# EXPLAINING SOCIAL BEHAVIOR

aca

More Nuts and Bolts for the Social Sciences

Jon Elster collège de france





### Introduction

& & &

This book is about explaining social behavior. In the first part, I spell out my conception of explanation, and in the remaining four parts, I construct a toolbox of concepts and mechanisms that apply to particular cases. Needless to say, it does not aspire to completeness. Rather than trying to spell out the gaps, which will be obvious, let me begin by enumerating a sample of the puzzles that, I submit, can be illuminated by the approach I am taking. In the Conclusion, I return to the same puzzles with brief references to the explanations I have cited in earlier chapters.

The examples and the explanations must be taken with two caveats. First, I do not claim that all the explananda are well-established facts. In an actual explanation, this is of course a crucial first step – it makes no sense to try to explain what does not exist. For the purpose of building a toolbox, however, one can be less rigorous. Second, even for the explananda whose existence is well documented I do not claim that the explanations I cite are the correct ones. I only claim that they satisfy a minimal condition for an explanation – that they logically imply the explananda. The puzzles and explanations are intended to show "if this kind of thing happens, here is the kind of mechanism that might explain it" as well as "if this mechanism operates, here is the kind of thing it can produce." Given these caveats, here are the puzzles, arranged somewhat arbitrarily (since many puzzles could fit in several categories) according to the four substantive parts of the book.<sup>I</sup>

#### I∽The Mind

• Why do some gamblers believe than when red has come up five times in a row, red is more likely than black to come up next?

<sup>&</sup>lt;sup>1</sup> Although the list overlaps somewhat with a list of puzzles presented in Chapter 12 as challenges to rational-choice theory, it has no polemical purpose, only that of inciting the reader's curiosity.



#### 2

#### INTRODUCTION

- Why do other gamblers believe than when red has come up five times in a row, black is more likely than red to come up next?
- Why do preferences sometimes change through the sheer passage of time?
- Why do many people who seem to believe in the afterlife want it to arrive as late as possible?
- Why are people reluctant to acknowledge, to themselves and others, that they are envious?
- Why are people reluctant to acknowledge, to themselves and others, that they are ignorant?
- Why, among sixteenth-century converts to Calvinism, did the belief that people were predestined either to heaven or to hell induce greater peace of mind than the belief that one could achieve salvation through good works?
- Why is it (sometimes) true that "Who has offended, cannot forgive"?
- Why is shame more important than guilt in some cultures?
- Why did the French victory in the 1998 soccer World Cup generate so much joy in the country, and why did the fact that the French team did not qualify beyond the opening rounds in 2002 cause so much despondency?
- Why do women often feel shame after being raped?
- Why do humiliating rituals of initiation produce greater rather than lesser loyalty to the group into which one is initiated?

#### II Action

- Why do more Broadway shows receive standing ovations today than twenty years ago?
- Why may punishments increase rather than decrease the frequency of the behavior they target?



#### INTRODUCTION

3

- Why are people unwilling to break self-imposed rules even when it makes little sense to follow them?
- Why is the pattern of revenge "Two eyes for an eye" instead of "An eye for an eye"?
- Why is the long-term yield on stocks much larger than that on bonds (i.e., why does not the value of stocks rise to equalize the yields)?
- Why do suicide rates go down when dangerous medications are sold in blister packs rather than bottles?
- Why did none of thirty-eight bystanders call the police when Kitty Genovese was beaten to death?
- Why did some individuals hide or rescue Jews under the Nazi regimes?
- Why did President Chirac call early elections in 1997, only to lose his majority in parliament?
- Why are some divorcing parents willing to share child custody even when their preferred solution is sole custody, which they are likely to get were they to litigate?
- Why are poor people less likely to emigrate?
- Why do some people save in Christmas accounts that pay no interest and do not allow for withdrawal before Christmas?
- Why do people pursue projects, such as building the Concorde airplane, that have negative expected value?
- Why, in "transitional justice" (when agents of an autocratic regime are put on trial after the transition to democracy), are those tried immediately after the transition sentenced more severely than those who are tried later?
- Why, in Shakespeare's play, does Hamlet delay taking revenge until the last act?



