

Karl-Dieter Opp

Explanations by mechanisms in the social sciences. Problems, advantages and alternatives

Received: 1 June 2004 / Accepted: 27 December 2004
© Fondazione Rosselli 2005

Abstract This paper discusses various problems of explanations by mechanisms. Two positions are distinguished: the narrow position claims that only explanations by mechanisms are acceptable. It is argued that this position leads to an infinite regress because the discovery of a mechanism must entail the search for other mechanisms etc. Another paradoxical consequence of this postulate is that every successful explanation by mechanisms is unsatisfactory because it generates new “black box” explanations. The second – liberal – position that is advanced in this paper regards, besides explanations by mechanisms, also the discovery of bivariate correlations as a first step of an explanation by mechanisms as meaningful. It is further argued that there is no contradiction between causal analysis and the explanation by mechanisms. Instead, explanations by mechanisms always presuppose the analysis of causal structures (but not vice versa). The final point is that an explanation by mechanisms is not inconsistent with the Hempel-Oppenheim scheme of explanation.

Keywords Explanation · Explanation by Mechanisms · Causality · Causal Analysis · Hempel-Oppenheim explanation

1 Introduction

Every social scientist will probably agree that the discovery of social mechanisms that explain or bring about certain phenomena is a fruitful procedure.

K.-D. Opp
Institut für Soziologie, Universität Leipzig, Emeritus

Present address: K.-D. Opp
Sulkyweg 22, 22159 Hamburg, Germany
E-mail: opp@sozio.uni-leipzig.de

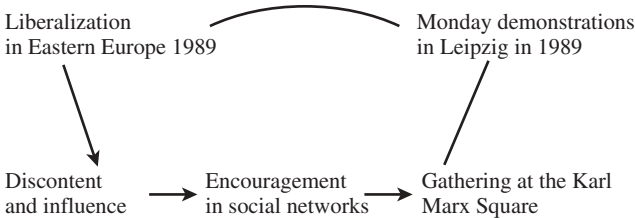
But what exactly is a “social mechanism” or an “explanation by mechanisms”? Why are explanations by mechanisms so universally hailed? What exactly is the difference between an explanation by mechanisms and other kinds of explanations? Is, for example, an explanation by mechanisms inconsistent with the Hempel-Oppenheim explanatory procedure? Is there a difference between explanations by mechanisms and the formulation of causal models? If there are differences: how are they to be judged? These are the questions this paper focuses on. But first it is useful to define what is to be understood by an explanation by mechanisms.

2 What is an explanation by mechanisms?

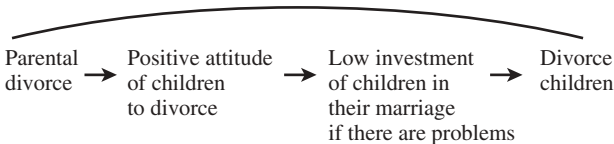
Let us begin with two examples (see the models in Figure 1). Our first example (see model 1a in Figure 1) claims that the liberalization in Eastern Europe (i.e. in the Soviet Union, Poland and Hungary) in 1989 was a cause for the Monday demonstrations in Leipzig in October 1989 (and, thus, contributed to the collapse of the German Democratic Republic - GDR). It is thus assumed that there is a causal relationship between the liberalization of Eastern Europe and the origin of the Monday demonstrations. Put differently, the Monday demonstrations are *explained* by the liberalization in Eastern Europe.

Although this explanation may sound plausible at first sight it is not satisfactory. One would like to know how the relationship between the two events (liberalization and Monday demonstrations) came into being, i.e. it

a Micro-macro model



b Micro-model



c Structure of an explanation by mechanisms



Fig. 1 Examples of mechanisms

would be interesting to find out what causal process or what causal processes can explain or generate this causal relationship.

What could these causal processes have been? The liberalization could have increased dissatisfaction of the GDR population as well as their belief that a liberalization could be achieved also in the GDR if oneself (and enough other citizens) participates in protests. Liberalization thus raised discontent and perceived personal political influence. This, in turn, had the effect that citizens discussed the possibilities of political action among each other and encouraged each other to engage in some political activity directed against the government. At that time, people who wanted to leave the GDR attended the Monday Prayers in the Nikolai Church (that took place since 1982) which is close to the Karl Marx Square and expressed their discontent. This was for many other citizens an incentive to go to the Nikolai church and the adjacent Karl-Marx Square to show their discontent to the communist regime. This, then, is the process that brought about the Monday demonstrations.¹

This explanation sketch shows how a relationship between two events – the liberalization of Eastern Europe and the Monday demonstrations – came about. A causal process is described that is set off by the liberalization in Eastern Europe and finally gives rise to the Monday demonstrations. It is important to note that the process is a causal chain of the kind “A leads to B, B leads to C”. Thus, intervening variables are specified that are inserted between “liberalization” and “Monday demonstrations”.

