CHAPTER 1

The *Science* in Social Science

1.1 Introduction

This book is about research in the social sciences. Our goal is practical: designing research that will produce valid inferences about social and political life. We focus on political science, but our argument applies to other disciplines such as sociology, anthropology, history, economics, and psychology and to nondisciplinary areas of study such as legal evidence, education research, and clinical reasoning.

This is neither a work in the philosophy of the social sciences nor a guide to specific research tasks such as the design of surveys, conduct of field work, or analysis of statistical data. Rather, this is a book about research design: how to pose questions and fashion scholarly research to make valid descriptive and causal inferences. As such, it occupies a middle ground between abstract philosophical debates and the hands-on techniques of the researcher and focuses on the essential logic underlying all social scientific research.

1.1.1 Two Styles of Research, One Logic of Inference

Our main goal is to connect the traditions of what are conventionally denoted "quantitative" and "qualitative" research by applying a unified logic of inference to both. The two traditions appear quite different; indeed they sometimes seem to be at war. Our view is that these differences are mainly ones of style and specific technique. The same underlying logic provides the framework for each research approach. This logic tends to be explicated and formalized clearly in discussions of quantitative research methods. But the same logic of inference underlies the best qualitative research, and all qualitative and quantitative researchers would benefit by more explicit attention to this logic in the course of designing research.

The *styles* of quantitative and qualitative research are very different. Quantitative research uses numbers and statistical methods. It tends to be based on numerical measurements of specific aspects of phenomena; it abstracts from particular instances to seek general description or to test causal hypotheses; it seeks measurements and analyses that are easily replicable by other researchers.

Qualitative research, in contrast, covers a wide range of approaches, but by definition, none of these approaches relies on numerical measurements. Such work has tended to focus on one or a small number of cases, to use intensive interviews or depth analysis of historical materials, to be discursive in method, and to be concerned with a rounded or comprehensive account of some event or unit. Even though they have a small number of cases, qualitative researchers generally unearth enormous amounts of information from their studies. Sometimes this kind of work in the social sciences is linked with area or case studies where the focus is on a particular event, decision, institution, location, issue, or piece of legislation. As is also the case with quantitative research, the instance is often important in its own right: a major change in a nation, an election, a major decision, or a world crisis. Why did the East German regime collapse so suddenly in 1989? More generally, why did almost all the communist regimes of Eastern Europe collapse in 1989? Sometimes, but certainly not always, the event may be chosen as an exemplar of a particular type of event, such as a political revolution or the decision of a particular community to reject a waste disposal site. Sometimes this kind of work is linked to area studies where the focus is on the history and culture of a particular part of the world. The particular place or event is analyzed closely and in full detail.

For several decades, political scientists have debated the merits of case studies versus statistical studies, area studies versus comparative studies, and “scientific” studies of politics using quantitative methods versus “historical” investigations relying on rich textual and contextual understanding. Some quantitative researchers believe that systematic statistical analysis is the only road to truth in the social sciences. Advocates of qualitative research vehemently disagree. This difference of opinion leads to lively debate; but unfortunately, it also bifurcates the social sciences into a quantitative-systematic-generalizing branch and a qualitative-humanistic-discursive branch. As the former becomes more and more sophisticated in the analysis of statistical data (and their work becomes less comprehensible to those who have not studied the techniques), the latter becomes more and more convinced of the irrelevance of such analyses to the seemingly nonreplicable and nongeneralizable events in which its practitioners are interested.

A major purpose of this book is to show that the differences between the quantitative and qualitative traditions are only stylistic and are methodologically and substantively unimportant. All good research can be understood—indeed, is best understood—to derive from the same underlying logic of inference. Both quantitative and qualitative research can be systematic and scientific. Historical research can be analytical, seeking to evaluate alternative explanations through a process of valid causal inference. History, or historical sociology, is not incompatible with social science (Skocpol 1984: 374–86).

Breaking down these barriers requires that we begin by questioning the very concept of "qualitative" research. We have used the term in our title to signal our subject matter, not to imply that "qualitative" research is fundamentally different from "quantitative" research, except in style.

Most research does not fit clearly into one category or the other. The best often combines features of each. In the same research project, some data may be collected that is amenable to statistical analysis, while other equally significant information is not. Patterns and trends in social, political, or economic behavior are more readily subjected to quantitative analysis than is the flow of ideas among people or the difference made by exceptional individual leadership. If we are to understand the rapidly changing social world, we will need to include information that cannot be easily quantified as well as that which can. Furthermore, all social science requires comparison, which entails judgments of which phenomena are "more" or "less" alike in degree (i.e., quantitative differences) or in kind (i.e., qualitative differences).

