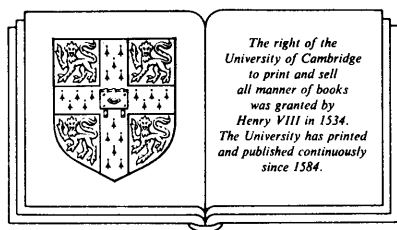


Nuts and Bolts for the Social Sciences



JON ELSTER



*The right of the
University of Cambridge
to print and sell
all manner of books
was granted by
Henry VIII in 1534.
The University has printed
and published continuously
since 1584.*

CAMBRIDGE UNIVERSITY PRESS

Cambridge

New York Port Chester Melbourne Sydney

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

© Jon Elster

First published 1989
Reprinted 1990 (twice)

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Elster, Jon, 1940–

Nuts and bolts for the social sciences / Jon Elster.

p. cm.

Bibliography: p.

ISBN 0-521-37455-3. – ISBN 0-521-37606-8 (pbk.)

1. Social sciences – Methodology. 2. Social interaction.

I. Title

H61.E434 1989

89-31449

300'.1 – dc20

CIP

British Library Cataloguing in Publication Data

Elster, Jon, 1940–

Nuts and bolts for the social sciences.

1. Social sciences. Concepts

I. Title

300

ISBN 0-521-37455-3 hardback

ISBN 0-521-37606-8 paperback

I

MECHANISMS

THE emphasis in this book is on *explanation by mechanisms*. It offers a toolbox of mechanisms – nuts and bolts, cogs and wheels – that can be used to explain quite complex social phenomena.

The social sciences, like other empirical sciences, try to explain two sorts of phenomena: events and facts. The election of George Bush as president is an event. The presence in the electorate of a majority of Republican voters is a fact, or a state of affairs. It is not immediately obvious what is more fundamental, events or facts. One might, quite plausibly, explain Bush's victory by the Republican majority. One might also, no less plausibly, explain the Republican majority as being the result of a series of events, each of which took the form of belief formation by an individual voter. The second perspective is the more fundamental: explaining events is logically prior to explaining facts. A fact is a temporal snapshot of a stream of events, or a pile of such snapshots. In the social sciences, the elementary events are individual human actions, including mental acts such as belief formation.

To explain an event is to give an account of why it happened. Usually, and always ultimately,¹ this takes the form of citing an earlier event as the cause of the event we want to explain, together with some account of the causal mechanism connecting the two events. Here is a simple, paradigmatic example. We want to know why someone changed his mind about a job he

¹ Sometimes people explain events by citing other events that occur later rather than earlier in time. When valid, such explanations ultimately conform to the main pattern. The topic is further discussed in chapters VIII and IX.

Introduction

previously held to be very desirable, but now finds utterly lacking in interest. The explanation has two elements. First, before changing his mind, he learned that he had no chances of getting the job. Second, there is a causal *mechanism*, often referred to as cognitive dissonance reduction, that makes people cease desiring what they cannot get, as in the story of the fox and the sour grapes. A more complex event might be a fall in average work tenure. The earlier event was legislation designed to enhance job security by requiring employers to give job tenure to all who had been employed for more than, say, two years. The causal mechanism is rational adaptation to the legislation by employers, who find it in their interest to dismiss workers just before the expiration of the two-year period.

Statements that purport to explain an event must be carefully distinguished from a number of other types of statement. First, causal explanations must be distinguished from true causal statements. To cite the cause is not enough: the causal mechanism must also be provided, or at least suggested. In everyday language, in most historical writings and in many social scientific analyses, the mechanism is not explicitly cited. Instead, it is suggested by the way in which the cause is described. Any given event can be described in many ways. In narrative explanations, it is tacitly presupposed that only causally relevant features of the event are used to identify it. If told that a person died as a result of having eaten rotten food, we assume that the mechanism was food poisoning. If told that he died as a result of eating food to which he was allergic, we assume that the mechanism was an allergic reaction. Suppose now that he actually died because of food poisoning, but that he was also allergic to the food in question, lobster. To say that he died because he ate food to which he had an allergy would be true, but misleading. It would suggest the wrong causal mechanism. To say that he died because he ate lobster would be true, but uninformative. It would suggest no causal mechanism at all and exclude very few. Indeed, the actual mechanism could be almost anything,

from being hit by a car to being hit by a bullet, if either of these events was triggered by the person's eating lobster.

