Mechanistic explanation in social contexts: Elster and the problem of local scientific growth

Persson, Johannes

Published in:
Social Epistemology

DOI:
10.1080/02691728.2011.604445

2012

Citation for published version (APA):
Jon Elster worries about the explanatory power of the social sciences. His main concern is that they have so few well-established laws. Elster develops an interesting substitute: a special kind of mechanism designed to fill the explanatory gap between laws and mere description. However, his mechanisms suffer from a characteristic problem that I will explore in this article. As our causal knowledge of a specific problem grows we might come to know too much to make use of an Elsterian mechanism but still lack a law. We might then find ourselves in the paradoxical position of knowing more relevant causal truths about the phenomenon we are interested in than we did before, but being able to explain less. If this possibility is realized in social science settings, as I argue it might well be, Elster’s mechanistic account is threatened. Moreover, even if the possibility is rarely realized in that way, it raises, simply as a possibility, a conceptual problem with Elster’s mechanistic framework.

Keywords: Elster; Explanation; Mechanism; Scientific Growth

Elster’s Mechanisms

It is primarily in the absence of laws that Jon Elster finds a use for mechanisms in social science. In one sense Elster is relatively modest, in that he does not challenge the law-perspective as an explanatory ideal. His claim is rather that, this ideal notwithstanding, mechanisms occupy an available position between laws and mere
descriptions. And in any case, since relevant laws are scarce in social contexts, the ideal is seldom useful. In practice, then, Elsterian mechanisms could be doing all the explanatory work—quite an ambitious proposal.

To make Elster’s suggestion precise one needs to specify both what a law is and what a mechanism is. Both have been conceived in a variety of ways, but for the purpose of this article I will follow Elster. I shall argue later, however, that although Elster’s suggestion is promising, his conception of mechanisms is unsuitable for the task.

First, the characterization of mechanisms (Elster refers to it occasionally as a definition of mechanism):

Roughly speaking, mechanisms are frequently occurring and easily recognizable causal patterns that are triggered under generally unknown conditions or with indeterminate consequences. They allow us to explain, but not to predict. (Elster 2007, 36; original emphasis)

Elster is careful to point out that he is not assuming objective, knowledge-independent indeterminacy. He is interested merely in consequences that are indeterminate relative to what we know. This is in keeping with his general aim to characterize mechanisms in a way that makes them depend on our epistemic condition: a mechanism belongs to the (partly ignorant) social scientist’s worldview rather than the world itself. This epistemic conception of mechanisms is somewhat unusual, and indeed it deserves special scrutiny precisely because it differs from several other conceptions in this respect.1

Elster illustrates explanation by mechanism with the case of alcoholism. For every child who becomes alcoholic, as he or she grows up, in response to an alcohol-abusive environment, another avoids alcoholism in response to the same environment. Whether or not we can predict a causal pattern, a demonstration that something is an instance of a causal pattern will satisfy many explanatory criteria. Mechanisms, it seems, are explanatory in their own right.

Most importantly, Elster notices that mechanisms often come in pairs, as is the case in the alcoholism example. From his characterization of mechanisms it follows that there are two general kinds of mechanism: those of type A are triggered under unknown conditions; by contrast, type-B mechanisms have indeterminate consequences. In a situation of type A, either this or that causal chain is triggered, but we do not know which one. For instance, the prey detects the predator. Sometimes this leads to escape, but sometimes the prey ignores the predator. The type-A mechanism has to do with what happens at the beginning of the causal chain. It is exclusively about triggering. In a situation of type B, on the other hand, two chains are triggered with an indeterminate overall result. Type B has to do with what happens at the end of the evolving causal pattern. It is about the indeterminacy of net effects. For all swimmers, probably, training increases swimming skill and increases confidence in swimming skill as well. Both might affect the person’s risk exposure. In some cases the risks increase with the swimmer’s skill (Elster 2007, 39). But certainly the risk of drowning does not invariably increase with...
swimming skill. Instead, training seems to activate two causal chains: in one, skill tends to depress the risk of misdemeanour; in the other, confidence tends to increase it. The net-effect of training on risk is often indeterminate. Sometimes, at least, confidence in swimming skill increases more rapidly than the skill itself.

