Chapter 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

Idiographic Causal Explanation

Causation in Qualitative Research
Explanation in Qualitative Research
Single-Subject Design

Conclusion

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 poverty rate for children under age six stay relatively stable in the 1970s, increase in the 1980s and early 1990s, and then begin to decline in 1993? Changes in public policy? Variations in the economy? Changing demographic patterns? To distinguish these possibilities, we must design our research strategies carefully.

This chapter considers the meaning of nomothetic causation, the criteria for achieving causally valid explanations, and the ways in which different research designs seek to meet these criteria. We will focus attention on the differences between the experimental and nonexperimental approaches to causation. We will also contrast the nomothetic approach with idiographic causal explanation. We 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.

**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 communities) or for events. In social work practice, interventions are thought to explain changes in client status; frequently, social work researchers test this assumption. Other social work researchers seek to explain the causes of social conditions. For example, some researchers seek to identify the causes of poverty so that they can then recommend policy changes that would deal with the identified causes. One explanation, or cause, may be that poverty is the result of differences in human capital, that is, differences in the skills and education that people bring to the marketplace. If that is the case, then there are programs to increase human capital. You should recognize that there is a hypothesis here: Adults with less education are more likely to be poor than adults with more education. (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). We 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. We 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).

**EXPERIMENTAL DESIGN AND THE CRITERIA FOR CAUSAL EXPLANATION**

Five criteria must be considered when deciding whether a causal connection exists. Research designs that allow us to establish these criteria require 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.

Elaine Walton, Mark W. Fraser, Robert E. Lewis, Peter J. Pecora, and Wendel K. Walton's 1993 child welfare study of the impact of intensive, in-home, family-based services on rates of family reunification illustrates how an experimental approach can be used to meet the criteria for establishing causal relations. Walton and her colleagues recruited 110 families using a computer-generated list of children in substitute care. The sample included only children who had been outside the home for more than 30 days, whose return home was not imminent but still the case goal, and whose parents were not dead or incarcerated. Fifty-seven families were in the experimental group while there were 53 families in the second group. In the original study (Walton et al., 1993), information was collected at the end of the 90 days, a 6-month follow-up, and a 12-month follow-up.

Families in the experimental group received intensive services that were home-based. The intervention model was based on the idea that creating a relationship between the caseworker and the family, providing concrete services, focusing on the family and not just the child, establishing or building a supportive network, and building skills would produce better client outcomes. Families in the experimental group received these services for 90 days and were seen in the home at least three times a week. The comparison group received the traditional package of routine services that were offered; these services were provided out-of-home, caseworkers were required to see each child only once per month, and the families were helped to locate services necessary to return the child to the home, such as mental health counseling and parenting skills training.

Did the intervention make any difference? As a short-term intervention, the program was quite successful in that 53 children (93%) in the experimental group were reunited with their families during the first 90 days of intervention whereas only 15 (28%) of the control group were returned to their homes. After 15 months, of the 57 children in the experimental group, 44 (77%) had returned home without any subsequent substitute care, 5 (9%) had been briefly placed out of the home again, 12 (21%) who had returned home were back in substitute care, and one never returned home. In comparison, among those children who received traditional services, only 25 (47%) returned home without any interruption, 5 (9%) who had returned home were back in substitute care, and 23 (43%) never returned home. Although Walton and her colleagues (1993) point out various concerns, they conclude that the “results are encouraging. Intensive in-home services that have a skills focus appear to be effective in the reunification of some families with children in out-of-home care” (p. 485).

Was this causal conclusion justified? How confident can we be about its internal validity? We will answer this question by reviewing how the experiment attempted to meet each of the causal criteria and what are the key features of a true experiment. We will study variations on experimental designs in Chapter 6.

Association

We say that there was an association between the intervention and family reunification because the rate of family reunification differed according to the type of intervention a family received. 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 intensive, in-home, family-based services, 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. Or, the other group may receive what is routinely provided so it is getting some form of treatment; this second group is termed the comparison group. The Walton study, as we have described it, compared two groups; other experiments may compare more groups, which represent multiple values of the independent variable or even combinations of the values of two or more independent variables. An empirical association was demonstrated by comparing the two groups using statistical techniques to show that the differences were not due to chance.

**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. The subjects in Walton’s study all started with the same status: All children had been placed out of the home with no immediate intent to reunite with their families. Walton and her colleagues (1993) then measured changes in that status, both during the 90 days of the treatment period and after the 90 days were finished. It was only during and after the interventions that children returned home.