Second, causal explanations must be distinguished from assertions about correlation. Sometimes we are in a position to say that an event of a certain type is invariably or usually followed by an event of another kind. This does not allow us to say that events of the first type cause events of the second, because there is another possibility: the two might be common effects of a third event. Consider the finding that children in contested custody cases suffer more than children whose parents have reached a private custody agreement. It could be that the custody trial itself explains the difference, by causing pain and guilt in the children. It could also be, however, that custody disputes are more likely to occur when the parents are bitterly hostile toward each other and that children of such parents tend to be more unhappy. To distinguish between the two interpretations, we would have to measure suffering before and after the divorce.

Here is a more complex example, my favorite example, in fact, of this kind of ambiguity. In *Democracy in America*, Alexis de Tocqueville discusses the alleged causal connection between marrying for love and having an unhappy marriage. He points out that this connection obtains only in societies in which such marriages are the exception and arranged marriages are the rule. Only stubborn people will go against the current and two stubborn persons are not likely to have a very happy marriage.² In addition, people who go against the current are treated badly by their more conformist peers, inducing bitterness and more unhappiness. Of these arguments, the first rests on a noncausal correlation between marrying for love and unhappiness. The second does point to a true causal connection, albeit not the one that the critics of love marriages to whom Tocqueville addressed

² Here the "third factor" is not an event, but a character trait: stubbornness. To explain the character trait, however, we would have to invoke (genetic and social) events. This illustrates the point, made earlier, that the priority of events over facts obtains ultimately, not immediately.

his argument had in mind. Marrying for love causes unhappiness only in a context where this practice is exceptional. Biologists often refer to such effects as "frequency dependent." I discuss this notion in chapter XI.

Third, causal explanations must be distinguished from assertions about necessitation. To explain an event is to give an account of why it happened *as* it happened. The fact that it might also have happened in some other way, and would have happened in some other way if it had not happened in the way it did, is neither here nor there. Consider a person who suffers from an incurable form of cancer, which is certain to kill him within one year. He is, however, killed in a car accident. To *explain* why he died within a certain time period, it is pointless to say that he *had to* die in that period because he had cancer. If all we know about the case is the onset of cancer, the limited life span of persons with that type of cancer and the death of the person, it is plausible to infer that he died because of the cancer. We have the earlier event and a causal mechanism sufficient to bring about the later event. But the mechanism is not necessary: it could be preempted by another. To find out what actually happened, we need more finely grained knowledge. The quest never ends: right up to the last second, something else could preempt the cancer. Yet the more we know, the more confident we are that we have the right explanation.³

The two problems we have just discussed add up to a weakness in the best-known theory of scientific explanation, that proposed by Carl Hempel. He argues that explanation amounts to logical deduction of the event to be explained, with general laws and statements of initial conditions as the premises. One objection is that the general laws might reflect correlation, not causation. Another is that the laws, even if genuinely causal,

³ Causal preemption should be distinguished from causal overdetermination. The latter is illustrated by a person being hit simultaneously by two bullets, each of which would have been sufficient to kill him. The former is illustrated by a person being hit by one bullet and as a result falling down, thereby avoiding being hit by another bullet, which would otherwise have killed him.

Mechanisms

might be preempted by other mechanisms. This is why I have placed the emphasis here on mechanisms, not on laws. This is not a deep philosophical disagreement. A causal mechanism has a finite number of links. Each link will have to be described by a general law, and in that sense by a "black box" about whose internal gears and wheels we remain ignorant. Yet for practical purposes – the purposes of the working social scientist – the place of emphasis is important. By concentrating on mechanisms, one captures the dynamic aspect of scientific explanation: the urge to produce explanations of ever finer grain.

Fourth, causal explanations must be distinguished from storytelling. A genuine explanation accounts for what happened, as it happened. To tell a story is to account for what happened as it might have happened (and perhaps did happen). I have just argued that genuine explanations differ from accounts of what had to happen. I am now saying that they also differ from accounts of what may have happened. The point may seem trivial, or strange. Why would anyone want to come up with a purely conjectural account of an event? Is there any place in science for speculations of this sort. The answer is yes – but their place must not be confused with that of explanations.

Storytelling can suggest new, parsimonious explanations. Suppose that someone asserts that self-sacrificing or helping behavior is conclusive proof that not all action is self-interested or that emotional behavior is conclusive proof that not all action is rational.⁴ One might conclude that there are three irreducibly different forms of behavior: rational and selfish, rational and nonselfish, and irrational. The drive for parsimony that characterizes good science should lead us to question this view.⁵ Could it not be in one's self-interest to help others? Could it not be

4 A well-known example from another domain is provided by the numerous biologists who have asserted that living organisms cannot possibly be explained by chemical and physical theories.