Elster’s Laws

Focusing on mechanisms, Elster has less to say about what kinds of law and law-governed explanation he is operating with. He characterizes explanation by law in the following way:

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. (Elster 2007, 8; original emphasis)

However, while laws are necessarily general and deterministic, qualified determinism is also acknowledged in the framework. Thus, Elster (2007, 36) discusses what he refers to as weak laws where only the direction of the change (and not the magnitude) can be predicted.

To complete this sketch of Elster’s characterization of laws we should perhaps add that he sometimes uses the expression “universal law” instead of “deterministic law” or “general law”. In many cases this notion is better suited to distinguish between mechanisms and laws than the more frequent “general law”. For, given that they are frequently occurring causal patterns, mechanisms are also general.

Among other things, this conception of laws situates Elster’s theoretical framework squarely within the traditional debate on prediction in social science. The scientific goal of prediction might be impossible to meet in many areas of social science, but Elsterian mechanisms nevertheless promise to deliver scientific understanding.

Comparing the Features of Mechanisms and Laws

In order to pinpoint the problem I am interested in, we should scrutinize the framework on offer. Only by learning just how different explanations by law and mechanisms are can we evaluate whether the latter fill the gap between mere descriptions and laws satisfactorily. The introduction to Elster’s framework I have given is admittedly brief. However, it puts us in a position to compare Elster’s mechanisms and laws in several respects.

It is clear that the primary difference between them is that the causal patterns corresponding to mechanisms have exceptions, whereas the causal patterns corresponding to laws do not. In both types of mechanism, the perceived cause occurs but the outcome varies. This partition of laws and mechanisms, together with the insight that mechanistic status is relative to what we know—that is, any
lack of determinism is the upshot of epistemic shortcomings—makes perfect sense on traditional accounts of explanation. Building our mechanistic knowledge to the point at which we can identify a general deterministic relation will entail that we have a covering-law explanation to hand.

Normally, of course, complete execution of the explanatory enterprise is not attempted within the social sciences, where it is not plausible that general, universal laws will be found to begin with. Many expressions that might, prima facie, be thought to be attempts at law formulations are therefore better thought of as elaborations of mechanism. I appreciate Elster’s insight; often enough science does indeed try to discover causal patterns that do not qualify as deterministic laws but can be understood as mechanisms (see, for instance, Mitchell 2009; Persson 2005).

However, a problem with Elster’s account is precisely that he pays excessive attention to laws when developing his concept of mechanism. This can be seen happening in two ways.

First, the limitations built into Elster’s mechanisms are clearly motivated by the distinction between law and mechanism: they help to uphold that distinction, and Elster does not explain why mechanisms have either unknown triggering conditions or indeterminate consequences in any other way. But if the scientific focus was really on mechanisms rather than laws, such limiting characteristics would be fairly uninteresting. There exists a rather rich body of mechanistic literature expanding on the nature of mechanisms nowadays, but on Elster’s view almost everything about mechanisms (except the fact that they can be superimposed, which is reflected in the two types of mechanism) is reduced to this alleged opposition between law and mechanism. I take this observation to demonstrate the intimate connection between laws and mechanisms in Elster’s framework.

Second, unless the traditional, essentially law-governed perspective on determinism is granted, there seems to be an additional problem with the proposed dichotomy between laws and mechanisms. Let us assume, for a moment, that it is an open question whether we need laws to be able to derive determinate causal consequences in a certain situation. For example, given the right kind of causal information we might be able to predict certain behavioural processes in response to pandemic flu risk communication in the United States. Success in this endeavour, however, would not imply that we can predict human behavioural processes in the general way presupposed by the covering-law model. On Elster’s account, a causal relationship with determinate consequences that is not assumed to be an instantiation of a general law will qualify neither as a mechanism nor a law.

Local Scientific Growth

It is time to substantiate the problem with Elster’s view referred to in the second point above. As already indicated, this problem has to do with what happens when we discover more about the triggering and/or net effects of a mechanism in a local context.
Here is a typical illustration of the scenario I have in mind. Although human behavioural processes are quite general, their expression clearly depends on the specific context. The difficulty in moving from general knowledge to specific predictions is often under-rated. This is one reason why people exaggerate their ability to predict other people’s behaviour. That includes, among much else, cases in which experts predict how emergency plans will work. This is in accordance with Elster’s account. The general pattern is clear enough, but the details on triggering and net effects are unknown. However, the situation of the risk managers can be improved by further research. The advice is often to make social scientists part of the planning team. Their contribution typically consists in adding crucial pieces of causal knowledge to the existing understanding of the specific situation. When this process is successful, predictions concerning the specific case become possible. “This is research that we know how to do”, says Baruch Fischhoff (2005) in a House Science Committee Briefing on gaps in the National Flu Preparedness Plan.