Two excellent recent studies exemplify this point. In *Coercive Cooperation* (1992), Lisa L. Martin sought to explain the degree of international cooperation on economic sanctions by quantitatively analyzing ninety-nine cases of attempted economic sanctions from the post–World War II era. Although this quantitative analysis yielded much valuable information, certain causal inferences suggested by the data were ambiguous; hence, Martin carried out six detailed case studies of sanctions episodes in an attempt to gather more evidence relevant to her causal inference. For *Making Democracy Work* (1993), Robert D. Putnam and his colleagues interviewed 112 Italian regional councillors in 1970, 194 in 1976, and 234 in 1981–1982, and 115 community leaders in 1976 and 118 in 1981–1982. They also sent a mail questionnaire to over 500 community leaders throughout the country in 1983. Four nationwide mass surveys were undertaken especially for this study. Nevertheless, between 1976 and 1989 Putnam and his colleagues conducted detailed case studies of the politics of six regions. Seeking to satisfy the "interocular traumatic test," the investigators "gained an intimate knowledge of the internal political maneuvering and personalities that have animated regional politics over the last two decades" (Putnam 1993:190).

The lessons of these efforts should be clear: neither quantitative nor qualitative research is superior to the other, regardless of the research problem being addressed. Since many subjects of interest to social scientists cannot be meaningfully formulated in ways that permit statistical testing of hypotheses with quantitative data, we do not wish to encourage the exclusive use of quantitative techniques. We are not trying to get all social scientists out of the library and into the computer center, or to replace idiosyncratic conversations with structured interviews. Rather, we argue that nonstatistical research will produce more reliable results if researchers pay attention to the rules of scientific inference—rules that are sometimes more clearly stated in the style of quantitative research. Precisely defined statistical methods that undergird quantitative research represent abstract formal models applicable to all kinds of research, even that for which variables cannot be measured quantitatively. The very abstract, and even unrealistic, nature of statistical models is what makes the rules of inference shine through so clearly.

The rules of inference that we discuss are not relevant to all issues that are of significance to social scientists. Many of the most important questions concerning political life—about such concepts as agency, obligation, legitimacy, citizenship, sovereignty, and the proper relationship between national societies and international politics—are philosophical rather than empirical. But the rules are relevant to all research where the goal is to learn facts about the real world. Indeed, the distinctive characteristic that sets social science apart from casual observation is that social science seeks to arrive at valid inferences by the systematic use of well-established procedures of inquiry. Our focus here on empirical research means that we sidestep many issues in the philosophy of social science as well as controversies about the role of postmodernism, the nature and existence of truth, relativism, and related subjects. We assume that it is possible to have some knowledge of the external world but that such knowledge is always uncertain.

Furthermore, nothing in our set of rules implies that we must run the perfect experiment (if such a thing existed) or collect all relevant data before we can make valid social scientific inferences. An important topic is worth studying even if very little information is available. The result of applying any research design in this situation will be relatively uncertain conclusions, but so long as we honestly report our uncertainty, this kind of study can be very useful. Limited information is often a necessary feature of social inquiry. Because the social world changes rapidly, analyses that help us understand those changes require that we describe them and seek to understand them contemporaneously, even when uncertainty about our conclusions is high. The urgency of a problem may be so great that data gathered by the most useful scientific methods might be obsolete before it can be accumulated. If a distraught person is running at us swinging an ax, administering a five-page questionnaire on psychopathy may not be the best strategy. Joseph Schumpeter once cited Albert Einstein, who said "as far as our propositions are certain, they do not say anything about reality, and as far as they do say anything about reality, they are not certain" (Schumpeter [1936] 1991:298–99). Yet even though certainty is unattainable, we can improve the reliability, validity, certainty, and honesty of our conclusions by paying attention to the rules of scientific inference. The social science we espouse seeks to make descriptive and causal inferences about the world. Those who do not share the assumptions of partial and imperfect knowability and the aspiration for descriptive and causal understanding will have to look elsewhere for inspiration or for paradigmatic battles in which to engage.

In sum, we do not provide recipes for scientific empirical research. We offer a number of precepts and rules, but these are meant to discipline thought, not stifle it. In both quantitative and qualitative research, we engage in the imperfect application of theoretical standards of inference to inherently imperfect research designs and empirical data. Any meaningful rules admit of exceptions, but we can ask that exceptions be justified explicitly, that their implications for the reliability of research be assessed, and that the uncertainty of conclusions be reported. We seek not dogma, but disciplined thought.

