothesis be restated so as to involve groups as the units of analysis? Would you expect the hypothesis to be supported at both levels? Why or why not? Repeat the exercise, this time starting with a different hypothesis involving groups as the units of analysis and then restating it so as to involve individuals as the units of analysis.

DEVELOPING A RESEARCH PROPOSAL

How will you try to establish the causal effects you hypothesize?

1. Identify at least one hypothesis involving what you expect is a causal relationship.
2. Identify key variables that should be controlled in your survey design in order to increase your ability to avoid arriving at a spurious conclusion about the hypothesized causal effect. Draw on relevant research literature and social theory to identify these variables.

3. Add a longitudinal component to your research design. Explain why you decided to use this particular longitudinal design.

4. Review the criteria for establishing a nomothetic causal effect and discuss your ability to satisfy each one. Include in your discussion some consideration of how well your design will avoid each of the threats to experimental validity.

WEB EXERCISES

   From the links supplied, find information regarding a subject of your choosing related to crime and/or violence (for example, youth violence, corporate crime, rape, etc.). Report on the prevalence and/or extent of the phenomenon you have identified. Propose a causal explanation for variation in this phenomenon. What research design would you propose to test this explanation? Explain.

   Check out “What is CSI” and “President’s Message.” How is CSI “fighting crime”? What does CSI’s approach assume about the cause of crime? Do you think CSI’s approach to fighting crime is based on valid conclusions about causality? Explain.

3. What are the latest trends in crime? Write a short statement after inspecting the FBI’s Uniform Crime Reports at www.fbi.gov
   You will need to use the Adobe Acrobat Reader to access these reports, most of which are in “pdf” format. Follow the instructions on the site if you’re not familiar with this.

SPSS EXERCISES

We can use the GSS98 data to learn how causal hypotheses can be evaluated with nonexperimental data.

1. Specify four hypotheses in which CAPPUN is the dependent variable and the independent variable is also measured with a question in the GSS98. The independent variables should have no more than 10 valid values (check the variable list).

2. Inspect the frequency distributions of each independent variable in your hypotheses. If it appears that any have little valid data or were coded with more than 10 categories, substitute another independent variable.
3. Generate crosstabs that show the association between CAPPUN and each of the independent variables. Make sure that CAPPUN is the row variable and that you select "Column Percents."

4. Does support for capital punishment vary across the categories of any of the independent variables? By how much? Would you conclude that there is an association, as hypothesized, for any pairs of variables?

5. Might one of the associations you have just identified be spurious due to the effect of a third variable? What might such an extraneous variable be? Look through the variable list and find a variable that might play this role. If you can’t think of any possible extraneous variables, or if you didn’t find an association in support of any of your hypotheses, try this: Examine the association between CAPPUN and WRKSTAT2. In the next step, control for sex (gender). The idea is that there is an association between work status and support for capital punishment, that might be spurious due to the effect of sex (gender). Proceed with the following steps:
   a. Select Analyze → Descriptive statistics → Crosstabs.
   b. In the Crosstabs window, click CAPPUN into Rows, WRKSTAT2 into Columns, and SEX into Layer 1 of 1.
   c. Select Cells → Percentages Column → Continue → OK.
   Is the association between employment status and support for capital punishment affected by gender? Do you conclude that the association between CAPPUN and WRKSTAT2 seems to be spurious due to the effect of SEX?

6. Does the association between support for capital punishment and any of your independent variables vary with social context? Marian Borg (1997) concluded that it did. Test this out for the association between attitude toward African Americans (RACPUSH2) and CAPPUN. Follow the procedures in Exercise 5, but click RACPUSH2 into columns and REGION4 into Layer 1 of 1. (You must first click the variables used previously back into the variables list.) Take a while to study this complex three-variable table. Does the association between CAPPUN and RACPUSH2 vary with region? How would you interpret this finding?

7. Now, how about the influence of an astrological sign on support for capital punishment? Create a crosstabulation in which ZODIAC is the independent (column) variable and CAPPUN is the dependent (row) variable (with column percents). What do you make of the results?