T. Petticrew, How to think like a social scientist? Harper Collins, 1995

Chapter

4

Searching for Causes and Changes

Like cats, social scientists focus on objects that move and change. Static phenomena are difficult to study, because they are difficult to measure. This situation presents few problems for social science. Social life, as discussed in Chapter 1, is always changing and in flux (dynamic).

This dynamic quality of social life allows social science to search for causes. What causes what? Why did that happen? What social policies would help solve a given societal problem? Most of the important questions addressed by social science involve the search for causes, but *cause* is a complicated concept. So the firm establishment of the causes of social phenomena poses a major challenge for social science.

We can glimpse at the complications by reviewing one of the most interesting investigations of college students ever conducted—the longitudinal (over time) study of Bennington College in Vermont. The research began in the 1930s. Theodore Newcomb (1943) collected extensive data on the political attitudes of all 527 students at the private women's college. It was an ideal setting for such research. Bennington was not only small, but in an isolated location and self-contained. The college took public issues seriously and had close faculty-student interaction. If changes in political attitudes take place at college, the Bennington campus would surely reveal them.

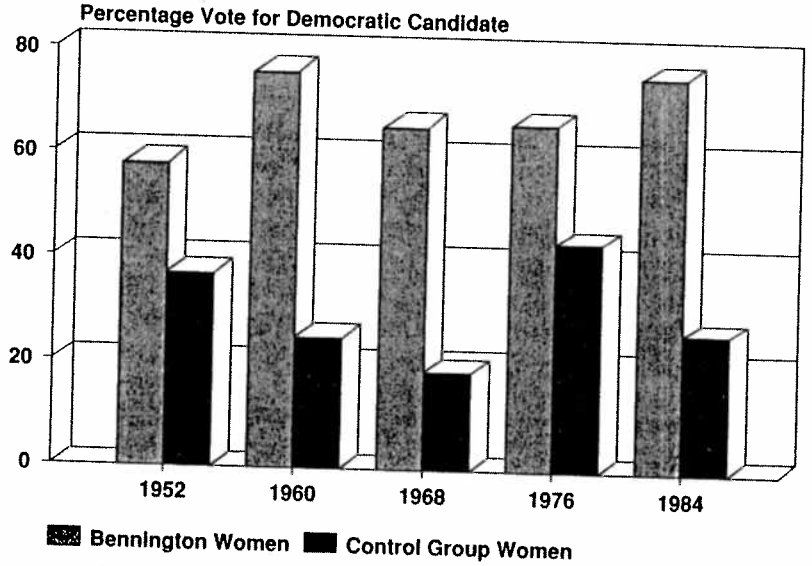
Tuition at this exclusive college was high. Most students came from wealthy families. This fact set up the principal tension in this bucolic setting—sharp differences in political views between the students' conservative family backgrounds and a liberal faculty.

The study's initial findings are consistent with the reference group concept described in Chapter 1. Most of the students became increasingly liberal as they progressed through their years at Bennington. There were differences, however, among the women. Those who most strongly identified with the college revealed the most marked shift toward liberalism. Those who continued to identify more strongly with their families than the college showed the least shift toward liberal attitudes. Hence, those students for whom Bennington became a major reference group moved to the political left. Those for whom their family remained as the dominant reference group did not.

Twenty-five (1959–60) and forty-nine (1984) years later, Newcomb and his colleagues reinterviewed the former Bennington students (Newcomb, Koenig, Flacks & Warwick, 1967; Alwin, Cohen & Newcomb, 1991). From our discussion in Chapter 3, this study used a *passive time series design.* That is, the researchers could not manipulate the independent variable (the Bennington experience). They compensated for this by measuring the dependent variable (political attitudes) on four occasions (on entering and leaving college as well as 25 and 49 years later). This extensive longitudinal research allows us to find out if the students' liberalism persisted in later life; or, did their social attitudes shift back toward those of their parents once they left college?

In general, their liberalism remained remarkably stable from their college days. Compared to a control group of comparably educated women of the same age, the Bennington graduates are significantly more liberal in their social attitudes. Figure 4.1 shows the differences between the two groups in presidential voting over the years. Even as they approach their 70th birthdays, the Bennington women in 1984 favored the liberal Democratic candidate,

Figure 4.1 Presidential Voting
Bennington Women Vs. Comparable Control



Source: Institute for Social Research, 1991–2.
Note: Control = Same Age College Grads.

Walter Mondale, over the conservative Republican candidate, Ronald Reagan, by a three-to-one margin. In sharp contrast, the control group supported the Republican by a three-to-one margin. (Note how the researchers developed a control group without a pre-test of what these women's attitudes were upon entering college a half-century earlier.)

The most interesting result reveals an important life-span principle of social influence. For those women who married liberal men, maintained liberal friends, and led their adult lives in liberal environments, the Bennington changes persisted. For those few women without such supportive settings, the changes of college did not persist. The significant finding is that most of the Bennington women who adopted a liberal political perspective as students went on as adults to shape supportive settings for their liberal views.

Thus, the college influence led to an active selection and shaping of future environments that in turn supported the original influence. It is not simply a matter, then, of "how the twig is bent," but also how the twig itself bends the environment. The women did not just react; the process also involved important proactive considerations. An investigation of students from colleges different from Bennington replicated this important proactive principle. Student attendance at Catholic colleges increases later adult religiosity largely because such attendance makes it more likely that the students will marry a highly-religious spouse (Greeley & Rossi, 1966).

Consider, however, the difficulties of specifying clearly what the causes were in this Bennington College phenomenon that unfolded over two generations. We know considerable change took place from the time these students entered college through their four years of study and into adult life. Yet what actually *caused* these changes? Was it their identification with Bennington that acted as the initial cause? Once the college served as an important reference group, were the women then open to the liberal social attitudes of the Bennington climate? This is the causal sequence suggested by Newcomb. Alternatively, could the liberal atmosphere of the college itself have been the initial cause? In this case, those most affected by the shift in political perspective could then have adopted Bennington as their reference group. Here the issue is causal order—which of the two measured variables came first? Most likely, the two processes acted jointly, each inducing the other as the students progressed through their college training (**reciprocal causation**).

A second causal problem in interpreting this famous study concerns unmeasured variables. No research can measure all the variables that might be important in such a change process. Could there have been key variables that Newcomb did not measure that were the principal causal agents in the change process? Here is where theory becomes essential. As discussed in Chapter 2, we can never measure all possible factors. We need a good theory to guide us in advance as to which variables we should measure. Newcomb chose reference group theory, and it served him well.

A healthy skepticism encourages you to speculate about what unmeasured variables might have been critical in the Bennington situation. Here, too, you should draw on your personal experience for ideas. You have experienced the role of the college student—though probably not in such an intense situation

as Bennington's. Can you think of any additional factors that might have been important?

A third set of causal complications involves the general persistence of the changes 25 and 49 years later. Now the women themselves appear as the chief causal agent. They chose and arranged their adult environments to be compatible with their liberal political orientations. These environments in turn fostered and advanced their views further.

The basic causal question, however, is why did the Bennington alumnae do this? Presumably, they did so because they had internalized political liberalism as a central part of their self-definition. These views were no longer simply those of the Bennington faculty. They were now *their* views, and so valued that they influenced their later life choices. Ultimate causes of such a process are difficult to discern without a causal sequence perspective, a type of domino effect. The Bennington atmosphere led the women to adopt, and in time internalize, liberal political attitudes. These attitudes in turn influenced how they shaped their adult environments. And these environments in turn maintained and advanced their liberal attitudes.

