violation by taking some protective action: "but I clean out my purse now since then and I leave very little of that kind of stuff in there."

Finally, when the victim is "Going On," he/she reflects on the changes the crime produced: "I don't think it made me stronger. It made me smarter."

Is this a causal explanation of the impact of crime on its victims? Not all would agree, but this illustration shows how the effort to "understand" what happened also inevitably gives us a better sense of why things happened as they did.

**RESEARCH DESIGNS AND CRITERIA FOR CAUSAL EXPLANATION**

In order to achieve causally valid nomothetic explanations, research designs must have certain features. In this section I will introduce these features and discuss what makes them so important.

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. Because experimental research designs have several features that are designed to meet these criteria, they are particularly strong designs for testing nomothetic causal explanations.

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. These criteria can be met in different ways by both experimental and nonexperimental designs.

In the following subsections, I will indicate how researchers attempt to meet the five criteria with both experimental and nonexperimental designs. Illustrations of experimental design features will use a 1999 study by Brad Bushman, Roy Baumeister, and Angela Stack on the effect of catharsis on aggression. Most illustrations of nonexperimental design features will be based on the study by Robert Sampson and Stephen Raudenbush (1999) of neighborhood social control, which I have already introduced.

Bushman and his colleagues recruited 707 undergraduate students from introductory psychology courses. The experiment had several stages, but I will focus on only a few (see Exhibit 6.7). 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 2 minutes. (This was the opportunity for catharsis—getting anger out of your system.)

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 procathearsis
message "blasted" their competitors more than did those who had read no message, while those who had read the anticatharsis message blasted their competitors the least (see Exhibit 6.8). 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? What about the conclusion by Sampson and Raudenbush (1999) that social and physical disorder does not directly cause neighborhood crime? I will answer these questions by reviewing how these two studies attempted to meet each of the causal criteria. I will also identify the key features of a "true experiment."
Exhibit 6.9  Experimental Conditions: A “3 × 2” Design

<table>
<thead>
<tr>
<th>Message Content (3):</th>
<th>Procatharsis</th>
<th>Control</th>
<th>Anticatharsis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X₁, X₂</td>
<td>X₃, X₄</td>
<td>X₅, X₆</td>
</tr>
</tbody>
</table>

Aggressiveness Target (2): Direct Displaced

Source: Based on Bushman, Baumeister, & Stuck, 1999:371.

Association

We say that there was an association between interpersonal aggression and type of message in Bushman’s experiment 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, for each of the three groups that differed in terms of message content, Bushman actually compared subgroups that differed in having direct or displaced targets for their aggression (see Exhibit 6.9).

In nonexperimental research, the test for an association between the independent and dependent variables is like that used in experimental research—seeing whether values of cases that differ on the independent variable tend to differ in terms of the dependent variable. The difference with nonexperimental research designs is that the independent variable is not a treatment to which the researcher assigns some individuals. In their nonexperimental study of neighborhood crime, Sampson and Raudenbush (1999) studied the association between the independent variable (level of social and physical disorder) and the crime rate, but they did not assign individuals to live in neighborhoods with low or high levels of disorder.

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 (1999) 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.

You have already seen that in nonexperimental research, time order can be readily established if a longitudinal design is used. Sampson and Raudenbush (1999) could not determine time order unambiguously in their cross-sectional study of the effect of social disorder on crime rates, but Sampson and Laub’s (1990) longitudinal study of the effects of childhood deviance on adult crime provided strong evidence of appropriate time order.

Nonspuriousness

Nonspuriousness is another essential criterion for establishing the existence of a causal effect of an independent variable on a dependent variable; in some respects it is the most important criterion. 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 other than an effect of the presumed independent variable on the dependent variable. If we measure children’s shoe sizes and their academic knowledge, for example, we will find a positive association. However, the association results 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. Many 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 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 more resources produce better student outcomes? 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 (1999) 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 6.10). 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 procatharsis message became more aggressive, while the others didn’t.
A nonexperimental study like Sampson and Raudenbush's (1999) 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 these other variables.

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 6.1, p. 167). Exhibit 6.11 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 visible disorder. Therefore, we must conclude that the apparent effect of "broken windows" was spurious due to level of informal social control. Neighborhoods with low levels of social control were more likely to have high levels of visible social 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.
Exhibit 6.11 The Use of Statistical Control to Reduce Spuriousness

Source: Based on Sampson & Raudenbush, 1999.

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.

We can strengthen our understanding of nomothetic causal connections, and increase the likelihood of drawing causally valid conclusions, by considering two additional criteria: causal mechanism and causal context. These two criteria are emphasized in the definition of idiographic causal explanation, with its attention to the sequence of events and the context in which they happen, but here I will limit my discussion of these criteria to research oriented to nomothetic causal explanations.

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, 1979:35; 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 described only a portion of Bushman’s (1999) 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 was 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.

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. For example, in a study that reanalyzed data from Sheldon Glueck and Elenor Glueck’s (1950) pathbreaking 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 6.12). 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
Exhibit 6.12 Intervening Variables in Nonexperimental Research: Structural Disadvantage and Juvenile Delinquency

![Diagram showing Intervening Variables]

Source: Based on Sampson & Laub, 1994.

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 6.13 the extraneous variable, joblessness, creates a spurious relationship. By contrast, in Exhibit 6.12, the intervening variable is part of the process that links the independent 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).

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 people had the same aggressive response to the cathartic message whether they competed against the person who was actually the source of their angry feelings (the person who wrote the negative evaluation) or against an innocent third person. An earlier experiment also showed that being angry was a precondition for the effect of the procatharsis and anticatharsis statements.
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).

**CONCLUSIONS**

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), as well as an alternative approach to explanation—case-oriented understanding—that does not involve causal reasoning. You also have learned how to establish the time order of effects in nonexperimental research, and you have been exposed to the problem of spuriousness and the ways that randomization and statistical control deal with it—and you have learned that the use of randomization in experimental designs makes this design preferable for establishing causal validity. Finally, you have studied the five criteria used to evaluate the extent to which particular research designs may achieve causally valid findings.

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, 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...