4

#### INTRODUCTION

#### III Lessons from the Natural Sciences

- Why are parents much more likely to kill adopted children and stepchildren than to kill their biological children?
- Why is sibling incest so rare, given the temptations and opportunities?
- Why do people invest their money in projects undertaken by other agents even when the latter are free to keep all the profits for themselves?
- Why do people take revenge at some material cost to them and with no material benefits?
- Why do people jump to conclusions beyond what is warranted by the evidence?

#### IV Interaction

- Why do supporters of a Socialist party sometimes vote Communist and thereby prevent their party from winning?
- Why do some newly independent countries adopt as their official language that of their former imperialist oppressor?
- Why are ice cream stalls often located beside each other in the middle of the beach, when customers would be better off and the sellers no worse off with a more spread-out location?
- Why does an individual vote in elections when his or her vote is virtually certain to have no effect on the outcome?
- Why are economically successful individuals in modern Western societies usually slimmer than the average person?
- Why do people refrain from transactions that could make everybody better off, as when they abstain from asking a person in the front of a bus queue whether he is willing to sell his place?



#### INTRODUCTION

5

- Why did President Nixon try to present himself to the Soviets as being prone to irrational behavior?
- Why do military commanders sometimes burn their bridges (or their ships)?
- Why do people often attach great importance to intrinsically insignificant matters of etiquette?
- Why do passengers tip taxi drivers and customers tip waiters even when visiting a foreign city to which they do not expect to return?
- Why do firms invest in large inventories even when they do not anticipate any interruption of production?
- Why, in a group of students, would each think that others have understood an obscure text better than he has?
- Why are votes in many political assemblies taken by roll call?
- Why is logrolling more frequent in ordinary legislatures than in constituent assemblies?

Suggested explanations for these phenomena will be provided at various places in the book and briefly summarized in the Conclusion. Here I only want to make a general remark about two types of explanation that are *not* likely to be useful. As readers will see in the very first chapter, with several reminders along the road, one of the aims of the book is to inculcate skepticism toward two common lines of reasoning. First, with very few exceptions the social sciences cannot rely on functional explanation, which accounts for actions or behavioral patterns by citing their consequences rather than their causes. Do norms of tipping exist because it is more efficient to have customers monitor waiters than to have the owner do it? I do not think so. Second, I now believe that rational-choice theory has less explanatory power than I used to think. Do real people act on the calculations that make up many pages of mathematical appendixes in leading journals? I do not think so.

On three counts at least, rational-choice theory is nevertheless a valuable part of the toolbox. If understood in a qualitative commonsense



6 INTRODUCTION

way, it is capable of explaining much everyday behavior. Even when it does not explain much, it can have immense conceptual value. Game theory, in particular, has illuminated the structure of social interaction in ways that go far beyond the insights achieved in earlier centuries. Finally, human beings *want* to be rational. The desire to have sufficient reasons for one's behavior, and not simply be the plaything of psychic forces acting "behind one's back," provides a permanent counterforce to the many irrationality-generating mechanisms that I survey in this book.

Even though I am critical of many rational-choice explanations, I believe the concept of *choice* is fundamental. In the book I consider several alternatives to choice-based explanation and conclude that although they may sometimes usefully supplement that approach, they cannot replace it. The fact that people act under different *constraints*, for instance, can often explain a great deal of variation in behavior. Also, in some cases one may argue that *selection of agents* rather than *choice by agents* is responsible for the behavior we observe. By and large, however, I believe that the subjective factor of choice has greater explanatory power than the objective factors of constraints and selection. This is obviously an intuition that cannot be proved in any rigorous sense, and in any case social scientists ought to have room for all the factors in their toolbox.



## I

## EXPLANATION AND MECHANISMS



This book relies on a specific view about explanation in the social sciences. Although not primarily a work of philosophy of social science, it draws upon and advocates certain methodological ideas about how to explain social phenomena. In the first three chapters, these ideas are set out explicitly. In the rest of the book they mostly form part of the implicit background, although from time to time, notably in Chapters 14 through 17 and in the Conclusion, they return to the center of the stage.