5 Yet the sense for realism that also characterizes good science should make us wary of the simplistic tendency to believe that all reductionist attempts will succeed.

rational to be swayed by one's emotions? The first step toward finding a positive answer is telling a *plausible story* to show how these possibilities could be realized. It could be, for instance, that people help others because they expect reciprocation or that people become angry because that helps them to get their way. By telling a story one can transform an issue from a meta-physical one into one that is amenable to empirical research. The question now is whether the premises of the story are true, not whether it is possible or impossible to explain one range of phenomena in terms of other, less complex phenomena.

At the same time, storytelling can be harmful if it is mistaken for the real thing. Much of social science is driven by the idea that "everything has a function." Even behavior that appears to be harmful and maladaptive should be shown to be useful and, moreover, be explained in terms of its usefulness. To demonstrate function and usefulness, scholars often resort to storytelling. They have a considerable number of devices at their disposal. Behavior that isn't optimal now may have been so under other circumstances in the past. Behavior that isn't optimal taken in isolation may be a necessary ingredient in an optimal package solution. What is maladaptive for the individual may be good for society. With some ingenuity – and many scholars have a great deal – one can *always* tell a story in which things are turned upside down. But that doesn't prove they really are that way, any more than Kipling's *Just So Stories* explain how the leopard got its spots or the Ethiopian his color.

Finally, causal explanations must be distinguished from predictions. Sometimes we can explain without being able to predict, and sometimes predict without being able to explain. True, in many cases one and the same theory will enable us to do both, but I believe that in the social sciences this is the exception rather than the rule.

To see why we can have explanatory power without predictive power, consider once again the reduction of cognitive dissonance. In many people, this mechanism coexists with the exactly opposite one, captured in homely sayings such as "The

Mechanisms

grass is always greener on the other side of the fence" and "Forbidden fruit tastes best." Sometimes it seems as if people want to be unhappy, by desiring objects demonstrably out of reach simply because they are out of reach. Build a fence around someone, and he immediately wants to get out, while before he had no such thought in his mind. As far as I know, we have no theories that tell us when one or the other of these mechanisms will operate. When one of them does operate, we recognize it immediately, and so we can explain the behavior it generates. But we cannot reliably predict when it will operate.

Another example will help to bring the point home. When people try to make up their mind whether to participate in a cooperative venture, such as cleaning up litter from the lawn or voting in a national election, they often look to see what others are doing. Some of them will think as follows: "If most others cooperate, I too should do my share, but if they don't I have no obligation to do so." Others will reason in exactly the opposite way: "If most others cooperate, there is no need for me to do so. If few others cooperate, my obligation to do so will be stronger." In fact, most individuals are subject to both of these psychic mechanisms, and it is hard to tell before the fact which will dominate.

It is sometimes said that the opposite of a profound truth is another profound truth.⁶ The social sciences offer a number of illustrations of this profound truth. They can isolate tendencies, propensities and mechanisms and show that they have implications for behavior that are often surprising and counterintuitive. What they are more rarely able to do is to state necessary and sufficient conditions under which the various mechanisms are switched on. This is another reason for emphasizing mechanisms rather than laws. Laws by their nature are general and do

6 "Opposite" must be taken in the sense of internal rather than external negation. The internal negation of "People prefer what they can have over what they cannot have" is "People prefer what they cannot have over what they can have." Both statements yield true and important insights. The external negation of the first statement is simply that "People do not prefer what they can have over what they cannot have," a statement that does not suggest any important insights.

Introduction

not suffer exceptions. One cannot have a law to the effect that "if p , then sometimes q ."⁷ Mechanisms, by contrast, make no claim to generality. When we have identified a mechanism whereby p leads to q , knowledge has progressed because we have added a new item to our repertoire of ways in which things happen.

Conversely, we may have predictive power without explanatory power. To predict that less of a good will be bought when its price goes up, there is no need to form a hypothesis about consumer behavior. Whatever the springs of individual action – rational, traditional or simply random – we can predict that people will buy less of the good simply because they can afford less of it. Here there are several mechanisms that are constrained to lead to the same outcome, so that for predictive purposes there is no need to decide among them. Yet for explanatory purposes the mechanism is what matters. It provides understanding, whereas prediction at most offers control.⁸

Also, for predictive purposes the distinction among correlation, necessitation and explanation becomes pointless. If there is a lawlike regularity between one type of event and another, it does not matter whether it is due to a causal relation between them or to their being common effects of a third cause. In either case we can use the occurrence of the first type of event to predict the occurrence of the second. Nobody believes that the first symptoms of a deadly disease cause the later death, yet they are regularly used to predict that event. Similarly, it does not matter for predictive purposes whether a necessitating mechanism might be preempted by another. Knowing that a person has incurable cancer allows us to predict that he will die, whether or not he in fact dies from cancer.

7 Although Sidney Morgenbesser has suggested as the "first law of Jewish logic": if p , why not q ?

8 "At most," for reasons discussed in chapter II.