One of the interesting ways this continuing process operated involved their children. The third wave of the Bennington study found that the children of these women were important contributors to their mothers' continued liberalism. Most of these children reached political maturity in the Civil Rights era of the 1960s. Their liberalism, shaped in part by their mothers, offered support for their mothers' maintaining their social views. Parents learn from their children just as children learn from their parents. Another longitudinal study, focusing on these activists of the 1960s, found that they, too, maintained their activism into later life (Franz & McClelland, 1994).

These difficulties of assessing causation in the Bennington study introduce us to the complexity of the idea of cause.

4.1 THE COMPLEX CONCEPT OF CAUSE

Thinking like a social scientist entails thinking about what causes social processes. Basic as this task is, it turns out to be difficult. To see why it is not easy to specify causes, we will discuss three aspects of the topic. First, assigning causes in everyday life is typically easy. We shall discuss why this is so, and also how this natural perception of causation features several consistent biases. Second, we turn to the logical complications underlying the concept of cause—a specialty of philosophy. Finally, we shall review the compromise that social science strikes between the rigorous strictures of philosophy's analysis of cause and the practical problems of studying social life.

4.1.1. The Natural Perception of Cause

In everyday life, assigning cause to events happening around us usually presents no problem at all. Take, for instance, when we play pool. As our cue stick hits the cue ball, it moves and hits the ten ball into the left corner pocket

Determining causation creates no problem for us here. We instantly regard the stick as having *caused* the cue ball to move, and the cue ball in turn having *caused* the ten ball to roll into the pocket. Research by Michotte (1963) shows that the immediacy and naturalness of this process is an integral part of human perception.

We carry over this process into our perception of human beings. With little difficulty, we constantly assign cause to human actions around us. Indeed, it is critical that we do so. Knowing why others act as they do toward us is necessary to understand the situation and react appropriately.

Valuable as it is, however, the human ability to attribute cause is often biased and misleading. In their intensive study of **causal attribution,** social psychologists have discovered consistent biases in the process. Two broad types roughly characterize causal attributions for human actions. With **situational causal attributions,** we explain human actions with special features of the situation. With **dispositional causal attributions,** we explain the actions with special traits of the individual. A principal bias in our perception is that we frequently exaggerate the importance of dispositional causes and disregard situational causes (Ross, 1977). People are salient. They are in the foreground of our visual field, and thus they impress us as causal agents. By contrast, situational factors are usually in the background and out of view. Social psychologists call this phenomenon the **fundamental attribution bias.**

This general preference for dispositional explanations is not the only human bias in causal perception. One study showed that we usually grant those we love the benefit of the causal doubt (Taylor and Koivumaki, 1976). If someone close to us acts negatively, we typically attribute it to a temporary situational cause. "Oh, she just lost her temper, because she's very tired after a frustrating day at work." If, however, the act is positive, we attribute it dispositionally. "What a thoughtful thing to do—she is such a sweet person!"

We are less generous with others (Pettigrew, 1979). At the extreme, with members of disliked outgroups, we often reverse our causal reasoning (the **ultimate attribution bias**). Following a negative act by an outgroup member—such as pushing ahead in line, we attribute it dispositionally. "That's the way those people are—always pushy and aggressive." If the outgroup member commits a positive act—such as holding a door open for us, we maintain our prejudice by denying it, regarding it as a "special case" or explaining it away situationally. We fail to notice the good Samaritan was a member of the disliked outgroup, or we regard it as "the exception that proves the rule." Otherwise we could say, "What else could he do with all those people watching, while I struggled with a big package?"

Such causal reasoning about human behavior has direct effects on how we respond. If we think a person yelled at us angrily for situational reasons, we are likely to understand and remain calm. If we think a person yelled at us as an expression of an obnoxious personality style, we may well act back in kind.

This differential reaction occurs for social action as well. If we think the poor living in wretched housing are not responsible for their condition, we are likely to favor efforts to improve the housing. If we think the slum dwellers brought the bad housing on themselves, we are content "to blame the victim" and take no action. This latter course of blaming the victim may be especially prevalent in Western cultures. Research shows that we typically see victims of accidents and other ill fortune as the causes of their plight. To maintain a "belief in a just world," we are prone to believe that people get what they deserve and deserve what they get (Lerner, 1980).

4.1.2. Philosophical Complications

Immediate and natural as causal attributions are for human beings, logical problems underlie the idea of cause. Philosophers make valuable contributions to scientific thinking on precisely such logical issues. Without reviewing the philosophical analyses of cause in depth, let's look at points of special relevance for social science.

David Hume, the 18th-century Scottish philosopher, raised the basic problem that we can never directly observe causation. Instead, we can only *infer* cause. When we see one pool ball hit another that then moves, we have not actually observed causation. All we have seen is a connection between the two balls in time and space. We then, without hesitation, *infer* that the first ball caused the second ball to roll.

Hume's point is well taken and worth remembering. Causation always involves inference—not direct observation. Yet his insight led him astray. In seeking some observable, objective means to determine causation, he decided that a correlation between two factors was sufficient evidence. If the presumed causal agent preceded and connected with the effect, then a correlation between them established cause.

Think about Hume's argument for a minute. Night always follows day and the correlation is complete ($r = +1.0$), yet day does not *cause* night. A positive correlation exists between weight and height, but genes and nutrition explain the relationship. Similarly, a high correlation exists between wearing jeans and liking rock music. One would doubt, however, that one caused the other. Correlation alone, even when one factor precedes the other, cannot establish cause. There may be additional variables (such as age and conformity to popular norms) that relate both to wearing jeans and liking rock music that account for their correlation. We also should look at error cases—people who wear jeans and abhor rock music, and those who adore rock music and never wear jeans.

Another problem involves the interval between the cause and effect. There is agreement that a causal factor must precede its effect. Yet, how long can this interval between cause and effect extend for us to consider it a causal relationship? In physics, theory can often specify such intervals in advance. Social science can rarely be so precise. The Bennington College effects, for instance, were still evident after a half century (Figure 4.1). In fact, delayed effects are common in social science. Though difficult to predict in advance, such effects demand explanation. A useful explanation of delayed effects must specify when (moderator variables) and how (mediator variables) such effects per-

sisted. In the Bennington study, the mediating variable involved the politically liberal environments the women chose and shaped for themselves after leaving college.

A third complication involves reciprocal causation. This happens when two or more variables mutually influence each other. Such a possibility is more than a concern for philosophers. Reciprocal causation may be the dominant pattern of social life. In economics, for example, what consumers spend depends on their income. It is also true that their income depends in part on their consumption. Such common causal patterns raise complex statistical issues (Ragosa, 1985). Single-equation predictions no longer suffice. Rather, systems of simultaneous equations are necessary, as developed by econometricians.

Reciprocal causation is best studied with longitudinal designs, like the Bennington research, in which the investigation measures the subjects at repeated intervals over time. The liberal political attitudes of the Bennington students and their later liberal environments offer an instance of reciprocal causation revealed by longitudinal research.

On another issue, some philosophers insist that cause and effect must involve perfect correlations. That is, when the cause is present, the effect must *always* follow. When the effect is present, the cause must *always* have preceded it. Such a criterion of cause simply does not reflect the messy nature of social life. Absolute relationships are rare. Instead, probability relationships are the rule. That is, cause A will be present with effect B under some conditions but not others at varying probabilities. Measurement error also will reduce our reading of the relationship between A and B. Correlational statistics measure this relationship, but in practice it can never attain the rigorous standard set by some philosophers. The strict absolutist position on causation is not realistic for social science.

