Causal Mechanisms in the Social Sciences

Article *in* Annual Review of Sociology · August 2010

DOI: 10.1146/annurev.soc.012809.102632

CITATIONS
914

READS
9,513

2 authors:

Peter Hedström
Linköping University
49 PUBLICATIONS **5,568** CITATIONS
SEE PROFILE

Petri Ylikoski
University of Helsinki
69 PUBLICATIONS **1,553** CITATIONS
SEE PROFILE

Causal Mechanisms in the Social Sciences

Peter Hedström
Nuffield College
University of Oxford
peter.hedstrom@nuffield.ox.ac.uk
(corresponding author)

Petri Ylikoski
Department of History and Philosophy
University of Tampere
petri.ylikoski@helsinki.fi

Abstract

During the last decade, social mechanisms and mechanism-based explanations have received considerable attention in the social sciences as well as in the philosophy of science. This article critically reviews the most important philosophical and social science contributions to the mechanism approach. The first part discusses the idea of mechanistic explanation from the point of view of philosophy of science and relates it to causation and to the covering-law account of explanation. The second part focuses on the ways in which the idea of mechanisms has been used in the social sciences. The final part discusses recent development in analytical sociology, covering the nature of sociological explananda, the role of theory of action in mechanistic explanation, Merton's idea of middle range theory, and the role of agent-based simulations in the development of mechanistic explanations.

Keywords: explanation, causal inference, realism, agent-based simulation, individualism, middle range theory

Running title: Causal Mechanisms

Table of Contents

Introduction

During the last decade, social mechanisms and mechanism-based explanations have received considerable attention in the social sciences as well as in the philosophy of science. Some writers have described this as a mechanism movement that is sweeping the social sciences (Norkus 2005). Although that is not an entirely correct description, the literature on mechanisms has grown rapidly and it touches upon many fundamental questions related to the aim of the social sciences.

The idea that science aims at providing mechanistic explanations of phenomena has a long history (Bechtel 2006), but only in recent decades has this idea been an object of more systematic study. In philosophy of science, mechanistic explanation has been mainly discussed in the context of biological sciences (Bechtel & Richardson 1993, Bechtel 2006,

Bechtel 2008, Craver 2007, Darden 2006, Glennan 2002, Thagard 1999, Wimsatt 2007), whereas in the social sciences the idea has been mostly discussed by social scientists themselves (Abbott 2007, Elster 1989, Elster 2007, Gross 2009, Hedström 2005, Hedström & Swedberg 1998, Manicas 2006, Mayntz 2004, Morgan & Winship 2007, Schmidt 2006, Tilly 2001, Wikström 2006). Although the basic ideas are quite similar, the discussions within philosophy and social science have proceeded mostly independently of each other. In both contexts, the development of the idea of mechanistic explanation has been partly motivated by the shortcomings of the once hegemonic covering-law account of explanation (Hempel 1965). The basic idea of mechanistic explanation is quite simple: At its core, it implies that proper explanations should detail the 'cogs and wheels' of the causal process through which the outcome to be explained was brought about.

In this article we critically review the most important philosophical and social science contributions to the mechanism approach. The article is organized as follows. In the next section, we discuss the ideas of causal mechanism and mechanistic explanation from a philosophy-of-science point of view. Thereafter, we review some of the uses of mechanistic ideas in the social sciences. Finally, we focus on analytical sociology and show how ideas about mechanistic explanations fit into a broader view of the aims and standards of sociology.

The idea of mechanistic explanation

As the entities and processes studied by different sciences are quite heterogeneous, it is difficult to propose a mechanism definition that would be both informative and cover all examples of mechanisms. Some disciplines, like cell biology (Bechtel 2006) and the

neurosciences (Craver 2007), study highly integrated systems, whereas others, like evolutionary biology and social science study more dispersed phenomena. For this reason, a characterization of a mechanism that applies to one field might not be informative when applied to another (Skipper & Millstein 2005, Kuorikoski 2009).

However, some general ideas are shared by most accepted mechanism definitions (See Table 1). First, a mechanism is identified by the kind of effect or phenomenon it produces. A mechanism is always a mechanism for something (Darden 2006, p. 273). The characterization of the effect often requires some care. For example, a standard roulette-wheel does not have different mechanisms for distributing the ball to pockets 16 and 17. Rather, the same mechanism produces all 37 outcomes.

Second, a mechanism is an irreducibly causal notion. It refers to the entities of a causal process that produces the effect of interest. Some definitions of mechanism make unnecessarily strong assumptions, however. For example, when Mahoney (2001, p. 580) characterizes causal mechanism as an unobservable entity that is sufficient to produce the outcome of interest, he makes two unnecessary commitments. First, there is nothing in the notion of a mechanism that would imply that it is by definition unobservable. Most of the mechanisms constituting an automobile's engine, for example, are quite visible when one opens up the hood (Mayntz 2004). Of course, when one appeals to mechanisms to make sense of statistical associations, one is referring to things that are not visible in the data, but this is different from it being unobservable in principle. Similarly, to require that the mechanism is sufficient for the effect is an all too strong requirement: a mechanism can involve irreducibly stochastic elements and thus affect only the probability of a given effect.

Third, the mechanism has a structure. When a mechanistic explanation 'opens the black box' it discloses this structure. It turns the black box into a transparent box and makes visible how the participating entities and their properties, activities, and relations produce the effect of interest. For this reason, the suggestion (e.g., Opp 2005) that a mechanism is just an intervening variable misses an important point. The focus on mechanisms breaks up the original explanation-seeking why-question into a series of smaller questions about the causal process: What are the participating entities, and what are their relevant properties? How are the interactions of these entities organized (both spatially and temporally)? What factors could prevent or modify the outcome? Etc.