The effects of the liberalizations were brought about in our example because certain conditions existed already. For example, without the peace prayers the Monday demonstrations would perhaps never originated. It is further important that the relationship between the macro-variables (liberalization and Monday demonstrations) in model 1a is regarded as a correlation (symbolized by the arc between “Liberalization...” and “Monday demonstrations...”²) and not as causal effect although at the beginning of the example a causal statement was made claiming that the liberalization *brought about* or *contributed to* the demonstrations. But the explanation shows that liberalization does not have a direct effect on the demonstrations; instead, the relationship between these variables is explained.² Furthermore, the relationship between “gathering at the Karl-Marx Square” and “Monday demonstrations” is no empirical (i.e. causal) but an analytical (i.e. definitional) relationship (symbolized by the line between “Gathering ...” and “Monday demonstrations...”): if many people convene at a certain place in order to express their discontent this is by definition a demonstration.

The previous example starts with a macro-relationship. It is then shown that the macro-cause (liberalization) changes certain properties of individual actors (the micro-level) and that these changes had a causal effect on the dependent macro-variable (demonstrations). A similar causal structure may not only exist as a micro-macro model, one may find such causal patterns

¹ For a more detailed explanation of the Monday demonstrations and its empirical test see Opp et al. 1995.

² This is, in the terminology of Lazarsfeld (1955) an “interpretation,” i.e. the variables added to the original correlation are intervening variables.

also on the micro-level, as our second example (model 1b) shows. This example refers to the divorce transmission proposition (see, e.g., Diefenbach 2000): if parents got divorced the likelihood is high that the children get divorced as well. In this case, one would also like to know how this relationship can be explained. It is plausible that the divorce of the parents has brought about a relatively positive attitude of the children to divorces in general; this, then, has the effect that investments in sustaining a problematic marriage will not be likely. This raises the likelihood of a divorce.

This explanation is similar to the previous one at least in two respects: the starting point is in both cases a relationship between two variables – in the latter case it is the relationship between divorce of the parents and divorce of the children – that is to be explained. The explanation consists of adding intervening variables and, thus, is a causal chain. Furthermore, the relationship between “divorce of parents” and “divorce of children” is a correlation and no causal relationship.

Both examples illustrate “mechanisms”³ which can now be characterized in a general way as follows:

An *explanation by mechanisms* exists if it can be shown how a relationship between variables is brought about.⁴

This definition implies the following: if there is a relationship between variables and if at least one intervening variable is specified an explanation by mechanisms is given. The minimal mechanism is thus a causal chain in which there is a correlation between an independent and a dependent variable and in which an intervening variable is added. Such a *minimal mechanism* exists, e.g., if in model 1a the intervening variable is only “discontent and influence” (i.e. an interaction term). The causal chain then reads:

Liberalization → Discontent and influence → Demonstrations

Our examples are simplified because only one single independent variable influences one single dependent variable. Models in actual research consist of several independent variables. Model 1a, e.g., would have to be extended by including in addition to liberalization events such as the emigration wave and the existence of the Monday prayers as independent variables which were mentioned but for the sake of simplicity not included in the model. It is also possible to formulate models with several dependent variables. The above definition of a mechanism implies that a causal chain is also to be called a mechanism if it is more complex than in the examples.

A second implication that was already mentioned is that an *explanation by mechanisms is not identical with a micro-macro explanation*. As our first example suggests, a micro-macro explanation is always an explanation by

³ A great many other examples with further references can be found in Hedström and Swedberg 1998a.

⁴ This definition is consistent with the characterization by Hedström and Swedberg (1998a) who begin the introduction to this book in the following way: “The main message of this book is that the advancement of social theory calls for an analytical approach that systematically seeks to explicate the social mechanisms that generate and explain observed associations between events” (Hedström and Swedberg 1998b, p. 1).

mechanisms. Our second example shows, however, that not every explanation by mechanisms is a micro-macro explanation. An explanation on the micro-level with intervening variables also counts as an explanation by mechanisms.