In accordance with Fischhoff’s promise, there should be a real possibility that the added causal knowledge—whether it concerns patterns of type A or type B—will clarify, and hence take us beyond, some of the previously unknown conditions (type A) and indeterminacies (type B) that were relied on in the initial characterization of the mechanism. (Remember that not even the indeterminacy of type-B mechanisms was supposed to be objective and knowledge-independent.)

However, if indeterminacy is resolved, it follows from Elster’s understanding of mechanisms that, instead of providing ourselves with an improved mechanistic explanation, we shall lose the mechanism in the context we are interested in. We shall no longer be dealing with situations of type A or type B. Elster partly acknowledges this:

\[(2007, 44)\]

What Elster is prepared to (and has to) accept is that his kind of mechanism is lost when, for instance, the triggering condition is identified. What he claims is, that in these instances, the mechanism is replaced by a law. This claim presupposes that the causal and explanatory information brought forward by the scientists, resulting in the loss of the local mechanism, enables us to give a covering-law explanation of the local explanandum. But this seems to presuppose too much. It may be found that the anxiety levels of a group of college students are affected by their beliefs that they are meeting the high standards of their teachers, but this does not imply that this relation can be turned into a law. For instance, if terror management theorists are right to think that self-esteem reduces anxiety because individuals believe they are a vital part of their own worldview, then it is only to be expected that meeting the high standards of your teachers is more important in some worldviews than in others.

Similarly, information about drinking habits, training conditions, water temperature, and so on, can probably be deployed to make the problem of how to
combine the facts that training increases the swimmer’s skill and confidence a whole lot easier to solve in a particular case. That would result in a vanished Elster mechanism, but not (necessarily) in the appearance of a law.

In other words, it turns out that the epistemic progress that allows us to improve our causal picture might result in the disappearance of both the mechanism and the explanation unless we are always willing to export causal facts and inferences widely. My problem with Elster’s view originates in the following observation about the differences between general and local explanatory (and predictive) advances.

The detailing of causal chains that may be followed by resolved causal indeterminacy, it seems to me, can be either a general or a local affair. There are various dimensions in which it can be general or local, but I will distinguish two: intended application and validity. We have a general affair of the first variety in cases where we assume that the causal knowledge we have acquired about the local causal chains holds wherever the causal pattern of the mechanism obtains. That such inferences are valid constitutes general affairs of the second variety. Both varieties are in line with Elster’s view of the intimate connection between mechanism and law.

However, the detailing of causal chains is clearly a local affair in contexts where we do not assume that the causal explanations we establish can be exported to other cases. The relationship between inflation and the rate of growth of real income may be found to be robust in some economies. We may nevertheless hesitate to assume it holds true everywhere. The identification of crucial components of causal chains is a local affair, of the second variety, in those cases where generalization takes place but where the inferences are not valid. These two local varieties are much more difficult to integrate with Elster’s view. Moreover, they are not mere “logical” possibilities. The second variety is well known to many social scientists as the distinction between internal and external validity (Campbell 1957). The tradition following Campbell’s use typically conceives of the problem of external validity in terms of the validity of causal generalizations (Shadish, Cook, and Campbell 2002).

It seems that in most scientific cases we improve explanations locally. For instance, when building causal knowledge of the specific causal context of our interest, we happen to acquire knowledge that resolves the indeterminacy expressed in the general causal mechanism of type A or type B (which may have guided our interests and efforts so far) as applied to the local case under scrutiny. As this happens we can describe the situation as one where we can make an internally valid causal inference. We have acquired knowledge that makes it possible to explain what happened in this context. Whether or not we are prepared to export or generalize this causal inference—and the causal facts making it possible—is an independent matter. Often the matter can be left as an open question for other studies to consider. Similarly, whether—irrespective of our thoughts and intentions on the matter—we could validly have generalized this causal inference may be an open question as far as our knowledge at the time is concerned. It depends, for example,
on whether the local causal knowledge we have relates to a situation that is similar enough to the situations covered by the intended generalization. To illustrate, Campbell’s well-known principle of proximal similarity says that we generalize an effect “with most confidence where treatment, setting, population, desired outcome, and year are closest in some overall way to the original program treatment” (Shadish, Cook, and Campbell 2002, 360).