I argue that all explanation is causal. To explain a phenomenon (an explanandum) is to cite an earlier phenomenon (the explanans) that caused it. When advocating causal explanation, I do not intend to exclude the possibility of intentional explanation of behavior. Intentions can serve as causes. A particular variety of intentional explanation is rational-choice explanation, which will be extensively discussed in later chapters. Many intentional explanations, however, rest on the assumption that agents are, in one way or another, irrational. In itself, irrationality is just a negative or residual idea, everything that is not rational. For the idea to have any explanatory purchase, we need to appeal to specific forms of irrationality with specific implications for behavior. In Chapter 12, for instance, I enumerate and illustrate eleven mechanisms that can generate irrational behavior.

Sometimes, scientists explain phenomena by their *consequences* rather than by their causes. They might say, for instance, that blood feuds are explained by the fact that they keep populations down at sustainable levels. This might seem a metaphysical impossibility: how can the



#### 8 EXPLANATION AND MECHANISMS

existence or occurrence of something at one point in time be explained by something that has not yet come into existence? As we shall see, the problem can be restated so as to make explanation by consequences a meaningful concept. In the biological sciences, evolutionary explanation offers an example. In the social sciences, however, successful instances of such explanation are few and far between. The blood-feud example is definitely not one of them.

The natural sciences, especially physics and chemistry, offer *explanations by law*; laws are general propositions that allow us to infer the truth of one statement at one time from the truth of another statement at some earlier time. Thus when we know the positions and the velocity of the planets at one time, the laws of planetary motion enable us to deduce and predict their positions at any later time. This kind of explanation is *deterministic*: given the antecedents, only one consequent is possible. The social sciences offer few if any law-like explanations of this kind. The relation between explanans and explanandum is not one-one or manyone, but one-many or many-many. Many social scientists try to model this relation by using *statistical* methods. Statistical explanations are incomplete by themselves, however, since they ultimately have to rely on intuitions about plausible causal *mechanisms*.



## EXPLANATION

& & &

#### Explanation: General

The main task of the social sciences is to explain social phenomena. It is not the only task, but it is the most important one, to which others are subordinated or on which they depend. The basic type of explanandum is an *event*. To explain it is to give an account of why it happened, by citing an *earlier event* as its cause. Thus we may explain Ronald Reagan's victory in the 1980 presidential elections by Jimmy Carter's failed attempt to rescue the Americans held hostage in Iran. Or we might explain the outbreak of World War II by citing any number of earlier events, from the Munich agreement to the signing of the Versailles Treaty. Even though in both cases the fine structure of the causal explanation will obviously be more complex, they do embody the basic *event-event* pattern of explanation. In a tradition originating with David Hume, it is often referred to as the "billiard-ball" model of causal explanation. One event, ball A hitting ball B, is the cause of – and thus explains – another event, namely, ball B's beginning to move.

Those who are familiar with the typical kind of explanation in the social sciences may not recognize this pattern, or not see it as privileged. In one way or another, social scientists tend to put more emphasis on *facts*, or states of affairs, than on events. The sentence "At 9 A.M. the road was slippery" states a fact. The sentence "At 9 A.M. the car went off the road" states an event. As this example suggests, one might offer a *factevent* explanation to account for a car accident. Conversely, one might propose an *event-fact* explanation to account for a given state of affairs, as when asserting that the attack on the World Trade Center in 2001

<sup>&</sup>lt;sup>1</sup> To anticipate a distinction discussed later, note that, Carter did not *fail to attempt* but *attempted and failed*. A nonaction such as a failure to attempt cannot have causal efficacy, except in the indirect sense that if others perceive or infer that the agent fails to act, they may take actions that they otherwise would not have.

<sup>&</sup>lt;sup>2</sup> The voter turnout example discussed later provides another illustration.