A third implication of our definition will be illustrated with an example. Many empirical studies indicate that members of the middle and upper class protest more frequently than members of the lower class. This correlation could be explained by pointing out that the middle and upper classes are politically dissatisfied to a high extent, that their political efficacy is relatively high and that these factors are causes for political action. In this explanation the correlation between class and protest is explained by the *correlation* of class with discontent and efficacy and not by an *effect* of class on discontent and efficacy. In general, thus, the intervening variable need only correlate with the other variables, there need thus not be causal effects.

Based on the previous examples, the most basic form of an explanation by mechanisms can be depicted in the following way (see also Hedström and Swedberg 1998b, pp.7–9). In model 1c in Figure 1 the starting point is a correlation between two entities I (input) and O (output) – see the arc between “I” and “O”. In order to explain this relationship one looks for a “mechanism”; this mechanism generates the effect or – equivalently – output O if input I obtains. This “generation” of O may be a causal effect but also – see the previous example – a correlation or an analytical aggregation. The latter is symbolized by the lines. Mechanism M is a “systematic set of statements that provide a plausible account of how I and O are linked to one another” (p. 7). Model 1c illustrates this definition. Is the intervening process M lacking a *black-box explanation* exists. This means that the relationship between input and output is not explained (or is not of interest, see p. 9). “Systematic set of statements” may refer to a clearly structured set of statements. It is not clear what a “plausible account” means – we will return to this later on.

The concept of “mechanism” is one of those concepts that is often used in a vague and inconsistent way. For example, the expression “market mechanism” refers to the functioning of the market, i.e. to *processes* such as price formation. Accordingly, a plea for the explanation by mechanisms would mean that social processes should be explained. Such a claim does not specify how complex the processes should be. A correlation such as “divorce of parents correlates with the divorce of children” is a process in the sense that an event obtains after another event has occurred. However, the claim to explain social processes probably means that more “complex” processes should be explained that extend to more than two points in time. But how many points in time are regarded as meaningful is not clear. It is worth noting that this kind of “explanation by mechanisms” does not start with a correlation between variables that is to be explained but with a causal statement.

We find many other definitions of “mechanism” or “explanation by mechanisms” in the literature. We do not intend to discuss these definitions. Such definitions are meaningful if they refer to a fruitful explanation strategy. This is the case with our definition above: the explanation of a given empirical relationship in the social sciences has always been regarded as an important scientific progress, and it is not important whether this is a

relationship on the micro- or macro-level – see the two examples above. In what follows we will therefore use our previous definition.

3 Are explanations by mechanisms meaningful?

It seems that the answer to this question is so obviously “yes” that no detailed arguments are necessary. The general argument is that explanations by mechanisms lead to a “deeper” (or more informative) explanation in the sense that we learn how a relationship between variables comes about. An explanation by mechanisms thus leads to a scientific progress.

Problems emerge, however, if a *narrow position* is taken. This is the claim that a mechanism must be found for each relationship between variables. In contrast, a *liberal position* values, among other things, also the mere discovery of a bivariate correlation positively. We will discuss this position in more detail below. At this point we will be concerned with the narrow position. This is burdened with two problems.

Let us illustrate the first problem with our first example. Starting with the original relationship “the liberalization in Eastern Europe generated the demonstrations in Leipzig” we succeeded in finding a mechanism that shows how this relationship was generated. If the discovery of a bivariate correlation for which a mechanism has not yet been found is valued negatively one must also explain the relationship “discontent/influence leads to the encouragement in social networks” (model 1a). This hypothesis is, by definition, a black box because it is not clear how this relationship is generated.

But assume that we have explained this relationship again in the following way: high discontent and high personal influence activate a norm that one should be politically active and that one should encourage others to participate. The explanation now reads:

Discontent/influence → Norm → Encouragement

We thus have found another mechanism and thus reached a scientific progress. However, according to the general claim that only explanations by mechanisms are to be accepted – i.e. according to the narrow position – this explanation is again not satisfactory. The reason is that we got again two black-box explanations:

Discontent/influence → norm

Norm → encouragement

Thus, if a new mechanism is discovered new causal chains and, thus, new black-box explanations originate which are unsatisfactory according to the narrow position. This position thus paradoxically implies:

Implication of the narrow postulate of an explanation by mechanisms: The more mechanisms are discovered, the more black-box explanations originate and the more unsatisfactory the explanation becomes.

Therefore, if the goal is to avoid black-box explanations it is better not to suggest any explanation by mechanisms! If a bivariate explanation exists only one mechanism is lacking; the more mechanisms are discovered and, thus, the longer the causal chain becomes, the more mechanisms are required – and the less satisfactory the explanation is.