Amidst such causal complications, social science is happy to accept the simplifications of another Scottish philosopher. John Stuart Mill advanced three conditions for inferring cause. The first two are easy to meet: *the cause must precede the effect,* and *the two factors must relate to each other* (though not necessarily perfectly). It is Mill's third criterion that frames the basic task for social research. To establish cause, *one must rule out other explanations for the cause-effect relationship under study.* We discussed this central issue in the last chapter by showing how controlled comparisons exclude plausible rival hypotheses.

Mill helps by proposing three "canons" to meet his third criterion for identifying cause. First, his *Method of Agreement* states that an effect will be present when the cause is present. Second, his *Method of Difference* states that the effect will be absent when the cause is absent. The third canon, the *Method of Concomitant Variation,* implies that when the first two canons hold true the causal inference is stronger. This is so because these canons remove rival explanations for the relationship.

Mill understood, as Hume did not, that mere correlations between variables were not sufficient by themselves to determine causal links. Nor did he require that the correlation coefficients—the *concomitant variation*—be near perfect ($r = +1.0$). His lasting point was that one must probe the relationship deeply to reach secure causal connections.

To follow Mill's reasoning fully, we would need to test all variables that could possibly provide a rival causal explanation. We saw in the Bennington study how this is impossible in practice. We need theory to guide us to select the most likely variables for testing. Since we can never be sure that we have exhausted all relevant variables or theories, our causal conclusions must remain tentative. For that matter, *all* science is tentative. Social science is as dynamic as its subject matter. New ideas, theories, and research overturn and improve what we call "knowledge" in a never-ending process.

The 20th-century Austrian philosopher Karl Popper (1959; Miller, 1985), best reflects this tentative perspective on science. Though Popper focused on physical science, much of his philosophy is directly useful for the social sciences (Campbell, 1974; Cook & Campbell, 1979; Pettigrew, 1991d). He maintains "truth" is a goal we can never attain. However, we at least can know when we are "getting warm." We know we are approaching "truth" by collecting research evidence that is not inconsistent with our causal theory. In short, we do not *confirm* a theory through research; instead, we *fail to falsify* the theory.

Thus, we do not discover "truth" with research. Instead, we gain confidence in our theory by learning that we have not yet falsified it. Following Mill, we do this by what Cook and Campbell (1979:22) call a *winnowing process.* We continue to test our causal theory repeatedly in ways that remove plausible rival hypotheses. Think back to the Archer and Gartner (1984) example in Chapter 3. Remember how these investigators cast doubt, one by one, on rival ways to explain their finding that homicide rates rise after wars. This emphasis on falsification may sound like "backing into" a causal inference. However, this cautious approach makes good sense when we realize all the logical problems involved.

Moreover, research observations—the *facts* of science—are themselves partly infused and shaped by the very theories they test. That is, scientific observations presuppose a theory and its conditions. "Even a simple reading of a watch presupposes that the hands will move uniformly," and minutes and hours are all constant according to astronomical theory (Von Bertalanffy, 1975:168).

Some modern philosophers go so far as to claim that this "theory-laden" quality of scientific observations means that theory testing is impossible. Social scientists rarely accept this argument. They join Cook and Campbell (1979:23–25) in regarding some observations as "stubborn facts." That is, these observations are often replicated. They recur across many settings by various researchers using different methods and testing different theories. To be sure, theory and methods partly shape facts. Yet not just a single theory or a single research design is usually involved. This overlap allows some degree of theory testing in social science, even if it is not as direct as in physics and chemistry.

Cook and Campbell (1979:25–28) remind us of a further point. Social scientists prefer *manipulable* causal factors. This preference has practical advantages. Causes that we cannot control are obviously of no direct value for applied work. Applying social research to alleviate the world's problems requires causal variables we can alter.

We usually can exercise this preference for manipulable causes. Actually, there is no one cause of a social effect. Since social phenomena operate at multiple levels (Chapters 1 and 6), any particular social effect will have causes at several levels of analysis. In addition, recall from Chapter 1 that many factors cause most social phenomena even at one level of analysis. With so many causal factors, it is likely some will be controllable.

Combatting inflation provides an example. Many factors at many levels of analysis cause goods and services to rise in cost. Such causal factors vary from governmental decisions about interest rates and the money supply to individual decisions on what to buy at what price. Some of these causes are easily altered, others are not. The U.S. Federal Reserve Board can decide to increase interest rates and reduce growth in the money supply to combat inflation throughout the nation. Trying directly to alter the buying patterns of millions of individual Americans is not so readily available.

Furthermore, we usually do not need to have complete information about all causal factors to act effectively. We turn on and off electric lights each day by knowing only that flipping a switch makes it happen. A detailed knowledge of electricity is not necessary. Our limited knowledge catches up with us, however, when the expected effect does not happen. We flip the switch, and the light fails to go on. We go one level deeper. We check the cord and plug and replace the bulb. If these do not work, we call an electrician. Now we need someone whose causal knowledge extends down to more basic levels.

4.1.3. The Social Science Compromise

Picking its way through this thicket of philosophical problems, social science adopts a compromise position to determine causation. This position varies across the various disciplines. Yet social science must strike a compromise between the most severe strictures of philosophy and the practical reality of studying an untidy world. To think like a social scientist, you must keep this compromise in mind.

Our brief sketch of the logical issues involved with causation suggests six major aspects of this compromise.

1. We cannot directly observe causation. We must always infer it.
2. Causes must always precede effects, yet different intervals may occur between them. Moreover, factors can reciprocally cause each other over time.
3. We never know "truth." We can approach it by a winnowing process that invalidates plausible rival hypotheses.
4. All observations partly reflect the theories and methods that led to them. Nevertheless, stubborn facts that replicate across settings, researchers, methods, and theories allow testing of causal theories.
5. Absolutes are not the rule in social life. One must use probabilities. So no findings and theories are final truths. All can and should be challenged in a continuing process of scientific development.
6. Finally, social science prefers manipulable causal factors, since they are of special importance for useful applications.

Several implications flow from this perspective on causation. Obviously, we must be tentative in our causal conclusions. Probabilities, not absolutes, are the rule. Moreover, causal inferences must carefully consider the intervals between cause and effect and the potential for reciprocal causation. We need to spread the net of rival explanations widely, so that we can thoroughly narrow down the most promising causal possibilities. To avoid having our observations become too theory and method laden, we also must measure our basic concepts with multiple indicators and test them from several perspectives. In all this, a healthy skepticism helps again in our approach to social science.

There are, however, two different strategies taken in the social sciences to carry out this causation compromise. We shall label these approaches the *building-block* and *bold strategies*.

The building-block approach comes from an influential school of philosophers of science called the Vienna Circle Positivists. They flourished in Vienna, Austria between World Wars I and II. They held to the view that research should try to *confirm* theory. Their arguments emphasized that research should accumulate in building-block fashion. Each study in a project should maintain the same measures and procedures as earlier studies except the one significant change under test. Then you can be confident that the one change caused any difference in the results. Bit by bit, each investigation provides one more block in the edifice that is the causal theory.

Sharply distinguishing his position from that of the Vienna Circle, Popper (1959) attacks the building-block strategy on two grounds. First, he argues it is not likely to tell us much, because it is too modest in scope. Even if supposedly confirmed, the hypotheses so tested bear little "content" and new information. Second, such a strategy risks having its results shaped by the repeated use of the same limited measures and research designs across many studies. Those who follow the building-block strategy risk having their results influenced by the narrow range of methods they use.