IO

Cambridge University Press
978-0-521-77744-5 - Explaining Social Behavior: More Nuts and Bolts for the
Social Sciences
Jon Elster
Excerpt
More information

EXPLANATION AND MECHANISMS

explains the pervasive state of fear of many Americans. Finally, standard social-science explanations often have a *fact-fact* pattern. To take an example at random, it has been claimed that the level of education of women explains per capita income in the developing world.

Let us consider the explanation of one particular fact, that 65 percent of Americans favor, or say that they favor, the death penalty.<sup>3</sup> In principle, this issue can be restated in terms of events: How did these Americans *come to favor* the death penalty? What were the formative events – interactions with parents, peers, or teachers – that caused this attitude to emerge? In practice, social scientists are usually not interested in this question. Rather than trying to explain a brute statistic of this kind, they want to understand *changes* in attitudes over time or *differences* in attitudes across populations. The reason, perhaps, is that they do not think the brute fact very informative. If one asks whether 65 percent is much or little, the obvious retort is, "Compared to what?" Compared to the attitudes of Americans around 1990, when about 80 percent favored the death penalty, it is a low number. Compared to the attitudes in some European countries, it is a high number.

Longitudinal studies consider variations over time in the dependent variable. Cross-sectional studies consider variations across populations. In either case, the explanandum is transformed. Rather than trying to explain the phenomenon "in and of itself," we try to explain how it varies in time or space. The success of an explanation is measured, in part, by how much of the "variance" (a technical measure of variation) it can account for. Complete success would explain all observed variation. In a cross-national study we might find, for instance, that the percentage of individuals favoring the death penalty was strictly proportional to the number of homicides per 100,000 inhabitants. Although this finding would provide no explanation of the absolute numbers, it would offer a perfect explanation of the difference among them. In practice, of course,

© Cambridge University Press

<sup>&</sup>lt;sup>3</sup> Answers fluctuate. Also, the number of people who favor the death penalty for murder goes down drastically when life imprisonment without parole is stated as the alternative.

<sup>&</sup>lt;sup>4</sup> As economists sometimes say, they are interested only in what happens "at the margin."

<sup>&</sup>lt;sup>5</sup> Strictly speaking, the causal chain might go in the other direction, from attitudes to behavior, but in this case that hypothesis is implausible.

perfect success is never achieved, but the same point holds. Explanations of variance do not say anything about the explanandum "in and of itself."

An example may be taken from the study of voting behavior. As we shall see later (Chapter 12), it is not clear why voters bother to vote at all in national elections, when it is morally certain that a single vote will make no difference. Yet a substantial fraction of the electorate do turn out on voting day. Why do they bother?

Instead of trying to solve this mystery, empirical social scientists usually address a different question: Why does turnout vary across elections? One hypothesis is that voters are less likely to turn out in inclement weather, because rain or cold makes it more attractive to stay home. If the data match this hypothesis, as indicated by line C in Figure 1.1, one might claim to have explained (at least part of) the variance in turnout. Yet one would not have offered any explanation of why the line C intersects the vertical axis at P rather than at Q or R. It is as if one took the first decimal as given and focused on explaining the second. For predictive purposes, this might be all one needs. For explanatory purposes, it is unsatisfactory. The "brute event" that 45

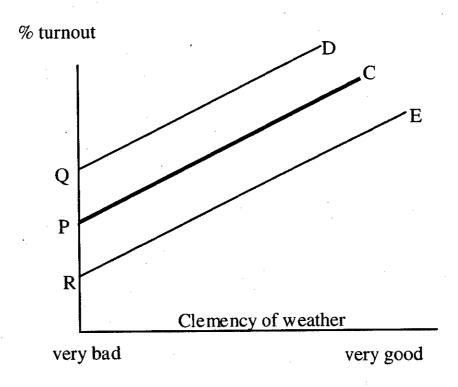


FIGURE I.I

percent or more of the electorate usually turn out to vote is an interesting one, which cries out for an explanation.