Another problem of the narrow position is that it leads to an *infinite regress*. If one has found a mechanism one must search for another mechanism; if this mechanism found one must try to find another one etc. In order to avoid such an infinite regress there are several possibilities: one is to stop the search.⁵ But where should the justification procedure be broken off? This question is not addressed in the literature and we will not discuss it because we do not adhere to the narrow position.

These implications of a narrow position are certainly not desirable. How could the postulate of an explanation by mechanisms be reformulated so that the consequences mentioned do not originate? The narrow postulate could be replaced by a *liberal postulate*. Such a heuristic principle holds that the discovery of mechanisms is principally positive. But finding a relationship between two variables is not valued negatively if a mechanism has not yet been found that can explain this relationship. Even if “only” a relationship between two variables is discovered scientific progress is achieved. One will further welcome an explanation of such a relationship (i.e. a mechanism).⁶ This is a rather *liberal* and not a *narrow* postulate of explanation by mechanisms.

Even if an explanation by mechanisms leads to a scientific progress the question arises whether there might be other procedures that also achieve this. What are these alternative explanation strategies? It is argued in the literature that these are in particular causal analysis and the covering law model.

4 Explanations by mechanisms and causal analysis

Hedström and Swedberg (1998b, p. 2) distinguish between “variable-based” and “mechanism-based” approaches. What this distinction means becomes clear by a quotation from a paper by A. B. Sørensen (*ibid.*, p. 4). He claims that sociologists dealing with the labor market regard theory as something that addresses the question

which variables should be included in the equations and how these variables relate to other variables – and not as something which is about which mechanisms produce the observed associations in the variables.

⁵ This problem is similar to the “Münchhausen-Trilemma” in the philosophy of science if philosophers subscribe to the principle of sufficient justification (i.e. for each statement a justification must be found). Such a principle leads to an infinite regress or to a logical circle (that is to say, in order to justify some statement one uses other statements that have been regarded as requiring justification before) or to an arbitrary breaking off of the justification. For details see Albert 1968, in particular pp. 11–15.

⁶ This would correspond to the replacement of the principle of sufficient justification with a principle of general criticism which rests on the philosophy of Karl R. Popper. See Albert 1968, pp. 37–41.

This statement seems to mean that explaining phenomena by mechanisms does not involve the formulation of equations, and it does not imply that it is specified which variables are related to other variables.

Hedström and Swedberg argue that a causal analysis provides only a black-box explanation (they refer to model 1c in Figure 1):

a regression coefficient linking I and O, and this regression coefficient (if the model includes all relevant variables) is supposed to describe the causal influence of I upon O. The approach advocated here does not rest with describing the strength and the form of the relationship between the entities of interest but addresses a further and deeper problem: how (i.e., through what process) was the relationship brought about?

The authors seem to maintain that a causal analysis only addresses a causal relationship between I and O. But the explanation of this relationship, i.e. the mechanism, apparently has nothing to do with causal relationships. This interpretation is supported by the claim “that it is actors and not variables who do the acting” (p. 24). However, another interpretation is suggested on pp. 21–25 where the authors seem to say that explanations by mechanisms imply causal structures. Whatever the authors mean, other authors clearly advance the view that explanations by mechanisms are identical with the formulation of complex causal models (see, e.g., Gambetta 1998, p. 102). Due to these different opinions it seems worthwhile to address the question to what extent an explanation by mechanisms differs from causal modeling.

5 Explanations by mechanisms as complex causal modeling

In this section we will argue that the claim explanations by mechanisms do not involve causal modeling is wrong. The opposite is true: each explanation is identical with a complex causal model. The examples of Figure 1 illustrate this. The mechanism described in model 1a indicates that liberalization affects certain incentives on the individual level. These incentives, along with certain given structures – e.g., social networks and institutions such as the Monday prayers – , had the consequence that many citizens convened at the Karl Marx Square. This collective action is, by definition, a demonstration. This argument is a complex causal model. It is possible to test it empirically. For example, data about the dissatisfaction of the Leipzig population, about their perceived political influence and about the extent to which friends of the respondents encouraged their participation in the demonstrations could be used to test the model.⁷ When such data are available the equations of the complex causal model can be estimated by regression or other multivariate techniques. It is also important to estimate coefficients about the extent to which the variables of the models affect

⁷ Such data were ascertained in our survey administered in 1990 referring to the situation in 1989. For details see Opp et al. 1995 (German version 1993).

participation in the Monday demonstrations or other types of political action.