By contrast, Popper (1959; Miller, 1985) calls for a bold strategy. His guiding analogy for testing scientific ideas is Darwin's "survival of the fittest" principle for the development of life forms. He urges the use of the widest range of possible measures and research designs to test broad theories. Such a bold approach leads to bold theories rich in content and new ideas. Hence, *falsification* of these bold efforts is more likely. Such tests, Popper insists, will teach us more, even if falsified, than will the confirmation of timid, incremental tests.

One popular application of the bold strategy is **meta-analysis.** This technique provides statistical analyses of the results of many independent studies of a single phenomenon. Meta-analysis estimates the size of a phenomenon across many studies and conditions. It is most effective when it reviews research that uses diverse types of subjects, measures, and methods. Meta-analysis summarizes all the known research on a given topic with greater precision and objectivity than subjective reviews of the same studies.

How successful have this compromise and these two strategies been in the social sciences? As with most sweeping questions, there are differences of opinion. I believe, however, that there has been considerable progress in the social sciences during the 20th century. The growth of the various knowledge bases reflects this progress. A quick review of published work across this century in any of the social sciences reveals this growth. I was convinced of this by a review I conducted of all research in race relations published in American journals of sociology from 1895 to 1979 (Pettigrew, 1980). It uncovered an enormous development in both methods and theory over these years.

There are many reasons for this progress. We will consider three specific procedures in particular. Evaluative statistics, replication, and longitudinal research designs have all helped social science to meet the terms of its causal compromise.

4.2 STATISTICS EVALUATE CHANGE

Chapter 3 stressed the importance of statistics in modern social science. It showed how statistics help us to achieve controlled comparisons with nonexperimental data from the field. Here we see a second vital role for statistics. We have talked about research results that support or falsify hypotheses and theories. But how do we evaluate these results as supporting or falsifying? How do we know if these results are not chance effects? For instance, how do we determine that the attitude changes of the Bennington students were large enough to conclude that a genuine effect had occurred?

These questions require statistics for answers. The various social sciences employ different, though related, approaches. Economics typically uses **effect sizes,** such as correlations, to assess the importance of its results. Effect sizes assess differences largely independent of the number of people (or other units) that make up the sample studied. They are especially useful for applications. They make sense for economics for two reasons. Economists typically deal with data sets that boast large numbers of subjects (or other units). Government agencies prepare many of these data sets. Economists are also often interested in applications of their research to policy.

The disadvantage of effect sizes is that they are unreliable if the research involves only a few subjects or observations. When there are many subjects, they are useful indicators of the strength and importance of effects. Think back to Figure 3.7 on European prejudice. There we had 4,000 subjects across seven samples. So Figure 3.7 used effect sizes to portray the results. With

4,000 observations and robust effect sizes, we can be certain of the **statistical significance** of the findings. That is, such results are unlikely to be due to chance.

If, however, the number of subjects or units is small, then even a large effect size may not be statistically significant. Hence, one must use statistical tests that take sample size directly into account. While needed for small sample studies, tests of significance can be misleading with large samples. Large samples can show a small effect to be statistically significant, even when the effect is too small to be of practical importance.

If you have taken a statistics course, you know that such tests of significance as student's *t*-test are among the first lessons. You also learned that two benchmarks—the 5% and 1% levels of confidence—mean that the relationship under test would not occur but five times or one time in a hundred by chance. Such tests characterize elementary assessments of data in most of the social sciences. Statisticians choose the 5% and 1% levels of confidence to strike a balance between two kinds of errors we can make using significance tests. **Type I errors** occur when researchers incorrectly accept their hypotheses as supported (the null hypothesis of no effect is true). **Type II errors** occur when researchers incorrectly reject their hypotheses as falsified (the null hypothesis of no effect is false).

There is a close relationship between effect sizes and significance tests. Only the sample size distinguishes them: significance test = effect size + sample size. Hence, studies with large sample sizes aimed at applications find effect sizes to be more useful. For example, meta-analyses, summarizing the results of many investigations, employ effect sizes. Studies with small samples that seek to determine the level of confidence for a result find tests of significance more useful. The point for our purposes is that statistics are essential. Since social life involves probabilities, we need evaluative statistics to assess them.

4.3 THE IMPORTANCE OF REPLICATION

An effective way to increase the power of statistical methods and to establish research findings more solidly is to repeat the study with a fresh sample of subjects (replication). It is surprising how often investigators do not take this straightforward precaution before they publicly announce exciting results. You may have noticed the many times investigators do not replicate pharmaceutical research, and must later withdraw their original claims for a new "wonder drug." This is also true for several celebrated social science studies that originally claimed to be of major significance for public policy. A careful second study before announcing findings of initial work could have prevented these embarrassments.

The complexities of social life cited in Chapter 1 make replication essential for social research. Recall again the study of European prejudice (Pettigrew & Meertens, 1994). Seven independent samples participated in this study—one from Germany, and two each from France, Great Britain, and the

Netherlands. Such multiple samples allow extensive replication of all findings. Special statistics allow researchers to combine the results of such multiple studies into particularly powerful tests of the hypotheses under test. In fact, the data in Figure 3.7 on the correlates of prejudice represent the combined results of the seven samples.

Sometimes repeating social research is not possible. Perhaps the study checked on the effects of a public event, such as the assassination of Dr. Martin Luther King, Jr. described in Chapter 1. It is impossible to repeat the event. Even here a form of replication is available. The researcher can split the sample in half and test each half as if it were a separate study. Though it reduces the sample size, this procedure can exploit the advantages of two separate studies.

The introduction of increasingly sophisticated mathematical models has heightened the importance of replication in social research in recent decades. Often the initial building of these models capitalizes in part on chance. That is, special characteristics of the initial sample inflate the model's "fit" with and prediction of actual data. Thus, replication with new samples is essential to test the model.

In addition to the original researchers, others may also wish to repeat an investigation. Such fresh replication is a major way social science has of correcting itself. This possibility explains why research articles in technical journals use so much space providing the details of the study's methods. Anyone attempting replication would require these details to proceed.

4.4 THE IMPORTANCE OF LONGITUDINAL STUDIES

Longitudinal research studies how social phenomena unfold over time. So, longitudinal data best meet the criteria for inferring causation.

The Bennington College study provides an excellent example of a longitudinal study. Such research collects its data at different times. Usually, like the Bennington study, it involves the same subjects. Social scientists call such research **panel studies**, with the same subjects forming the *panel*. The National Longitudinal Surveys (NLS) of the U.S. Department of Labor are among the most extensive on-going panel studies in the United States (U.S. Bureau of Labor Statistics, 1992). One of these surveys focuses on youth who ranged in age from 14 to 21 in January 1979. It drew a probability sample of 12,686, including a special military sub-sample. Each year the survey reinterviews these respondents about their schooling, jobs, work attitudes, marital history, income and assets, health, and other matters.

Recall the unrealistic assumption underlying the basic form of multiple regression analysis: namely, that all the predictor (independent) variables operate *simultaneously* in influencing the dependent variable. One major advantage of longitudinal studies is that they allow us to investigate how social processes unfold over time. We saw this in the Bennington College study. Many types of research demonstrate this singular asset of longitudinal research. We shall look at a variety of such studies drawn from (1) laboratory experiments, (2) field studies, (3) trend analyses, and (4) investigations using archival data.