The ideal procedure, in an event-event perspective, would be the following. Consider two elections, A and B. For each of them, identify the events that cause a given percentage of voters to turn out. Once we have thus explained the turnout in election A and the turnout in election B, the explanation of the difference (if any) follows automatically, as a by-product. As a bonus, we might also be able to explain whether identical turnouts in A and B are accidental, that is, due to differences that exactly offset each other, or not. In practice, this procedure might be too demanding. The data or the available theories might not allow us to explain the phenomena "in and of themselves." We should be aware, however, that if we do resort to explanations of variance, we are engaging in a second-best explanatory practice.

Sometimes, social scientists try to explain nonevents. Why do many people fail to claim social benefits they are entitled to? Why did nobody call the police in the Kitty Genovese case? Considering the first question, the explanation might be that the individuals in question decide not to claim their benefits, because of fear of stigma or concerns with self-image. Since making a decision is an event, this would provide a fully satisfactory account. If it fails, social scientists would, once again, look at the differences between those who are entitled to benefits and claim them and those who are and do not. Suppose the only difference is that the latter are unaware of their entitlement. As an explanation, this is helpful but insufficient. To go beyond it, we would want to explain why some entitled individuals are unaware of their entitlement. To discover that because they are illiterate, they are unable to read the letters informing them about their rights would also be helpful but insufficient. At some point in the explanatory regress, we must either come to a positive event, such as a conscious decision not to become literate or a conscious

<sup>&</sup>lt;sup>6</sup> For more than half an hour on March 27, 1964, thirty-eight respectable, law-abiding citizens in Queens, New York, watched a killer stalk and stab a woman in three separate attacks in Kew Gardens. Twice their chatter and the sudden glow of their bedroom lights interrupted him and frightened him off. Each time he returned, sought her out, and stabbed her again. Not one person telephoned the police during the assault; one witness called after the woman was dead.

decision by officials to withhold information, or turn to those who do seek the benefits to which they are entitled. Once we have explained the behavior of the latter, the explanation why others fail to seek their benefit will emerge as a by-product.

Considering the Kitty Genovese case, there is no variation in behavior to explain, since nobody called the police. Accounts of the case indicate that several of the observers decided not to call the police. In terms of proximate causes this provides a fully satisfactory account, although we might want to know the reasons for their decision. Was it because they feared "getting involved" or because each observer assumed that someone else would call the police ("Too many shepherds make a poor guard")? Some of the observers, however, apparently did not even think about calling the police. One man and his wife watched the episode for its entertainment value, while another man said he was tired and went to bed. To explain why they did not react more strongly one might cite their shallow emotions, but that, too, would be to account for a negative explanandum by citing a negative explanans. Once again, their behavior can only be explained as a by-product or residual. If we have a satisfactory explanation of why some individuals thought about calling the police, even if in the end they decided not to, we shall have the only explanation we are likely to get of why some did not even think about it.

In the rest of this book I shall often relax this purist or rigorist approach of what counts as a relevant explanandum and an appropriate explanation. The insistence on event-focused explanations is a bit like the principle of methodological individualism, which is another premise of the book. In principle, explanations in the social sciences should refer only to individuals and their actions. In practice, social scientists often refer to supraindividual entities such as families, firms, or nations, either as a harmless shorthand or as a second-best approach forced upon them by lack of data or of fine-grained theories. These two justifications also apply to the use of facts as explananda or as explanantia, to explanations of variance rather than of the phenomena "in and of themselves," and to the analysis of negative explananda (nonevents or nonfacts). The purpose of the preceding discussion is not to hold social scientists to pointless or impossible standards, but to argue that at the level of first principles the

event-based approach is intrinsically superior. If scholars keep that fact in mind they may, at least sometimes, come up with better and more fruitful explanations.