The same holds for the second example: the mechanism also consists of explaining a relationship between variables by adding intervening variables. This is again a causal model that can and has been tested by regression analysis (Diefenbach 2000, Chapt. 5).

Our claim that explanations by mechanisms are causal models becomes still more convincing when we look at the explanatory argument put forward by Hedström and Swedberg themselves – see model 1c in Figure 1. We add to their model the starting point of the explanation by mechanisms, viz. the correlation between two variables (symbolized by the arc between I and O). The authors assume that there are causal effects from I on M and from M on O that show how the correlation between I and O is generated (the authors thus do not take into account that there may exist only correlations between I and O on the one hand and the intervening variables and M on the other; we added these correlations in model 1c in Figure 1 – see the lines between the variables).

The previous examples clearly falsify the claim that an explanation by mechanisms is something different than the analysis of causal structures. Thus, if an explanation by mechanisms is provided then causal relationships are proposed.

However, the opposite is not true: not every causal model is an explanation by mechanisms. For example, if only the relationship between divorce of the parents (and perhaps other independent variables) on the one hand and the divorce of the children on the other is analyzed causal modeling occurs, but there is no explanation by mechanisms.

Hedström and Swedberg provide evidence for their position described above by citing authors such as Raymond Boudon and James S. Coleman. However, a closer look at those references shows that these authors do not deny that an explanation by mechanisms is an analysis of causal structures. Instead, they complain that in the literature given statistical correlations are often not further explained, i.e. that the causal structures that generate such correlations are not analyzed – an argument every reader will probably agree with.

This is illustrated by two contributions which the authors refer to in support of their position. Boudon (1979) criticizes that “the interpretation of a statistical table or of a set of statistical tables seldom ends with a causal analysis” (p. 51). As an example he refers to a table from “The Americana Soldier” by S. Stouffer and others that shows an interaction effect. Such a table, Boudon notes, does not yet explain the relationship found. Only the explanation for this statistical relationship makes it understandable. Boudon then continues: “This example is paradigmatic: *understanding* a statistical structure means in many cases building a generating theory or model (in the case where the theory needs to be formalized and has actually been) that includes the observed empirical structure as one of its consequences” (p. 52). It is important to note that this argument is not directed against the analysis of causal structures. Instead, it is claimed that a mere bivariate causal structure should not be regarded as given but that it should be explained.

Coleman (1986) does not criticize the analysis of causal structures either but that “causal explanation based on statistical evidence has replaced purposive explanation. One way of describing this change is to say that statistical association between variables has largely replaced meaningful connection between events as the basic tool of description and analysis. The ‘meaningful connection’ was ordinarily provided by the intentions or purposes of an actor or combination of actors” (pp. 1327–1328). Again, this argument does not denounce the analysis of causal structures but criticizes that explanations are often not based on the goals and intentions of the actors. I find these arguments absolutely convincing.

Let us finally look at the previous argument “that it is actors and not variables who do the acting” (p. 24). Does this mean that an explanation by mechanisms is not identical with the analysis of causal structures? It cannot be denied that individuals and not variables act, but “acting” of actors is a variable that is ascribed to individuals. If, thus, the effects of individual action are explained then variables (that are features of individuals) are causally related to other variables.

6 A plea for causal analysis

Is the critique of causal modeling referred to before justified? It can hardly be denied that many causal models are not concerned with the causal structures of mechanisms. Models analyze, e.g., macro-relationships between variables without showing how such relationships are brought about. Whatever the weaknesses of existing causal models may be: these flaws do not imply that causal analysis can be dismissed as a fruitful instrument for a formulation or empirical test of mechanisms or that the estimation of regression coefficients is to be criticized. The opposite is true, as we tried to show before: if, e.g., macro-relationships are explained by indicating how certain macro-factors changed individual incentives which, in turn, affected individual action and, again, macro-events (such as demonstrations – see model 1a in Figure 1) it is of interest how strong the (perceptions of) macro-factors affect the individual incentives and how strong their impact is on individual action. Even if one has to admit that causal analysis is often not meaningfully applied it is an important instrument that helps to clarify and test explanations by mechanisms.

Even if explanations by mechanisms are regarded as fruitful the question is whether it is questionable to formulate only models with a few independent and a dependent variable (or a group of dependent variables). If a relationship between variables is to be explained by mechanisms such a relationship must first be discovered. It often requires detailed causal analyses to find out whether a relationship on the macro- (or micro-) level exists. An example is the relationship between corporatism and unemployment (see, e.g., Kenworthy 2002). Is the hypothesis correct that corporatism diminishes unemployment? If yes: what kind of corporatism generates a diminution of unemployment? It is by no means superfluous or meaningless if as a first step macro-relationships are established. Only then the next step can be taken to ask which mechanism generated this relationship. It often happens that data

do not exist that substantiate the existence of the presumed mechanism, or the effort to test the one-level causal model is already so high that the next step has to be left to further research.