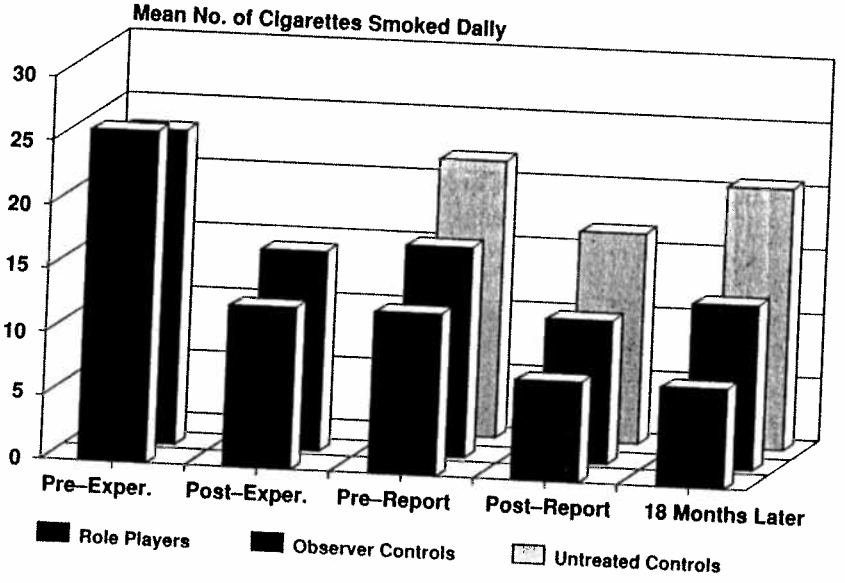
4.4.1. A Longitudinal Laboratory Experiment

Most laboratory research is not longitudinal; that is, the experimenter measures the dependent variable only once. Yet over time, research designs with laboratory experiments can be particularly powerful. An interesting investigation of role playing and smoking by Mann and Janis (1968) offers a pointed example.

In July 1963, these investigators had heavy-smoking females play one of two roles. The experimental group played the role of smokers whose doctor had just informed them that they had developed lung cancer. The control group watched the role playing as passive observers. As you can see in Figure 4.2, both groups reported marked declines in cigarette smoking one month later (pre- and post-experiment bars in Figure 4.2). The experimental group, however, showed a statistically significant greater decline. Playing the role of a cancer victim had a marked effect. Even just observing the role playing had some effect.

Six months later, the reduced smoking continued. Despite minor increases in both groups, a large difference remained between the two original groups as well as between them and a second, new and untreated control group (the

Figure 4.2 Longitudinal Lab Study
Role Playing and Smoking Reduction



Source: Mann

pre-report bars in Figure 4.2). So both the role playing and observer effects persisted.

Right after this measurement, a significant event occurred. This is a further advantage, as well as a hazard, of longitudinal designs. Often a fortuitous real-world intervention occurs that provides special insight into how the experimental variables operate. Immediately after the third measurement in the study, the U.S. Surgeon General released a highly-publicized report on smoking and health that held cigarettes to be especially dangerous.

The study then tested the smokers again. The three groups were equally aware of the Surgeon General's report and its contents. All three reported similar declines in smoking (the post-report bars in Figure 4.2). Clearly, the Surgeon General's report had its desired effect. Yet this result did not change the differences among the experimental and two control groups. The cumulative effects of the treatments and the new health report simply added together. The groups remained different in their decreased smoking.

Finally, Mann and Janis (1968) tested the subjects 18 months after the experiment. As Figure 4.2 shows, the untreated controls had almost returned to their pre-report cigarette smoking levels. For them, the Surgeon General's report apparently had only a short-lived effect. The two treated groups also reported small average increases in smoking over the year. This increase, however, was particularly minor for the role-playing subjects. They continued after 18 months to reflect on average the additive effects of both the experimental treatment and the Surgeon General's report.

4.4.2. A Longitudinal Field Study

In addition to the Bennington study, another famous longitudinal field study shows how small effects can accumulate over time into large effects. Such effects would go undetected in cross-sectional designs; only longitudinal research can uncover them.

Muzafer Sherif's (1966) Robbers' Cave study at a summer camp began by forming two comparable groups of young boys—called the Eagles and the Rattlers. He first kept the groups apart and unaware of each other in different parts of the camp. They each engaged in popular activities, and soon each group developed a sense of "we-ness" and high morale. Then the second phase of the field study began. Sherif brought the two groups together and had them compete in a series of contests. These contests, such as tug-of-war and softball, all featured **zero-sum** payoffs. That is, one side had to win and the other lose. These contests soon led to intense conflict between the Eagles and Rattlers.

Sherif then set out to reduce the hostility that had arisen between them. He tested the efficacy of cooperative contact in the pursuit of superordinate goals—goals unattainable by one group alone. The two groups engaged in such joint experiences as fixing the camp's broken water supply, working to pay for a motion picture, and pulling a rope to start up a truck for a picnic. In time, Sherif succeeded in sharply reducing the intergroup conflict and prejudice between the Eagles and the Rattlers.

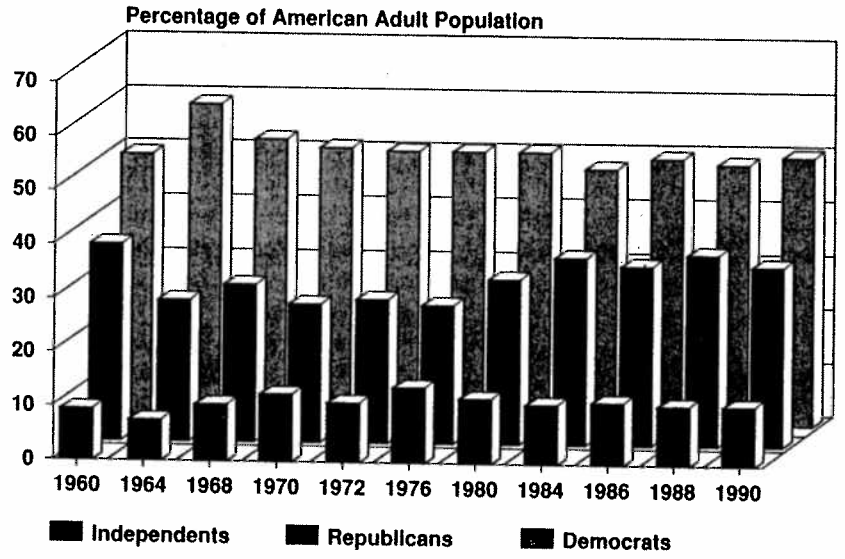
However, these cooperative efforts considered singly did not achieve dramatic change. It was a gradual, additive effect. The attainment of the first superordinate goal, the plumbing repairs, aroused "good spirits" at the time, but "the groups fell back on their old recriminations once the immediate crisis was over" (Sherif, 1966:89). A cross-sectional study would have ended with this one measurement of the dependent variable, and incorrectly concluded that cooperative experiences only temporarily improve intergroup conflict (Pettigrew, 1991c). Sherif persisted, however, with further cooperative experiences. With each new situation, the combined effects added up until little intergroup animosity remained. Only longitudinal research could have uncovered this phenomenon.

4.4.3. Trend Analyses

Political scientists and economists often study long-term societal trends in their primary variables. This work is non-experimental and makes use of various quasi-experimental designs discussed in Chapter 3. This longitudinal perspective focuses attention on sharp deviations that occur from these trends. This non-experimental approach, then, uses past data as control comparisons for interpreting present data.