Sometimes, we might want to explain an event (or rather a pattern of events) by its consequences rather than by its causes. I do not have in mind explanation by intended consequences, since intentions exist prior to the choices or actions they explain. Rather, the idea is that events may be explained by their actual consequences, typically, their beneficial consequences for someone or something. As a cause must precede its effect, this idea might seem to be incompatible with causal explanation. Yet causal explanation can also take the form of explanation by consequences, if there is a loop from the consequences back to their causes. A child may initially cry simply because it feels pain, but if the crying also gets it attention from the parents, it may start crying more than it would have otherwise. I argue in Chapters 16 and 17 that this kind of explanation is somewhat marginal in the study of human behavior. In most of the book, I shall be concerned with the simple variety of causal explanation in which the explanans - which might include beliefs and intentions oriented toward the future - precedes the occurrence of the explanandum.<sup>7</sup>

In addition to the fully respectable form of functional explanation that rests on specific feedback mechanisms, there are more disreputable forms that simply point to the production of consequences that are beneficial in some respect and then without further argument assume that these suffice to explain the behavior that causes them. When the explanandum is a token, such as a single action or event, this kind of explanation fails for purely metaphysical reasons. To take an example from biology, we cannot explain the occurrence of a neutral or harmful mutation by observing that it was a necessary condition for a further, advantageous one. When the explanandum is a type, such as a recurrent pattern of behavior, it may or may not be valid. Yet as long as it is not supported by

<sup>&</sup>lt;sup>7</sup> For some purposes, it may be useful to distinguish among causal, intentional, and functional explanation. Physics employs only causal explanation; biology additionally admits functional explanation; and the social sciences further admit intentional explanation. At the most fundamental level, though, all explanation is causal.

a specific feedback mechanism, we should treat it as if it were invalid. Anthropologists have argued, for instance, that revenge behavior has beneficial consequences of various kinds, ranging from population control to decentralized norm enforcement. (Chapter 22 offers many other examples.) Assuming that these benefits are in fact produced, they might still obtain by accident. To show that they arise nonaccidentally, that is, that they sustain the revenge behavior that causes them, the demonstration of a feedback mechanism is indispensable. And even when one is provided, the initial occurrence of the explanandum must be due to something else.

## The Structure of Explanations

Let me now turn to a more detailed account of explanation in the social sciences (and, to some extent, more generally). The first step is easily overlooked: before we try to explain a fact or an event we have to establish that the fact is a fact or that the event actually did take place. As Montaigne wrote, "I realize that if you ask people to account for 'facts,' they usually spend more time finding reasons for them than finding out whether they are true. . . . They skip over the facts but carefully deduce inferences. They normally begin thus: 'How does this come about?' But does it do so? That is what they ought to be asking."

Thus before trying to explain, say, why there are more suicides in one country than in another, we have to make sure that the latter does not tend, perhaps for religious reasons, to underreport suicides. Before we try to explain why Spain has a higher unemployment rate than France, we have to make sure that the reported differences are not due to different definitions of unemployment or to the presence of a large underground economy in Spain. If we want to explain why youth unemployment is higher in France than in the United Kingdom, we need to decide whether the explanandum is the rate of unemployment among young people who are actively searching for jobs or the rate among young people overall, including students. If we compare unemployment in Europe and the United States, we have to decide whether the explanandum is the unemployed in the literal sense, which includes the

incarcerated population, or in the technical sense, which only includes those searching for work. Before we try to explain why revenge takes the form of "tit for tat" (I kill one of yours each time you or yours kill one of mine), we should verify that this is actually what we observe rather than, say, "two tits for a tat" (I kill two of yours each time you or yours kill one of mine). Much of science, including social science, tries to explain things we all know, but science can also make a contribution by establishing that some of the things we all think we know simply are not so. In that case, social science may also try to explain *why* we think we know things that are not so, adding as it were a piece of knowledge to replace the one that has been taken away. 9

Suppose now that we have a well-established explanandum for which there is no well-established explanation – a puzzle. The puzzle may be a surprising or counterintuitive fact, or simply an unexplained correlation. One small-scale example is "Why are more theology books stolen from Oxford libraries than books on other subjects?" Another small-scale example, which I shall explore in more detail shortly, is "Why do more Broadway shows receive standing ovations today than twenty years ago?"

Ideally, explanatory puzzles should be addressed in the five-step sequence spelled out in the following. In practice, however, steps (1), (2), and (3) often occur in a different order. We may play around with different hypotheses until one of them emerges as the most promising, and then look around for a theory that would justify it. If steps (4) and (5) are carried out properly, we may still have a high level of confidence in the preferred hypothesis. Yet for reasons I discuss toward the end of the next chapter, scholars might want to limit their freedom to pick and choose among hypotheses.