If the discovery of a mechanism is endorsed it is strange if the formulation and test of causal models is rejected that are a first step in an explanation by mechanisms. If the *explanation* of correlations is considered desirable then the *discovery* of such correlations should be valued positively as well. But, as emphasized before, the correlations are only a first step in an explanatory argument.

It is further meaningful to *test empirically* models with some independent and a dependent variable or a group of dependent variables (i.e. single-stage models) that are part of an explanation by mechanisms. This may be illustrated with our first example. A central part of the explanation of revolutions is a micro-theory of the participation of individuals in political protest: if the goal is to explain how macro-events (such as the repressive behavior of security forces) affect the size of demonstrations it must be examined how such events change the incentives for participation at the micro-level. What are these incentives? To answer this (and similar) questions empirical research is necessary in which the incentives are independent and participation in demonstrations (or other forms of political action) dependent variables. In analyzing the data it is also of interest to estimate the weight of the single incentives, i.e. the Beta- or B-coefficients.

The formulation and empirical test of single-stage causal models is important for another reason: such models could become components of explanations by mechanisms without knowing it at the time when such models are proposed and tested. Assume, for example, the hypothesis is tested whether the identification with a region has an effect on participating in political action designed to promote the welfare of the region. According to the narrow position this hypothesis is not meaningful because no mechanism is specified that explains how this correlation is generated. However, the correlation mentioned could be part of an explanation by mechanisms. For example, assume it is found that the unification of Europe diminishes political action in sub-regions (such as countries). This relationship could be explained in the following way: the unification of Europe diminishes identification with sub-regions. If the lower identification with a region raises the costs of political action for the region then this explains why the unification of Europe diminishes political action in sub-regions. To summarize:

Original proposition: Regional identification → decrease of political action.

Intervening variables (mechanism): Identification with Europe → lower identification with sub-regions → higher costs of political action for increasing the welfare of the region → decrease of political action.

Proponents of causal modeling do often not engage in formulating and testing explanations by mechanisms, and they are criticized for this. However, the often almost emotional aversion of proponents of rational choice theory against causal modeling (and, in general, against surveys that are often used to test causal models) is to be criticized as well. In general,

contemporary rational choice theory should be more often tested in natural situations, and for this purpose causal modeling and surveys are useful. The general critique of surveys and causal modeling has the consequence that a useful instrument for the test of propositions is too rarely used (for details see Opp 1998).

7 The explanation by mechanisms and the covering-law model

According to the Hempel-Oppenheim scheme – henceforth HO-scheme – an explanation consists of an explanandum (i.e. a phenomenon or phenomena to be explained) that is derived from at least one lawful statement and the pertinent initial conditions (for details see, e.g., Hempel 1965). Let us illustrate this procedure with our first example:

Law: If political discontent and perceived personal influence are relatively intense, the frequency of participation in demonstrations is high.

Initial conditions: In October 1989 discontent and perceived influence of the population of Leipzig increased.

Explanandum: The participation in the demonstrations in Leipzig increased in October 1989.

The initial question in this explanation is: why did the number of participants in the Monday demonstrations increase in October 1989? The answer is: because discontent and perceived influence (initial conditions) increased *and* because the following law holds: increasing discontent and influence raise participation in demonstrations. This is a deterministic explanation because it is assumed that if the initial conditions exist the explanandum will always (i.e. without exception) occur. But most of the time laws are not deterministic. The above law might have a non-deterministic form such as “the higher discontent and influence, the more likely is participation in demonstrations”. In this case, the explanandum can no longer be logically deduced from the law and the initial conditions but can only be confirmed (for details see Hempel 1965).

It is not possible and also not necessary in this article to enter the extensive discussion about the structure of scientific explanations. One major point of criticism is directed against applying a law in an explanation. However, assume the law is missing. The explanation would then read: the size of the demonstrations increased because people became more dissatisfied and influential. This explanation is highly problematic because we don't dispose of any information telling us why discontent and influence and why not other factors are relevant. Without a law it is thus possible to suggest any other factor which preceded the demonstration in time. The selection of causal factors is thus arbitrary. Only a law provides a *selection criterion* for the factors that have caused a phenomenon. In the extensive discussion about the logic of explanation there is no convincing alternative to how the problem of the arbitrary selection of causal factors can be solved unless by using lawful statements.