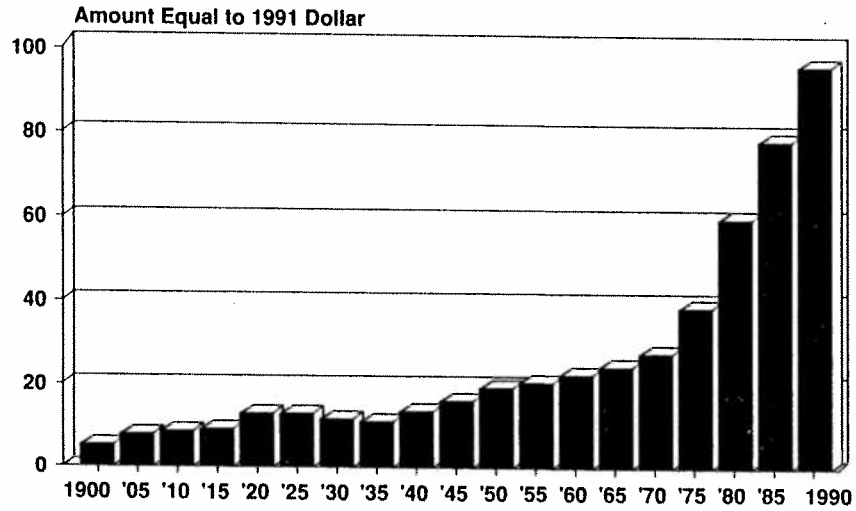
Some important indicators are surprisingly stable over time. Figure 4.3 provides an example of importance to students of American politics. Over the 30-year period, 1960–1990, political party identifications given by adult Americans to survey interviewers remained remarkably constant (U.S. Bureau

Figure 4.3 Party Trends
Political Party Identification, 1960–90



Source: U.S. Bureau of the Census, 1992a.

Figure 4.4 Consumer Price Index
Inflation Trend for All Items, 1900–1991



Source: U.S. Bureau of the Census, 1975, 1992a.
Note: All amounts in 1991 dollars.

of the Census, 1992a:270). In 1960, half of the respondents called themselves Democrats of varying strength, two-fifths Republicans, and the remainder Independents. Three decades later, 51% called themselves Democrats, 37% Republicans, and 11% Independents.

Other indicators show considerable variation. Figure 4.4 graphs the cost of living in the United States for all items across the 20th century (U.S. Bureau of the Census, 1975:210; 1992a:469). The benchmark comparison is the U.S. dollar in 1991. We read from Figure 4.4 that what cost one dollar in 1991 cost only 60 cents in 1980, and 28 cents in 1970. These data give a broad picture, though there are technical problems in combining data from the beginning and close of the century. In addition, by using composite averages for all items, Figure 4.4 obscures differences in the trends for various expenditures. Since 1960, for example, housing and medical costs have climbed more rapidly than telephone and transportation costs (U.S. Bureau of the Census, 1992a:469).

While such crude data cannot prove causation, a glimpse at this trend suggests hypotheses for further test. Note the decade of the Depression experienced a drop in prices (1930–1940). Marked increments occurred during decades of major wars (1910–1920, 1940–1950, and 1960–1970). The steep rise in prices during the 1980s suggests additional factors are also important. The 1980s experienced no major war, but it was a time of massive deficit spending on armaments. For those Americans who think their dollar does not buy what it once did, Figure 4.4 shows how right they are.

Observe also that this one indicator of inflation does not chart the economic standard of living. The inflation index improves at a time of economic

disaster (the 1930s) and rises during periods of economic prosperity. Obviously, broader indices of economic prosperity need to consider how much money people have to spend, as well as the costs of goods and services.

4.4.4. A Longitudinal Study Using Archival Data

Elder's (1974) sociological study of the effects of the Depression illustrates the use of archival data in longitudinal research. Archival data are records and materials gathered earlier and maintained for later research use. Elder did **secondary analyses** of data from a longitudinal study that earlier researchers had collected for other purposes. Such analyses are "secondary," because he reanalyzed the old data with fresh new hypotheses and methods.

The sweep of Elder's investigation provides a rare perspective on issues normally studied over short periods. With data on 167 families followed from 1932 to 1964, he looked at the effects of unemployment and economic deprivation on family structure and personality. The Depression led to economic loss that generated severe social strains and altered family life. Economic deprivation caused changes in marital relations, parent-child relations, and the division of labor within the family.

For example, the need for girls to take responsibility for home duties and for boys to earn outside income reinforced traditional attitudes about gender roles. Taking early jobs for the boys meant "an accelerated movement toward the adult world." At the personality level, these changes led to a lasting need for stability and security among these Americans who were children in the 1930s. Elder inferred that economic deprivation as children was the cause of these effects. The fact that the effects were greatest for those families whose income loss during the Depression was especially profound strengthens Elder's inference.

With these varied examples, we see the advantages and importance of longitudinal research designs. Now we must review the problems of interpretation that accompany these designs.

4.5 INTERPRETING LONGITUDINAL EFFECTS

Elder's research introduces us to the types of interrelated effects that occur over time. The Depression study revealed **cohort** effects. That is, a particular cluster of Americans who experienced the Depression as children show the effects. That is what a cohort is: *an age cluster of people who experience a significant event at about the same time in their lives* (Ryder, 1968).

One theorist, Karl Mannheim (1952), held that the political events of early life decisively shape every cohort's political orientation. Thus, the events that occur as it reaches voting age (the Depression, a war, the Civil Rights Movement, the Clinton era) have a special influence on each cohort's political perspective throughout life. Americans who entered the electorate during the 1930s provide an example. Their gratitude for President Roosevelt's efforts against the Depression led them to be unusually loyal to the Democratic Par-

in their voting ever since. So cohort is a useful concept in many fields. We will illustrate its use further in Chapter 5.

Causes of long-term changes are often difficult to untangle. Three overlapping possibilities arise. In addition to cohort effects, there are also **generation** and **life-cycle** effects. The three processes are intertwined, since each relates to age.

Generation is the broadest term. In popular usage generation includes all three effects, yet it has proven helpful in social science to distinguish between them. Social scientists limit generation effects to *the structural and cultural aspects of family lineage and parent-child relationships.* They commonly employ it, for instance, in referring to immigration patterns. They call those American citizens born elsewhere first-generation Americans. The children of immigrants, born in the United States, are second-generation Americans.

Often social scientists use generation in explanations that involve parents passing on cultural values to their children. Hence, Flacks (1967) explained the young political activists of the 1960s as a generation effect. He found these activists typically came from families with liberal, humanitarian values. Their demands for social change, Flacks argued, represented their application of these values to social injustices in American society.

If you are familiar with William Shakespeare's *As You Like It,* you already know the basic idea behind the life-cycle process. Shakespeare provides this most famous of life-cycle theories in his soliloquy about "All the world's a stage, and all the men and women merely players" (Act II, Scene VII). "And one man in his time plays many parts. . . ": the infant, "the whining schoolboy," the lover, the soldier, the middle-aged professional ("the justice"), late-middle age ("with spectacles on nose and pouch on side"), and old age (in "second childishness and mere oblivion, sans teeth, sans eyes, sans taste, sans everything").

Erik Erikson (1963, 1968) advanced today's most popular outline of life stages. He lists eight stages, each with its own "psychosocial crisis": infancy, early childhood, play age, school age, adolescence, young adulthood, adulthood, and old age. Erikson (1968:286) emphasizes two sides of the life cycle. One refers to how an individual's life "round[s] itself out as a coherent experience." The other side is how the life cycle forms "a link in the chain of generations." The life-cycle process, then, is *the course of human life moving through a sequence of developmental stages from birth to death.* All the social sciences make use of the life-cycle idea, but it is a particular specialty of developmental psychology.