**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. Correlation does not prove causation as 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 a **spurious relationship**. In fact, both the number of storks and the birth rate are higher in rural districts than in urban districts. 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 Walton’s study of intensive in-home, family-based services uses a technique called **randomization** to reduce the risk of spuriousness. If subjects were assigned to only two groups, a coin toss could have been used (see Exhibit 5.1). The families in Walton’s experiment were randomly assigned to either the experimental group or the traditional treatment group. This means that chance determined into which group a family would be placed. **Random assignment** ensures that neither the families’ other characteristics or attitudes could influence which of the treatment methods they received. 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 families, these could not explain why the group that received intensive, in-home, family-based services had higher rates of reunification than the families receiving the traditional form of services.

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.

**Mechanism**

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, p. 35; Marini & Singer, 1988). Many social work researchers (and scientists in other fields) argue that no nomothetic causal explanation is adequate until a causal mechanism is identified.

Up to this point, we have actually described only a portion of Walton’s research. In a subsequent article, Walton and her colleagues (Lewis, Walton, & Fraser, 1995) explored the process: that is, the essential actions leading to the successful outcomes. To do this, they focused only on the treatment group. They compared the kinds of activities, goals, and clinical services received by families that were reunified and those in which the child did not return to the home. They found that successful outcomes occurred when caseworkers could focus on treatment goals involving skill building, such as communication skills, parenting skills, and anger...
management, goals about improving school performance, and increased compliance with family rules. Service failure occurred when caseworkers had to devote their time to providing transportation or spending time on the phone with the client-family. In addition, failure to reunify was more likely when caseworkers had to focus on clarifying problems, defusing crises, and crisis or conflict management. In general, then, when caseworkers had to deal with crises, failure was more likely; when caseworkers could spend time on skill building, success was more likely.

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, Walton and her colleagues (Lewis et al., 1995) suggest that future research should identify the factors associated with goal achievement and success, “for goal achievement appears to be a proximal outcome measure that mediates family risk factors and more distal outcomes such as child out-of-home placement” (p. 279).

**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.

Walton and colleagues (1993) noted the possibility that contextual factors influenced their findings. The state in which the study was done, Utah, has unique religious and social characteristics that might influence the findings (and that limited the generalizability of the findings). Furthermore, the researchers suggest that the decision to place a child is not just the result of the family situation but of many system-level factors as well; these factors could play a role in placement.

**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**

Roberta Sands and Robin Goldberg-Glen (2000) used a cross-sectional design to study the relationship of social supports to psychological stress (or distress) with a sample of
grandparents raising grandchildren. Their theoretical framework was based on stress theory, or the idea that stressors such as individual or family changes generate stress in people’s lives, but the extent or nature of that stress may be altered by the resources available to the individual or family (see Exhibit 5.2). In this study, the stressful life event was having to raise a grandchild, and the stress was psychological anxiety. One type of resource that varies among individuals or families is social support, and the researchers hypothesized that the strength of social support would explain variation in the level of psychological anxiety.

Sands and Goldberg-Glen (2000) gathered their data using face-to-face interviews. Trained female interviewers matched by race to the respondent did the interviewing. Sands and Goldberg-Glen found that a lack of social support was related to higher levels of psychological anxiety.

How well does the study meet the criteria for establishing a causal connection? Sands and Goldberg-Glen (2000) showed an association between variation in the independent variable, the level of informal social support, and the dependent variable, psychological anxiety. However, their design could not establish directly that the variation in anxiety occurred after variation in social support. Maybe people who are less anxious to begin with attract family members and use other resources. It is difficult to discount such a possibility when only cross-sectional data are available.

A nonexperimental study like Sands and Goldberg-Glen’s cannot use random assignment to comparison groups to minimize the risk of spurious effects. We cannot randomly assign people to many of the social conditions they experience, such as raising a grandchild or living in a particular neighborhood. 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 when we hold constant the values of other variables. If it does, the relationship could not be caused by variation in that other variable.
Statistical control  A technique used in nonexperimental research to reduce the risk of spuriousness. The effect of one or more variables is 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.

Sands and Goldberg-Glen (2000) designed their study to control for other factors that might explain the relationship between social support and stress. They included contextual factors such as race, length of caregiving, age of the respondent, and employment status. In fact, when social support was not included in their analysis, several of these variables were associated with anxiety, specifically, race, age, and number of years of caregiving. When additional stressors were added to the analysis, they found that race of the caregiver was no longer related to anxiety, and when social support variables were added, the length of caregiving was no longer related to anxiety. We might conclude then that the relationship of race was spurious due to other stressors and the length of caregiving was spurious due to social supports.

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 research, help us to understand how variation in the independent variable results in variation in the dependent variable. In the above example, social supports served as an intervening variable for contextual variables such as race, employment status, length of caregiving, and age.