Hedström und Swedberg (1998b, pp. 8–9) claim that an explanation by mechanisms “differs in important respects from the classical covering-law model” (p. 8) – they are referring here to the HO-scheme. What are these differences?

- (1) The authors correctly note that the quality of an explanation depends on the quality of the law. If this law is merely a statistical relationship – they seem to refer to a spurious correlation – then the explanation is not acceptable. However, a spurious correlation is no candidate for a law in an explanation (if it is known that a statement is a spurious correlation). A statement qualifies as a law only if it has empirical content and is empirically corroborated.⁸ Only explanations that include such a statement are adequate. This is well known for a long time (see the first detailed study of the logic of explanation by Hempel and Oppenheim 1948).
- (2) A second argument of the authors against the covering law model reads:

The covering-law model provides justification for the use of “black-box” explanations in the social sciences because it does not stipulate that the mechanism linking explanans and explanandum must be specified in order for an acceptable explanation to be at hand (p. 8).⁹

In other words, the HO-scheme does not demand explanations by mechanisms and it does not prohibit an explanation consisting of a single-stage model (initial conditions and explanandum); this means – the authors assert – that a black-box explanation (i.e. an explanation without a mechanism) can be justified by the HO-scheme.

Is it really possible to assert that if something is not explicitly demanded (such as an explanation by mechanisms) the opposite is automatically justified (a black-box explanation)? If a scientist does not explicitly endorse explanations by mechanisms does this imply that he or she justifies a black-box explanation? This argument may be written in the following way – the line between the sentences means, as before, that the sentence under the line follows logically from the sentence above the line:

Premise : An explanation without mechanisms is neither adequate nor inadequate.

Conclusion : An explanation without mechanisms is justified.

We do not know of any logical rule that allows the derivation of the conclusion from the premise. Rather, if the HO-scheme does not explicitly demand explanation by mechanisms it is left open which procedure is meaningful. It can thus not be maintained that the HO-scheme justifies black box explanations.

⁸ Strictly speaking, the law must be true. But because the truth of a law can never be established we can only apply the criterion of empirical confirmation.

⁹ It seems that they refer to the “link” between the initial conditions (and not explanans which consists of the law and the initial conditions) and the explanandum.

Can't we nevertheless criticize the HO-scheme because it does not explicitly demand explanations by mechanisms? Not at all. The scheme describes and demands a general procedure that holds for all sciences. It is thus not possible and probably also not desirable to list all features of laws or explanations that are specific for each discipline. One could even say that it is a strength of the HO-scheme that it does neither demand nor exclude explanations by mechanisms. This has the consequence that the HO-scheme is not burdened with the problems mentioned before (if a narrow position is advanced).

If social scientists agree that explanations by mechanisms are desirable why not *complement the HO-scheme by a methodological postulate referring to explanations by mechanisms?* This would be a better strategy than dismiss the whole scheme because no better alternative to the HO-procedure is available.

- (3) Hedström and Swedberg (1998b) seem to believe that explanations by mechanisms are completely different from explanations with the HO-scheme. In their extensive discussion of explanations by mechanisms it is not mentioned that an explanation by mechanisms also uses laws. Maybe the authors believe that this is not the case or not necessary. The above quotation claims that explanations by mechanisms are a "plausible account". Thus, explanations by mechanisms seem to be different from the covering law model. Is this claim tenable?

If it is argued that explanations by mechanisms don't require laws one would like to know how the factors that are claimed to be causal are selected. The authors don't answer this question. For example, if research shows that state repression has sometimes positive and sometimes negative effects on collective political action: how do we find the intervening variables between repression and collective action? Without applying a lawful statement or a theory in the sense of a set of laws this question can only be answered ad hoc. If the authors invoke a "plausible account" (see the quotation above) the question arises how it can be judged whether an account is more or less plausible. What decision is taken if different mechanisms are plausible or if there is a scientist A for whom one mechanism is plausible whereas there is another scientist B who finds another mechanism plausible? Laws are, as said before, a selection device for the relevant factors. If it is claimed to dispense with laws without proposing a better alternative, the implication is that explanations by mechanisms become arbitrary.

Proposing a "plausible account" may also be (probably mis-) understood in the way that rigorous research is not needed in judging the validity of an explanation by mechanisms. However, in applying the HO-scheme implies that the conditions of adequacy of an explanation are to be applied as well. One of these conditions is that all statements of an explanation must be empirically corroborated.