1.1.2 Defining Scientific Research in the Social Sciences

Our definition of "scientific research" is an ideal to which any actual quantitative or qualitative research, even the most careful, is only an approximation. Yet, we need a definition of good research, for which we use the word "scientific" as our descriptor.[1] This word comes with many connotations that are unwarranted or inappropriate or downright incendiary for some qualitative researchers. Hence, we provide an explicit definition here. As should be clear, we do not regard quantitative research to be any more scientific than qualitative research. Good research, that is, scientific research, can be quantitative or qualitative in style. In design, however, scientific research has the following four characteristics:

1. **The goal is inference.** Scientific research is designed to make descriptive or explanatory *inferences* on the basis of empirical information about the world. Careful descriptions of specific phenomena are often indispens-

[1] We reject the concept, or at least the word, "quasi-experiment." Either a research design involves investigator control over the observations and values of the key causal variables (in which case it is an experiment) or it does not (in which case it is nonexperimental research). Both experimental and nonexperimental research have their advantages and drawbacks; one is not better in all research situations than the other.

able to scientific research, but the accumulation of facts alone is not sufficient. Facts can be collected (by qualitative or quantitative researchers) more or less systematically, and the former is obviously better than the latter, but our particular definition of science requires the additional step of attempting to infer beyond the immediate data to something broader that is not directly observed. That something may involve *descriptive inference*—using observations from the world to learn about other unobserved facts. Or that something may involve *causal inference*—learning about causal effects from the data observed. The domain of inference can be restricted in space and time—voting behavior in American elections since 1960, social movements in Eastern Europe since 1989—or it can be extensive—human behavior since the invention of agriculture. In either case, the key distinguishing mark of scientific research is the goal of making inferences that go beyond the particular observations collected.

2. **The procedures are public.** Scientific research uses explicit, codified, and *public* methods to generate and analyze data whose reliability can therefore be assessed. Much social research in the qualitative style follows fewer precise rules of research procedure or of inference. As Robert K. Merton ([1949] 1968:71–72) put it, "The sociological analysis of qualitative data often resides in a private world of penetrating but unfathomable insights and ineffable understandings. . . . [However,] science . . . is public, not private." Merton's statement is not true of all qualitative researchers (and it is unfortunately still true of some quantitative analysts), but many proceed as if they had no method—sometimes as if the use of explicit methods would diminish their creativity. Nevertheless they cannot help but use some method. Somehow they observe phenomena, ask questions, infer information about the world from these observations, and make inferences about cause and effect. If the method and logic of a researcher's observations and inferences are left implicit, the scholarly community has no way of judging the validity of what was done. We cannot evaluate the principles of selection that were used to record observations, the ways in which observations were processed, and the logic by which conclusions were drawn. We cannot learn from their methods or replicate their results. Such research is not a *public* act. Whether or not it makes good reading, it is not a contribution to social science.

   All methods—whether explicit or not—have limitations. The advantage of explicitness is that those limitations can be understood and, if possible, addressed. In addition, the methods can be taught and shared. This process allows research results to be compared across separate researchers and research projects studies to be replicated, and scholars to learn.

3. **The conclusions are uncertain.** By definition, inference is an imperfect process. Its goal is to use quantitative or qualitative data to learn about the world that produced them. Reaching perfectly certain conclusions

from uncertain data is obviously impossible. Indeed, uncertainty is a central aspect of all research and all knowledge about the world. Without a reasonable estimate of uncertainty, a description of the real world or an inference about a causal effect in the real world is uninterpretable. A researcher who fails to face the issue of uncertainty directly is either asserting that he or she knows everything perfectly or that he or she has no idea how certain or uncertain the results are. Either way, inferences without uncertainty estimates are not science as we define it.

4. **The content is the method.** Finally, scientific research adheres to a set of rules of inference on which its validity depends. Explicating the most important rules is a major task of this book.[2] The content of "science" is primarily the methods and rules, not the subject matter, since we can use these methods to study virtually anything. This point was recognized over a century ago when Karl Pearson (1892: 16) explained that "the field of science is unlimited; its material is endless; every group of natural phenomena, every phase of social life, every stage of past or present development is material for science. The unity of all science consists alone in its method, not in its material."