<sup>&</sup>lt;sup>8</sup> In either of the last two cases, some individuals may take up a career as criminals or students because they do not think they would get a job if they tried. For some purposes, one might want to count these among the unemployed; for other purposes, not.

<sup>&</sup>lt;sup>9</sup> Just as science can help explain popular beliefs in nonfacts, it can help explain popular beliefs in false explanations. For instance, most of those who suffer from arthritis believe arthritic pain is triggered by bad weather. Studies suggest, however, that there is no such connection. Perhaps we should drop the search for the causal link between bad weather and arthritic pain and instead try to explain why arthritics believe there is one. Most likely they were once told there was a connection and subsequently paid more attention to instances that confirmed the belief than to those that did not.

- 1. Choose the theory a set of interrelated causal propositions that holds out the greatest promise of a successful explanation.
- 2. Specify a hypothesis that applies the theory to the puzzle, in the sense that the explanandum follows logically from the hypothesis.
- 3. Identify or imagine plausible accounts that might provide alternative explanations, also in the sense that the explanandum follows logically from each of them.
- 4. For each of these rival accounts, refute it by pointing to additional testable implications that are in fact *not* observed.
- 5. Strengthen the proposed hypothesis by showing that it has additional testable implications, preferably of "novel facts," that are in fact observed.

These procedures define what is often called the *hypothetico-deductive* method. In a given case, they might take the form shown in Figure 1.2. I shall illustrate it by the puzzle of increasing frequency of standing ovations on Broadway. It is not based on systematic observations or controlled experiments, but on my casual impressions confirmed by newspaper reports. For the present purposes, however, the shaky status of the explanandum does not matter. If there are in fact more standing ovations on Broadway than there were twenty years ago, how could we go about explaining it?

I shall consider an explanation in terms of the rising prices of Broadway tickets. One newspaper reports the playwright Arthur Miller as saying, "I guess the audience just feels having paid \$75 to sit down, it's their time to stand up. I don't mean to be a cynic but it probably all changed when the price went up." When people have to pay seventy-five dollars or more for a seat, many cannot admit to themselves that the show was poor or mediocre, and that they have wasted their money. To confirm to themselves that they had a good time, they applaud wildly.

More formally, the explanation is sought in the hypothesis "When people have paid a great deal of money or effort to obtain a good, they tend (other things being equal) to value it more highly than when they paid less for it." Given the factual premise of rising prices, this

A similar idea is sometimes used to defend the high fees of psychotherapists: patients wouldn't believe in the therapy unless they paid a lot for it. But no therapists to my knowledge state that they donate 50 percent of their fee to Red Cross.

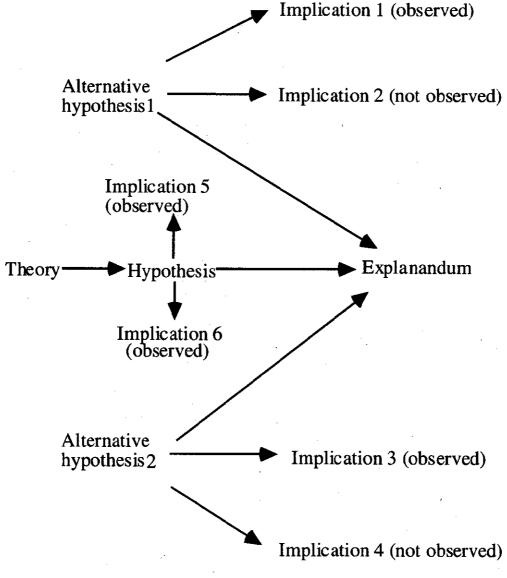


FIGURE 1.2

proposition passes the minimal test that any explanatory hypothesis must satisfy: If it is true, we can infer the explanandum. But this is a truly minimal test, which many propositions could pass. To strengthen our belief in this particular explanation, we must show that it is supported from below, from above, and laterally.