- (4) Elster (1998, p. 48) also claims that there is a difference between an explanation with laws and an explanation by mechanisms: "a law has the form 'If conditions $C_1, C_2, C_3, \dots, C_n$ obtain, then *always* E.' [...] a

statement about mechanisms might be ‘If $C_1, C_2, C_3, \dots, C_n$ obtain, then *sometimes E*’ (italics not in the original). Everybody is free to restrict the meaning of the expression “explanation with laws” to cases where laws are deterministic and define an explanation by mechanisms as an explanation with non-deterministic laws. However, as said before, the HO scheme also holds for non-deterministic laws. If Elster pleads for an explanation with mechanisms this means that he favors applying non-deterministic laws if deterministic laws are not available. Who would disagree with this?

8 Conclusion

Reading this paper might convey the impression that this is a plea against the explanation by mechanism. Such an impression would be completely wrong. The paper only discusses some problems that come up if a narrow postulate of explanations by mechanisms is put forward. This postulate states that each explanation that does not specify a mechanisms is to be rejected out of hand. Those problems do not exist if a liberal position is adopted which is principally in favor of an explanation by mechanisms but does not reject any other explanation out of hand. Furthermore, some claims are criticized that assign explanations by mechanisms a special methodological status: allegedly, explanations by mechanisms differ from causal models on the one hand and from explanations according to the Hempel-Oppenheim scheme on the other. I tried to show that both claims are not tenable. Finally, it is argued that explanations by mechanisms require the same rigorous empirical control as any other explanatory argument.

It would be important for further research that those social scientists who test causal models would concentrate more on explanations by mechanisms. A deficiency of the present state of the theory of rational action is that there is in general too little empirical research that tests rational choice propositions and that there is in particular too little research in natural situations. It would be important for the progress of our discipline if both “camps” – the causal modelers and those favoring exclusively explanations by mechanisms – would more extensively formulate and test complex causal models that are explanations by mechanisms, based on a general theory of action.

References

- Albert H (1968) *Traktat über kritische Vernunft*. Mohr, Tübingen
- Boudon R (1979) *Generating models as a research strategy*. In: Merton RK, Coleman JS, Rossi PH (eds) *Qualitative and quantitative social research*. Free Press, New York, pp 51–64
- Coleman JS (1986) *Social theory, social research, and a theory of action*. *Am J Sociol* 91:1309–1335
- Diefenbach H (2000) *Intergenerationale Scheidungstransmission in Deutschland. Auf der Suche nach dem “missing link” zwischen Ehescheidung in der Elterngeneration und Ehescheidung in der Kindgeneration*. Ergon, Würzburg

-
- Elster J (1998) A plea for mechanisms. In: Hedström P, Swedberg R (eds) *Social mechanisms. An analytical approach to social theory*. Cambridge University Press, Cambridge, pp 45–73
- Gambetta D (1998) Concatenations of Mechanisms. In: Hedström P, Swedberg R (eds) *Social mechanisms. An analytical approach to social theory*. Cambridge University Press, Cambridge, pp 102–124
- Hedström P, Swedberg R (eds) (1998a) *Social mechanisms. An analytical approach to social theory*. Cambridge University Press, Cambridge
- Hedström P, Swedberg R (1998b) *Social mechanisms: An introductory essay*. In: Hedström P, Swedberg R (eds) *Social mechanisms. An analytical approach to social theory*. Cambridge University Press, Cambridge, pp 1–31
- Hempel CG (1965) *Aspects of scientific explanation and other essays in the philosophy of science*. Free Press, New York and London
- Hempel CG, Oppenheim P (1948) Studies in the logic of explanation. *Phil Sci* 15:135–175
- Kenworthy L (2002) Corporatism and unemployment in the 1980s and 1992. *Am Sociol Rev* 67:367–388
- Lazarsfeld PF (1955) Interpretation of statistical relations as a research operation. In: Lazarsfeld PF, Rosenberg M (eds) *The language of social research*. Free Press, Glencoe, IL, pp 115–125
- Opp K-D (1998) Can and should rational choice theory be tested by survey research? The example of explaining collective political action. In: Blossfeld H-P, Prein G (eds) *Rational choice theory and large-scale data analysis*. Westview Press, Boulder, CO, pp 204–230
- Opp K-D, Voss P, Gern C (1993) *Die volkseigene Revolution*. Klett-Cotta, Stuttgart
- Opp K-D, Voss P, Gern C (1995) *The origins of a spontaneous revolution. East Germany 1989*. Michigan University Press, Ann Arbor