It might be claimed, in Elster’s defence, that what is at issue is not causal generalization as such but the exportability of causal explanations. Certainly, we are in a much better situation when exporting causal explanations than when exporting causal knowledge in general. In particular, we know through which causal mediating process the outcome was affected by the cause, and this puts us in a good position “to accurately describe the components that need to be transferred to other situations to reproduce the effect” (Shadish, Cook, and Campbell 2002, 358).

However, our knowledge is always less than perfect, and we can only assume that the many components we do not include in the explanation are in fact irrelevant to it. The situation is typically such that, in comparison with the mechanistic explanation that guided our prior research, we are prepared and entitled to generalize less after the indeterminacy of the type-A or type-B mechanism has been locally resolved. After all, “If P then (sometimes) Q” (made true by the type-A or type-B mechanism) is true of more situations than “In environment E, If P then Q” (made true by the local explanatory advance in E), and there is no reason to assume that such local universal generalizations will add to the number of laws in social science. Social scientists do not appear to think so.

Two brief illustrations support this last point. First, terror management theory (TMT) claims that reminders of one’s mortality lead to a host of responses to secure symbolic immortality. To date, empirical research into TMT has been confined largely to the USA, and the countries outside the USA in which studies have been conducted are mainly in the West. So does TMT (potentially) explain part of what is generally true about humans, or does it rather capture “a motivational system of individualists” (Heine, Harihara, and Niiya 2002, 189)? Does the TMT explanation generalize, say, to Japan? Or even to western social contexts beyond those scientifically researched?

Second, the monetarist thesis is that inflation is an exclusively monetary phenomenon. The rate of growth of consumer prices, it claims, is a function of growth rates in money supply, real income and the expected cost of holding money. Interestingly, monetarist models perform better in some economies than in others. For instance, in a study by Saini, consumer prices in Sri Lanka developed in accordance with the monetarist model but that model was found to be “totally inadequate in explaining the variation in prices in South Korea” (1982, 878). Saini (1982, 880) conjectures that the monetarist conception is too rigid to apply to “developing countries where structural and institutional changes may make part of the monetary growth endogenous”.

The notion that explanatory advances are sometimes local affairs in the sense we have been discussing is strengthened by both illustrations. An explanation may arguably be satisfactory, and is certainly more than “mere description”, in a given context even if it cannot be exported to certain other contexts. In addition, TMT and monetarism point to different reasons why external validity often is perceived to be a problem for the exportability of explanations in social science. In TMT it is acknowledged that background properties in the North American context might be absent, or differently distributed, in other cultures (giving a corresponding cultural variation in where and when the determinate causal relation obtains, and hence in where the explanation applies). In the monetarist explanation, the limited exportability seems to emerge from the fact that the proposed explanatory mechanism does not include all of the relevant explanatory variables. The explanation adequately captures, or so we may assume, what happens in otherwise stable contexts, yet it breaks down with relevant structural changes.

How would an advocate of Elster’s position respond to this? The most obvious reply would involve pressing the point that Elster’s account builds on the assumption that mechanisms are replaced by laws when what generally triggers the mechanism is identified. What we get in the local cases we have been discussing are not general but local triggers. However, this reply only enforces the problem by widening the explanatory gap between Elster’s mechanisms and laws.

What should we do with the explanations in the local cases? Elster’s position offers three options: (1) the explanations are law-governed; (2) the explanations are mechanistic; or (3) there are no longer any explanations to be had in these local cases. Certainly, option (1) is not an option for Elster. His abiding worry, recall, is that laws are scarce in the social sciences. But they would not be if internally valid conclusions were sufficient to establish (local) laws. Nor is option (2) an attractive option. We have assumed that indeterminacy is locally resolved. Hence there cannot be any local Elsterian mechanisms to do the explaining. To the extent that we want to exploit Elster’s mechanisms in such cases, we have to deploy the more general mechanisms with unknown triggering conditions or indeterminate consequences that we had at the beginning. But that seems utterly strange. If I find that in environment E, P causes Q, I presumably have what I need. I am certainly not helped by citing the earlier type-A or type-B mechanism in explaining Q in E. (The earlier mechanism is now relevant to the explanation of situations that are not covered by the new explanation alone.)