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 social supports tend to result in lower levels of psychological anxiety or whether different forms of social supports are related to anxiety. In fact, Sands and Goldberg-Glen (2000) reported that although higher social support in general was related to less anxiety, not all forms of social support were related to level of anxiety. They found that the more family cohesion, the lower the anxiety level, but other indicators of social support, such as use of community services or belonging to a support group, were unrelated to anxiety level. You could then conduct research to identify the mechanisms that link, for example, family cohesion to stress. (Perhaps the respondents who perceived the highest levels of cohesion were also receiving the most concrete support from their families so that they were less anxious.) 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.

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. We will describe survey research in Chapter 8.

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 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 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 5.3).

**Exhibit 5.3 Three Types of Longitudinal Design**

![Diagram of three types of longitudinal design: Repeated Cross-Sectional Design (Trend Study), Fixed-Sample Panel Design (Panel Study), Event-Based Design (Cohort Study)]
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, p. 1).

A study using a repeated cross-sectional design is 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? Have the characteristics of nursing home residents changed? Has client satisfaction with the services provided by a family service agency changed? Has the extent of drug and alcohol use changed in a community? These questions concern changes in the population as a whole or in defined populations, not changes in individuals within the population. We want to know whether current clients of the agency are more likely to be satisfied than clients were last year or 10 years ago, not whether this change is due to changes in the composition of clients or changes in their receptivity to help-seeking. When we need to know whether individuals in the population changed, we must turn to a panel design.

**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 Designs

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

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.
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.

Because a panel design follows the same individuals, it is better than a repeated cross-sectional design for testing causal hypotheses. For example, Knowlton Johnson (1997) used a fixed-sample panel design to investigate whether professional services help reduce psychological symptoms among victims of crime. He drew his sample from a larger sample of respondents to a statewide study of victimization and psychological distress. His sample included households that had experienced either a violent crime or a property crime as well as a sample of households that had not been victimized. Johnson (1997) found that the use of legal services (private lawyers, legal aid, or prosecutor’s office) was associated with higher levels of distress in the short term, but 6 to 12 months later, it was associated with reduced symptoms. On the other hand, among those people using health-related services (medical, clergy, or mental health professionals), symptoms of distress increased except for those who received the services immediately and on a continuous basis. In this study, the level of psychological distress was measured before and after the use of services. If the researchers had used a cross-sectional design, it would have been impossible to establish a baseline measure of psychological distress.

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 5.4, there is a hypothesized association between delinquency in the 11th grade and grades obtained in the 12th grade (the dependent variable).

Exhibit 5.4  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 assumed. In reality, grades in the 7th grade also play a role.
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, they 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, p. 170). However, subject attrition can be reduced substantially if sufficient staff can be used to keep track of panel members. In his panel study, Johnson (1997) lost only 18% of the respondents in the original wave of data collection. 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 helps 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 people with mental illness who had been homeless, only three or four 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, and participants received 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 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.

**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.

**Cohort** Individuals or groups with a common starting point. Examples include the 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”).

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, families, households, groups, communities, or towns.

**Individual and Group Units of Analysis**

In most social work studies (and 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, households, schools, human service organizations, neighborhoods, towns, or states. For example, a researcher may want to learn about the relationship between client satisfaction and agency auspices (public, private nonprofit, private for-profit). The researcher collects data on client satisfaction from different agencies serving similar target groups, such as the mentally ill. The researcher can then analyze the relationship between average client satisfaction and the auspices of the agency. Because the data describe the agency, agencies are the units of analysis.

We can distinguish the concept of units of analysis from the units of observation. In a study comparing family income, data were collected from one member of the family, the unit of observation, but the data were used to describe the family. In some studies, the units of observation and the units of analysis are the same. The important point is to know the unit of analysis because the conclusions that are made are about it. A conclusion that teen pregnancies increase with poverty rates could imply either that teenage girls who are poor are more likely to become pregnant or that a community with a high poverty rate is likely to have a high teenage pregnancy rate—or both. Whether we are drawing conclusions from data or interpreting others’ conclusions, it is important to be clear about the relationship to which we refer.

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, pp. 29–31).

**The Ecological Fallacy and Reductionism**

Researchers should make sure that their causal conclusions reflect the units of analysis in their studies. 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 5.5). The conclusions may or may not be correct, but we must recognize that group-level data do not describe individual-level processes. 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 he or 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).
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.