These four features of science have a further implication: science at its best is a *social enterprise*. Every researcher or team of researchers labors under limitations of knowledge and insight, and mistakes are unavoidable, yet such errors will likely be pointed out by others. Understanding the social character of science can be liberating since it means that our work need not to be beyond criticism to make an important contribution—whether to the description of a problem or its conceptualization, to theory or to the evaluation of theory. As long as our work explicitly addresses (or attempts to redirect) the concerns of the community of scholars and uses public methods to arrive at inferences that are consistent with rules of science and the information at our disposal, it is likely to make a contribution. And the contribution of even a minor article is greater than that of the "great work" that stays forever in a desk drawer or within the confines of a computer.

1.1.3 Science and Complexity

Social science constitutes an attempt to make sense of social situations that we perceive as more or less complex. We need to recognize, however, that what we perceive as complexity is not entirely inherent in phenomena: the world is not naturally divided into simple and com-

[2] Although we do cover the vast majority of the important rules of scientific inference, they are not complete. Indeed, most philosophers agree that a complete, exhaustive inductive logic is impossible, even in principle.

plex sets of events. On the contrary, the perceived complexity of a situation depends in part on how well we can simplify reality, and our capacity to simplify depends on whether we can specify outcomes and explanatory variables in a coherent way. Having more observations may assist us in this process but is usually insufficient. Thus *"complexity" is partly conditional on the state of our theory*.

Scientific methods can be as valuable for intrinsically complex events as for simpler ones. Complexity is likely to make our inferences less certain but should *not* make them any less scientific. Uncertainty and limited data should not cause us to abandon scientific research. On the contrary: the biggest payoff for using the rules of scientific inference occurs precisely when data are limited, observation tools are flawed, measurements are unclear, and relationships are uncertain. With clear relationships and unambiguous data, method may be less important, since even partially flawed rules of inference may produce answers that are roughly correct.

Consider some complex, and in some sense unique, events with enormous ramifications. The collapse of the Roman Empire, the French Revolution, the American Civil War, World War I, the Holocaust, and the reunification of Germany in 1990 are all examples of such events. These events seem to be the result of complex interactions of many forces whose conjuncture appears crucial to the event having taken place. That is, independently caused sequences of events and forces converged at a given place and time, their interaction appearing to bring about the events being observed (Hirschman 1970). Furthermore, it is often difficult to believe that these events were inevitable products of large-scale historical forces: some seem to have depended, in part, on idiosyncracies of personalities, institutions, or social movements. Indeed, from the perspective of our theories, chance often seems to have played a role: factors outside the scope of the theory provided crucial links in the sequences of events.

One way to understand such events is by seeking generalizations: conceptualizing each case as a member of a *class of events* about which meaningful generalizations can be made. This method often works well for ordinary wars or revolutions, but some wars and revolutions, being much more extreme than others, are "outliers" in the statistical distribution. Furthermore, notable early wars or revolutions may exert such a strong impact on subsequent events of the same class—we think again of the French Revolution—that caution is necessary in comparing them with their successors, which may be to some extent the product of imitation. Expanding the class of events can be useful, but it is not always appropriate.

Another way of dealing scientifically with rare, large-scale events is to engage in counterfactual analysis: "the mental construction of a course of events which is altered through modifications in one or more 'conditions'" (Weber [1905] 1949:173). The application of this idea in a systematic, scientific way is illustrated in a particularly extreme example of a rare event from geology and evolutionary biology, both historically oriented natural sciences. Stephen J. Gould has suggested that one way to distinguish systematic features of evolution from stochastic, chance events may be to imagine what the world would be like if all conditions up to a specific point were fixed and then the rest of history were rerun. He contends that if it were possible to "replay the tape of life," to let evolution occur again from the beginning, the world's organisms today would be a completely different (Gould 1989a).

A unique event on which students of evolution have recently focused is the sudden extinction of the dinosaurs 65 million years ago. Gould (1989a:318) says, "we must assume that consciousness would not have evolved on our planet if a cosmic catastrophe had not claimed the dinosaurs as victims." If this statement is true, the extinction of the dinosaurs was as important as any historical event for human beings; however, dinosaur extinction does not fall neatly into a class of events that could be studied in a systematic, comparative fashion through the application of general laws in a straightforward way.