The trilemma should be obvious. The new explanation cannot be mechanistic, on Elster’s conception, if it does not involve indeterminacy. And it cannot be an explanation by law, on almost any conception, since it is less general than the earlier mechanistic explanation. But then, on Elster’s account, it seems that it can be no explanation at all; that is, option (3) is correct.

There is perhaps something to be said for option (3). Detail is often needed when we move from a mechanistic sketch to an explanation (see Machamer, Darden, and Craver 2000; see also Dray 1957, 166–167), or from a partial explanation to a potential explanation (see Hempel 1970). However, the notion that the
more detail we have, the better the explanation is no truism. Sometimes detail has a negative impact on explanatory power. Thus it is well known that irrelevant information can destroy an explanation.

In spite of this it is clear that option (3) leads to paradox. There is explanatory power to begin with, when the mechanism is in place. Then, through the addition of relevant causal truths, that power is annihilated (since the mechanism has to be replaced in the face of locally resolved indeterminacy). This cannot be right. The causal detail we are adding in these cases is much more important than the potentially explanation-annulling detail we hinted at in the paragraph above. The suggestion that crucial pieces of causal knowledge always obliterate explanatory power is plainly false. Consider one of Elster’s hypothetical examples: high unemployment causes wars of aggression. An inevitable explanation-seeking question, following this finding, will be how this is possible. Could it be that the unemployed tend to vote for populist leaders who are likely to give up on diplomacy and instead resort to war to deal with conflicts? This speculation is a first attempt to detail the causal relations and resolve the indeterminacy of the attempted mechanistic explanation. If it is successful, it will inspire the investigations of the further causal chain just suggested: why would the unemployed vote for such leaders? And so on. It is intuitively apparent that discoveries about the relevant causal chains will generally promote explanatory progress. Whichever concept of explanation one starts with, rendering an explanation more fine-grained by tracking the actual causal chains will potentially add explanatory value (as long as the explanation is still understandable).

This idea should be applicable to many of the other mechanisms Elster suggests as well. Engaged in the particular task of finding more causal truths associated with the mechanism indicated by the proverb “the best swimmers drown”, we would be very likely to find reasons to fill in the causal picture between confidence and risk-taking, and also between skill and risk management. Even a layperson following reports of drowning accidents in the media may be able to suggest one or two relevant causal relationships: not only does it seem to be a fact that men swim into open water and occasionally drown, while women swim along the shore and survive, but men seem also to be the victims of over-confidence engendered by unsuitable training. Some men who once swam 2 km in an indoor pool believe they will have the same capacity in colder, rougher, unsafe lakes and seas—according to reports in the media, at least. It cannot be doubted that the mechanistic picture ought to improve in a number of ways through the discovery of causal relations like these, so plainly, in general, option (3) is not a viable position.

Briefly, then, the correct position is that explanatory power does not necessarily decrease as a result of the successful detailing of causal chains. On the contrary, added causal knowledge resolving the uncertainty and indeterminacy depicted by Elster’s type-A and type-B situations often, and radically, improves the explanation. But Elster’s current concept of mechanism cannot accommodate this truth—a good reason to resist Elster’s characterization of mechanisms.
Notes

[1] The epistemic conception of mechanisms is clearly not the only option. Ontic conceptions, where mechanisms relate real, worldly existents (see Persson 2010), are probably more common. An example would be the definition of mechanisms in McAdam, Tarrow, and Tilly (2001, 24): “Mechanisms are a delimited class of events that alter relations among specified sets of elements in identical or closely similar ways over a variety of situations”. Ontic conceptions are seldom vulnerable to the problem discussed here. The mechanistically causal properties of the ontic mechanism are not dependent on our knowledge.


[4] However, such a reductive strategy is often found to be problematic in its own right. There are several examples from theorizing in psychology where the introduction of simple dichotomies—such as between System 1 and System 2 in dual process theories of judgement and decision-making—reduces theoretically fruitful complexity but retains the difficulties that led to the development of competing theories in the field to begin with (see Sahlin, Wallin, and Persson 2010; Gigerenzer 2009). The analogous charge in Elster’s case would be that his concept of mechanism might lack the precision and potential needed to guide further developments in the field, and that his concept of deterministic law is already known to be inapplicable to most contexts in the social sciences.

References