Readers of research must also beware not to confuse the unit of analysis, even when the researcher has done so correctly. The term underclass first referred to neighborhoods or communities with certain characteristics such as high rates of unemployment, poverty, out-of-wedlock births, and welfare recipiency and lower educational attainment. The term began to be misused when the people living in such communities were described as members of the underclass.

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, pp. 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 5.5). For example, William Julius Wilson (1987, p. 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 suggests 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.

**IDIOGRAPHIC CAUSAL EXPLANATION**

**Causation in Qualitative Research**

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, 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, pp. 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 concerned with context, with understanding the particular outcome as part of a larger set of interrelated circumstances.
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).

Explanation in Qualitative Research

When qualitative researchers seek to develop causal explanations, they often take an idiographic approach. The rich detail about events and processes that field research generates (see Chapter 9) can be the basis for a convincing idiographic, narrative account of why things happened as they did. 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. (pp. 86–87)

An idiographic explanation like Anderson’s (1990) 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.

Single-Subject Design

An alternative to group designs available to social work researchers is to use a single-subject design. The name aptly describes the sample size of these designs, for typically they involve a single case, whether an individual, an agency, or a community. Single-subject designs involve closely monitoring the impact of an intervention on a particular client. For example, a researcher may examine the impact of cognitive behavioral therapy with a depressed, 50-year-old male. The close monitoring allows the researcher to identify
the point at which improvement or change occurs and to examine the factors that are associated with the change. You can see how this is a form of idiographic research; the researcher starts with a particular client having a particular condition, and data are collected to see if an event (the intervention or other life changes) results in a new or changed condition.

Single-subject designs may be used to better specify findings from experimental group designs. Researchers using experimental group designs typically report aggregate findings, so we do not know what happened with each of the individuals. Single-subject designs examine individual changes rather than group changes. Single-subject designs can be used to test findings produced from such a nomothetic method with specific clients in specific contexts.

Making a case for causality using single-subject designs is much more difficult. While often an association and a time-order can be demonstrated, it is much more difficult to establish nonspuriousness. And the nature of the design—a sample size of one, the specific context, the specific provider—makes it very difficult to establish generalizability. We will describe in detail single-subject designs in Chapter 7.

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.

We 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. 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 (nomothetic perspective)
Causal effect (idiographic perspective)
Causal mechanism
Ceteris paribus
Cohort
Cohort study
Context
Cross-sectional research design
Ecological fallacy
Event-based design
Fixed-sample panel design
Idiographic causal explanation
Intervening variables
Longitudinal research design

Nomothetic causal explanation
Nonspuriousness
Random assignment
Randomization
Reductionism
Reductionist fallacy
Repeated cross-sectional design
Spurious relationship
Statistical control
Subject fatigue
Time order
Trend studies
True experiment
Units of analysis
Units 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 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 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 to reduce the risk of spurious effects due to extraneous variables.
- Ethical and practical constraints often preclude the use of experimental designs.
- 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.
- 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.
- Units of analysis refer to the level of social life about which we can generalize our findings and include such levels as individuals, groups, families, communities, or organizations.
- 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.

- 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.

**Discussion Questions**

1. In cross-sectional research designs, satisfying the time order criterion for causal explanations can be very challenging. There are, however, four special circumstances in which cross-sectional data can provide adequate information to infer time order of effects. Review these special circumstances, and discuss the potential difficulties present in each of these situations.

2. Compare and contrast the nomothetic and idiographic explanations of causation. How do the approaches or goals of each vary? Review the Walton et al. (1993) study of the effects of intensive family-based services on family reunification described at the beginning of the chapter. What were the advantages of using a nomothetic perspective? What advantages might have been gained using an idiographic perspective? If you were to conduct the study, which approach would you choose?

3. Develop an explanation of the relationship between juvenile delinquency and the poverty rate by specifying intervening variables that might link the two. Is your proposed causal mechanism more compatible with a “culture of poverty” explanation of this relationship or with a conflict theory explanation? Explain your answer.

**Practice Exercises**

1. Review articles in several newspapers, copying down all causal assertions. These might range from assertions that the increasing divorce rate is due to women’s liberation from traditional gender roles to explanations about why welfare rates are decreasing or child abuse reports are increasing. 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. Search *Social Work Abstracts* 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?

**Web Exercises**

2. What are the latest trends in crime? Write a short statement after inspecting the FBI’s Uniform Crime Reports at www.fbi.gov (go to the Library and References section). You will need to use the Adobe Acrobat Reader to access some of these reports (those in PDF format). Follow the instructions on the site if you’re not familiar with this.

To assist you in completing the Web exercises, please access the study site at http://www.sagepub.com/prsw where you will find the Web Exercises reproduced and suggested links for online resources.

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 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. 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.