Nevertheless, dinosaur extinction can be studied scientifically: alternative hypotheses can be developed and tested with respect to their observable implications. One hypothesis to account for dinosaur extinction, developed by Luis Alvarez and collaborators at Berkeley in the late 1970s (W. Alvarez and Asaro, 1990), posits a cosmic collision: a meteorite crashed into the earth at about 72,000 kilometers an hour, creating a blast greater than that from a full-scale nuclear war. If this hypothesis is correct, it would have the observable implication that iridium (an element common in meteorites but rare on earth) should be found in the particular layer of the earth's crust that corresponds to sediment laid down sixty-five million years ago; indeed, the discovery of iridium at predicted layers in the earth has been taken as partial confirming evidence for the theory. Although this is an unambiguously unique event, there are many other observable implications. For one example, it should be possible to find the metorite's crater somewhere on Earth (and several candidates have already been found).[3]

The issue of the cause(s) of dinosaur extinction remains unresolved, although the controversy has generated much valuable research. For

[3] However, an alternative hypothesis, that extinction was caused by volcanic eruptions, is also consistent with the presence of iridium, and seems more consistent than the meteorite hypothesis with the finding that all the species extinctions did not occur simultaneously.

our purposes, the point of this example is that scientific generalizations are useful in studying even highly unusual events that do not fall into a large class of events. The Alvarez hypothesis cannot be tested with reference to a set of common events, but it does have observable implications for other phenomena that can be evaluated. We should note, however, that a hypothesis is not considered a reasonably certain explanation until it has been evaluated empirically and passed a number of demanding tests. At a minimum, its implications must be consistent with our knowledge of the external world; at best, it should predict what Imre Lakatos (1970) refers to as "new facts," that is, those formerly unobserved.

The point is that even apparently unique events such as dinosaur extinction can be studied scientifically if we pay attention to improving theory, data, and our use of the data. Improving our theory through conceptual clarification and specification of variables can generate more observable implications and even test causal theories of unique events such as dinosaur extinction. Improving our data allows us to observe more of these observable implications, and improving our use of data permits more of these implications to be extracted from existing data. That a set of events to be studied is highly complex does not render careful research design irrelevant. Whether we study many phenomena or few—or even one—the study will be improved if we collect data on as many observable implications of our theory as possible.

1.2 Major Components of Research Design

Social science research at its best is a creative process of insight and discovery taking place within a well-established structure of scientific inquiry. The first-rate social scientist does not regard a research design as a blueprint for a mechanical process of data-gathering and evaluation. To the contrary, the scholar must have the flexibility of mind to overturn old ways of looking at the world, to ask new questions, to revise research designs appropriately, and then to collect more data of a different type than originally intended. However, if the researcher's findings are to be valid and accepted by scholars in this field, all these revisions and reconsiderations must take place according to explicit procedures consistent with the rules of inference. A dynamic process of inquiry occurs within a stable structure of rules.

Social scientists often begin research with a considered design, collect some data, and draw conclusions. But this process is rarely a smooth one and is not always best done in this order: conclusions rarely follow easily from a research design and data collected in accordance with it. Once an investigator has collected data as provided by a research design, he or she will often find an imperfect fit among the main research questions, the theory and the data at hand. At this stage, researchers often become discouraged. They mistakenly believe that other social scientists find close, immediate fits between data and research. This perception is due to the fact that investigators often take down the scaffolding after putting up their intellectual buildings, leaving little trace of the agony and uncertainty of construction. Thus the process of inquiry seems more mechanical and cut-and-dried than it actually is.

Some of our advice is directed toward researchers who are trying to make connections between theory and data. At times, they can design more appropriate data-collection procedures in order to evaluate a theory better; at other times, they can use the data they have and recast a theoretical question (or even pose an entirely different question that was not originally foreseen) to produce a more important research project. The research, if it adheres to rules of inference, will still be scientific and produce reliable inferences about the world.

Wherever possible, researchers should also improve their research designs before conducting any field research. However, data has a way of disciplining thought. It is extremely common to find that the best research design falls apart when the very first observations are collected—it is not that the theory is wrong but that the data are not suited to answering the questions originally posed. Understanding from the outset what can and what cannot be done at this later stage can help the researcher anticipate at least some of the problems when first designing the research.

For analytical purposes, we divide all research designs into four components: the *research question*, the *theory*, the *data*, and the *use of the data*. These components are not usually developed separately and scholars do not attend to them in any preordained order. In fact, for qualitative researchers who begin their field work before choosing a precise research question, data comes first, followed by the others. However, this particular breakdown, which we explain in sections 1.2.1–1.2.4, is particularly useful for understanding the nature of research designs. In order to clarify precisely what *could* be done if resources were redirected, our advice in the remainder of this section assumes that researchers have unlimited time and resources. Of course, in any actual research situation, one must always make compromises. We believe that understanding the advice in the four categories that follow will help researchers make these compromises in such a way as to improve their research designs most, even when in fact their research is subject to external constraints.