To apply these three types of effects, consider a finding from political science. Jennings and Niemi (1975) studied a large national sample of Americans as they developed politically after their high school graduation in 1965. Their longitudinal research followed the respondents for eight years, from the age of 18 until they were 26. Their findings differ from those of the Bennington College women. These young adults changed their political attitudes extensively over this period back toward those of their parents' generation.

All three age effects offer possible explanations for this finding (Smelser & Smelser, 1981). Did this group shift because of their special experience with a particular historical event—such as the return to power of the Republican Party in 1968 under Richard Nixon? This would be a cohort effect. Did they shift back to their parents' more conservative views because their youthful rebelliousness receded? This would be a generation effect. Did they shift because of their passage from adolescence to young adulthood? This would be a life-cycle effect. Most likely, it was a combination of these effects. Obviously, longitudinal data are necessary to grapple with these questions—even then the interpretation of the results can be difficult.

Cohort, generational, and life-cycle effects all involve changes across age groupings. A related set of effects involve the breadth of social changes taking place across a society. Do major shifts in American public opinion, for example, reflect only the replacement of older cohorts by younger cohorts? Or do these shifts represent attitude changes across all cohorts?

These questions involve two separate processes. The first is a demographic cohort replacement process. Younger cohorts reach maturity and replace older cohorts that are dying. Due to educational advances during the 20th century, these younger cohorts are better educated than the older ones they replace. Thus, for this and other reasons, the replacement process has provided a slow but continuous liberalizing trend in American public opinion.

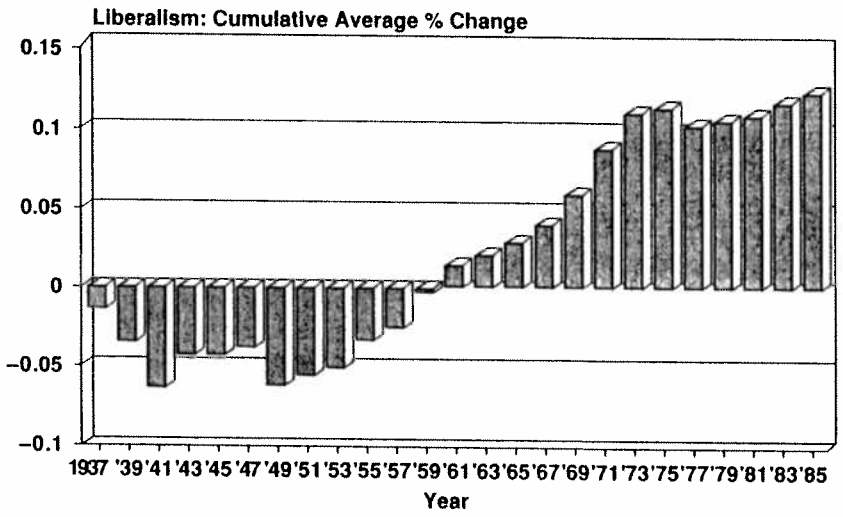
By contrast, the conversion process represents a shift in the cultural norms across all cohorts in varying degrees. Unlike the replacement process, conversion changes are sensitive to immediate events and thus can be rapid and volatile. Davis (1975, 1980, 1992) offers a metaphor using climate and the weather. Climate patterns, like the cohort replacement process, are comparatively stable and change slowly. However, the weather, like the conversion process, can change suddenly.

Sociologists and political scientists extensively analyze shifts in American social attitudes (Chafetz & Ebaugh, 1983; Davis, 1975, 1980, 1992; Kiecolt, 1988; Smith, 1982, 1990). These trend analyses use massive amounts of national survey data and are complex. They would not be possible without the computing and statistical advances mentioned in Chapter 3. There is general agreement in this work on the broad trends of American opinion since World War II. Interestingly, these conclusions differ from popular interpretations by the media of these trends.

As Figure 4.5 (adapted from Smith, 1990) illustrates, there was a marked shift in American public opinion toward liberal beliefs from 1953 until 1973. Though this period covers the Civil Rights era of the 1960s, the shift toward more liberal social attitudes is broader. The trend appears on a range of topics—from abortion to foreign affairs. There are, however, differences among topics. Leading the way were changes toward more liberal positions on individual choice, the treatment of minorities, and civil liberties in general.

Most of the steep liberal climb shown in Figure 4.5 represents a conversion effect. That is, most cohorts and social locations (sex, social class, educational and occupational subtypes) reveal these shifts in varying degrees. This finding suggests a sharp alteration of American cultural norms during this two-decade period. Yet this is not the whole story. As the weather was growing more liberal, so was the climate. Underlying the normative changes was a

Figure 4.5 Social Attitude Trend
Social Attitude Changes, 1937–1985



Source: Smith (1990).
Note: Positive = Liberal; Negative = Conservative.

smaller but continuous cohort replacement effect. Young cohorts educated after World War II were steadily replacing old cohorts educated at the turn of the century. Thus, a large conversion effect, bolstered by a smaller replacement effect, accounts for the liberalizing trend through this period.

Figure 4.5 also shows that this trend stopped in the mid-1970s. Overall, it did not become a general swing toward conservative positions, for the flat plateau in Figure 4.5 masks different trends in different opinion areas. Attitudes on race, free speech, and gender roles continued to become more liberal, though at a slower rate than before. Two areas, however, reveal conservative trends—opinions about taxation, and crime and the treatment of criminals.

Most of this change in public opinion reflects a definite shift in the conversion effect. Influenced by much publicized rises in crime, drug use, and taxation, cultural norms on these topics moved to the political right. However, the cohort replacement process was also involved. While it still produced a small net shift in the liberal direction, its size markedly decreased from previous years.

Finally, in results only partially shown in Figure 4.5, a rebound back to the liberalizing trend occurred in the mid-1980s (Davis, 1992). This well-documented phenomenon was neither expected nor detected by the mass media in the midst of the Reagan presidency. It represented a shift again in the conversion process, for it occurred among a wide variety of Americans. The cohort replacement effect, however, continued to get smaller and did not account for this liberal shift. Using his metaphor, Davis (1992) called this complicated picture "changeable weather in a cooling climate."

One final point about the replacement effect concerns its value in prediction. Since it is both stable and gradual, one can project its effects into the future with greater confidence than the cross-cutting currents of the conversion effect. Moreover, these effects can be large when there is a marked difference on the dependent variable between the newest and the oldest cohorts.

Consider again the data on political party identification shown in Figure 4.3. Recall that these data revealed overall stability in the percentages identifying with each major party. Additional data not shown in Figure 4.3 exhibit a future shift based on cohort replacement (U.S. Bureau of the Census, 1992a:270). In 1990, 43% of respondents born before 1911 (thus at least 80 years old) identified as "strong" Democrats, another 23% leaned toward the Democrats, only 5% called themselves Independents, and 30% Republicans. This is the cohort who were young adults during the Depression. As predicted by Mannheim, their strong loyalty to the Democratic Party that came from that special experience never waned. The youngest cohort interviewed in 1990 (born after 1958) looks very different. Only 11% regard themselves as "strong" Democrats, 35% leaned toward the Democrats, while 15% are Independents and 39% are Republicans. Not only will cohort replacement cost the Democratic Party support, it will especially cost them ardent followers.