The human mind seems to have a tendency to turn this minimal requirement into a sufficient one. Once we have hit upon an account that may be true, we often do not pause to test it further or to consider alternative accounts. The choice of an account may be due to the idea of post hoc ergo propter hoc (after it, therefore because of it), or to an inference from the fact that a given account is more plausible than others to the conclusion that it is more likely than not to be correct.

An explanation is supported *from below* if we can deduce and verify observable facts from the hypothesis over and above the fact that the hypothesis is intended to explain. It must have "excess explanatory power." In the case of the Broadway shows, we would expect fewer standing ovations in shows whose prices for some reason have not gone up.<sup>12</sup> Also, we would expect fewer standing ovations if large numbers of tickets to a show are sold to firms and given by them to their employees. (This would count as a "novel fact.") Even if these tickets are expensive, the spectators have not paid for them out of their own pocket and hence do not need to tell themselves that they are getting their money's worth.

An explanation is supported *from above* if the explanatory hypothesis can be deduced from a more general theory.<sup>13</sup> In the present case, the explanatory proposition is a specification of the theory of cognitive dissonance proposed by Leon Festinger. The theory says that when a person experiences an internal inconsistency or dissonance among her beliefs and values, we can expect some kind of mental readjustment that will eliminate or reduce the dissonance. Typically, the adjustment will choose the path of least resistance. A person who has spent seventy-five dollars to see a show that turns out to be bad cannot easily make herself believe that she paid less than that amount. It is easier to persuade herself that the show was in fact quite good.

Although not without problems, the theory of cognitive dissonance is pretty well supported. Some of the support is from cases that are very different from the one we are considering here, as when a person who has just bought a car avidly seeks out ads for that very brand of car, to bolster his conviction that he made a good decision. Some of the support arises from quite similar cases, as when the painful and humiliating initiation

We would *not* necessarily expect fewer people to rise to their feet in the cheaper sections. They might feel foolish sitting when others are rising; also, they might have to get up to see the actors who would otherwise be blocked from view by those standing in front of them.

<sup>&</sup>lt;sup>13</sup> More accurately: if it is a *specification* of a more general theory. The relation between a general theory and a specific explanatory hypothesis is rarely a deductive one. For one thing, there may be some slack in the theory itself (see Chapter 2). For another, a given theory can usually be operationalized in many different ways.

rituals of college fraternities and sororities induce strong feelings of loyalty. I am not saying that people would consciously tell themselves, "Because I suffered so much to join this group, it must be a good group to belong to." The mechanism by which the suffering induces loyalty must be an unconscious one.

An explanation receives *lateral support* if we can think of and then refute alternative explanations that also pass the minimal test. Perhaps there are more standing ovations because today's audiences, arriving in busloads from New Jersey, are less sophisticated than the traditional audience of blasé New York denizens. Or perhaps it is because shows are better than they used to be. For each of these alternatives, we must think of and then disconfirm additional facts that would obtain if they were correct. If standing ovations are more frequent because audiences are more impressionable, we would expect them also to have been frequent in out-of-town performances twenty years ago. If shows are better than they used to be, we would expect this to be reflected in how well they are reviewed and how long they play before folding.

In this procedure, the advocate for the original hypotheses also has to be the devil's advocate. One has consistently to think against oneself - to make matters as difficult for oneself as one can. We should select the strongest and most plausible alternative rival explanations, rather than accounts that can easily be refuted. For similar reasons, when seeking to demonstrate the excess explanatory power of the hypothesis, we should try to deduce and confirm implications that are novel, counterintuitive, and as different from the original explanandum as possible. These two criteria - refuting the most plausible alternatives and generating novel facts - are decisive for the credibility of an explanation. Support from above helps but can never be decisive. In the long run it is the theory that is supported by the successful explanations it generates, not the other way around. Emilio Segrè, a Nobel Prize winner in physics, said that some winners confer honor on the prize whereas others derive honor from it. The latter are, however, parasitic on the former. Similarly, a theory is parasitic on the number of successful explanations it generates. If it is able to confer support on a given explanation, it is only because it has received support from earlier explanations.