These broad processes—cohort, generation, life-cycle, replacement, and conversion—aid in interpreting longitudinal data. Yet they only suggest, rather than specify precisely, the basic causes of the changes uncovered in longitudinal research. To know that a liberal rebound during the 1980s was a conversion effect suggests that a normative shift occurred in America. That leaves open the causal question of *why* did norms change toward liberalism right when an extremely conservative presidential administration governed the country. Similarly, the decline in the liberalizing cohort replacement effect suggests that those born after 1950 are not much more liberal than the old cohorts they are replacing. Why? These processes narrow and set the causal questions. We need theory to guide us in interpreting these interesting trends in American public opinion.

4.6 SUMMING UP

A primary goal of social science is to measure and explain social change. To achieve this goal, the field must grapple with the complexities of causation.

Assigning cause to events around us is usually natural, immediate, and necessary. To be sure, human perception of causation is biased. We see people as salient, causal agents (dispositional attributions), and neglect the power of situations to influence social behavior (situational attributions). We also typically give those we love the causal benefit of the doubt, while attributing the worst to those we do not like (the ultimate attribution bias).

Establishing cause in social science, however, is not so simple. Indeed,

cause becomes one of the field's most difficult tasks. We have reviewed some complications that underlie the idea of causation. Considering the messy nature of the social world, philosophy's strictest criteria for cause are not possible for social science to meet. For instance, perfect correlations between cause and effect are not realistic in social research.

Thus, a compromise on causation is necessary. This compromise balances the philosophical criteria for identifying causation with the realities of conducting social research. Recall the six interrelated contentions. (1) You cannot directly observe cause; you must always infer it. (2) Causes must precede effects. Widely different intervals between cause and effect can occur, as can reciprocal causation. (3) Since you can never ascertain "truth," you must approach it by winnowing out plausible rival hypotheses. (4) All observations reflect in part the theories and methods that lead to them. Still, stubborn facts exist that replicate widely and allow testing of causal theories. (5) Social life lacks absolutes, so we must take a probability approach to causation. Finally, (6) social science prefers manipulable causal factors, because they are useful in practical applications. These contentions provide further support for the need of a healthy skepticism in social science.

Two contrasting strategies in social science carry out this causal compromise. One strategy believes research and theory should develop in building-block fashion. This strategy maintains the same concepts and measures across studies, so we can more easily trace differences in results to their cause.

The bold approach adopts a different strategy. It uses a broad range of measures and research designs to cast as wide a net as possible. In this way, it tries to avoid overlooking new and rival explanations. This strategy is especially useful for meta-analysis—the procedures for statistically summarizing the results of many studies on a specific topic.

There has been considerable progress in social science over the 20th century in both research sophistication and theoretical rigor. Developments discussed in Chapter 1, such as the surge of computing and statistical capability, fueled this progress. We noted in this chapter the use of statistics in assessing change. Both effect sizes and significance tests effectively assess the importance and reliability of social research results. Replication of research, using new subjects, offers another important tool. It increases the power of statistical methods, and grounds important results more solidly before we apply them to practical problems.

Longitudinal studies are especially important for determining cause. Such research, measuring the dependent variable at two or more times, tracks how social phenomena unfold over time. We reviewed examples of longitudinal research of many types—laboratory and field experiments, trend analyses, and studies using archival data. In each case, we noted how the over-time design uncovered results that cross-sectional research would have missed.

Causal interpretations of longitudinal results are often difficult. For this task, it is useful to distinguish three processes related to age: cohort, generation, and life-cycle effects. A cohort is an age cluster of people who experience a significant event at about the same time in their lives. A generation refers to the structural and cultural aspects of family lineage and parent-child relationships. Finally, the life-cycle process involves the course of human life moving through a sequence of developmental stages from birth to death. While intertwined in practice, these effects have distinctive characteristics that aid analysis of longitudinal data.

Another set of processes gauges the breadth of social change. The cohort replacement process involves younger cohorts replacing older cohorts in society. It introduces slow, continuous change when the young differ from the old in key characteristics other than age.

The rival process entails participation by most cohorts and social locations in social change. This conversion process can be swift and abrupt in its effects. The two processes do not always point in the same direction. In fact, we noted the two form a varied pattern in describing the liberal shift in American social attitudes since World War II. Helpful as these processes are for guiding the interpretation of longitudinal results, more precise causal specification requires theoretical guidance.

ISSUES FOR DISCUSSION AND REVIEW

A> Newcomb's study of Bennington College students is a classic example of longitudinal research. Suppose a social scientist presently conducted a study such as Newcomb's on your campus. What do you think the results would be? Do you think the social attitudes of students at your college change over their years of attendance? If so, how and why? If not, why not? How is campus life at your college different from that of Bennington College's in the 1930s?

B> What makes "cause" such a difficult idea? Are the complications raised by philosophers directly relevant for social science? Why?

C> Choose a cross-sectional study that you know. Suppose it had been a longitudinal study. How might the findings and interpretations have differed with the dependent variable measured more than once? What would be some advantages of the longitudinal study? What are some problems of interpretation introduced by the longitudinal research design?

D> Try to think of further examples of the three principal types of over-time age effects—cohort, generational, and life cycle.

RECOMMENDATIONS FOR FURTHER READING ON ISSUES RAISED IN THIS CHAPTER

On The Logic of Social Research:

For readers new to the subject:

C. M. Judd, E. R. Smith & L. H. Kidder. 1991. *Research Methods in Social Relations.* New York: Holt, Rinehart & Winston. Chapters 3 and 4.

L. A. Davis. 1985. *The Logic of Causal Order.* Beverly Hills, CA: S

On Karl Popper's Philosophy of Science:

For readers new to the subject:
D. Miller. (Ed.) 1985. *Popper Selections.* Princeton, NJ: Princeton University Press.
For readers with a background in the philosophy of science:
K. R. Popper. 1959. *The Logic of Scientific Discovery.* New York: Basic Books.

On Meta-Analysis to Summarize Research Results:

For readers new to the technique:
C. M. Judd, E. R. Smith & L. H. Kidder. 1991. *Research Methods in Social Relations.* New York: Holt, Rinehart & Winston. Chapter 18.
For readers with statistical and methodological training:
R. Rosenthal. 1991. *Meta-Analytic Procedures for Social Research.* Rev. Ed. Newbury Park, CA: Sage.

Chapter 5

Sampling, Selecting, and Socializing

For many years, the *New York Daily News* conducted large pre-election polls with thousands of participants. The results regularly covered the paper's front pages. The problem, however, was that these polls did not allow for differences between those who responded and the voters on election day. Everyone was welcome to join the fun and list the candidates they favored. Thus, the *Daily News* did not conduct a survey with a **probability sample** that carefully mirrored the voters.

The paper's editors argued that their rough-and-ready polls compensated for their lack of sampling sophistication with their massive size. In truth, these polls were many times larger than any probability survey of New York voters that any agency could afford. Sometimes the *Daily News* polls correctly picked the election winners. Sometimes they failed miserably. Surveys using far smaller probability samples of respondents did much better.

Why? What was wrong with these enormous *Daily News* polls? Just what are probability samples, and why do they perform better? These are the questions we will consider in this chapter. Then we will show that sampling is a general social science issue, not one confined to election surveys. In fact, concern for sampling issues is an integral part of how social scientists think.

5.1 WHY ELECTION POLLS GO WRONG

The "everybody is welcome" technique of the *New York Daily News* results in massive constant errors. That is, such polls attract only particular types of people. These types are different in many ways from the total voting population. The target population to which you wish to generalize the results of a