Fourth, mechanisms form a hierarchy. While a mechanism at one level presupposes or takes for granted the existence of certain entities with characteristic properties and activities, it is expected that there are lower level mechanisms that explain them. (For a discussion of different notions of level, see Craver 2007, Ch. 5.) It is an inherent feature of the mechanistic view that the entities and mechanisms of various sciences are ultimately related to each other. The entities and activities that one theory postulates should be mechanistically explainable by another theory. Although it is sometimes (Kincaid 1996, Norkus 2005, Opp 2005) suggested that the idea of mechanistic explanation leads to an infinite regress, this is not the case. First, for a mechanism to be explanatory, it is not required that the entities, properties, and activities that it appeals to are themselves explained. The only requirement is that such entities, properties, and activities really exist; their explanation is a separate question. Second, the mechanistic account is fully compatible with the notion that there are no mechanistic explanations for fundamental (physical) processes.

Fifth, mechanisms can be combined. Two or more mechanisms can be combined to form a more complicated mechanism. For this reason, one can distinguish between atomic and molecular mechanisms (Elster 1999, 2007). Atomic mechanisms are the simplest mechanisms analyzed within a scientific field, whereas molecular mechanisms combine two or more atomic mechanisms (see also Gambetta 1998).

There is an ambiguity in the use of the notion of mechanisms. Sometimes it is used to refer to a causal process that produces the effect of interest and sometimes to a representation of the crucial elements of such a process. This should not be a cause of concern, however, since the latter presupposes the former. When one makes a claim that a certain mechanism explains some real world events, one commits to the existence of the entities, properties, activities, and relations that the description of the mechanism refers to.

In the following, we will follow a long philosophical tradition and assume that explanations are answers to questions (Hempel 1965, Salmon 1998, Woodward 2003). The why or how question one is addressing determines what the representation of the mechanism should include in order to be explanatory. Only by knowing the nature of the explanatory task at hand can one determine which details of a mechanism are relevant to include and the appropriate degree of abstraction.

Roughly, mechanisms have two kinds of *explananda*. First, they might address particular empirical facts. In such cases the description of the mechanism is often a modified adaptation and combination of more general mechanism schemes. Second, they might address stylized facts. Although the explanation of particular empirical facts is the ultimate goal of mechanistic theory development, most of the time theorists are addressing highly stylized theoretical *explananda* that do not necessarily have close resemblance to any particular empirical fact. Consider for example the small-world problem of Milgram (1967). Watts and Strogatz (1998) identified some salient and abstract features of the small-world problem, average path length and local clustering, and developed a model that could explain them. Simplified and heavily idealized facts like those identified by Watts and Strogatz call for mechanism schemes with wide application domains. By establishing how stylized facts can in principle be explained, theorists contribute to the toolbox of possible mechanistic explanations.

Following Lindley Darden (2006, p. 281), a mechanism scheme can be defined as ”a truncated abstract description of a mechanism that can be filled with more specific descriptions of component entities and activities”. In itself, a mechanism scheme provides a how-possible explanation; it tells us how the effect could in principle be produced. As it is possible that similar effects can be produced by a number of different (known or unknown) mechanisms, a crucial element in any mechanistic explanation of empirical facts is the collection of empirical evidence about the assumed entities, activities, relations, etc. The empirical evidence turns a possible mechanism to a plausible mechanism and may eventually lead to the identification of the actual mechanism. By presenting evidence in support of the assumptions of one mechanism and showing the absence of evidence for the assumptions of

competing mechanisms, we increase the plausibility of the explanatory hypothesis. What separates proper mechanistic explanations from mere mechanistic storytelling is this kind of rigorous checking of the assumptions upon which the mechanism schemes rest.

A mechanistic explanation describes the causal process selectively. It does not aim at an exhaustive account of all details but seeks to capture the crucial elements of the process by abstracting away the irrelevant details. The relevance of entities, their properties and their interactions is determined by their ability to make a relevant difference to the outcome of interest. If the presence of an entity or changes in its properties or activities truly does not make any difference to the effect to be explained, it can be ignored. This counterfactual criterion of relevance implies that mechanistic explanations involve counterfactual reasoning about possible changes and their consequences.

Mechanisms and causality

Although some authors (e.g. Glennan 1996) have attempted to define causation in terms of mechanisms, it is more prudent to keep these two things apart as most characterizations of mechanisms employ causal notions. A mechanistic definition of causation would be circular, and it would face a thorny question about the causal powers of fundamental (physical) entities. If causal relations at the fundamental level are not mechanical, the definition is false, and if they are mechanical, we end up with an infinite regress, which is regard by many as an unhappy consequence.

The metaphysics of causation is still hotly debated among philosophers, so it is a strength that the mechanistic account of explanation is compatible with a number of different accounts of causation. However, the mechanism perspective sets some important constraints. For example, the idea of productive causal activities associated with the mechanism perspective implies a commitment to the locality of causal processes: whether *a* is a cause of *b* depends on facts about spatiotemporally restricted causal process, not on what would happen in other similar situations. This means that theories that attempt to define causality in terms of regularities (like Hume's constant conjunction theory, and many probabilistic theories of causation) are not compatible with mechanistic theories. (Obviously, this does not mean that the supporters of the mechanistic perspective would have to disregard regularities as an important source of evidence about causal relations.)

Many supporters of mechanistic explanations find Wesley Salmon's (1984, 1998) conserved quantities theory of causal processes intuitively appealing. This is understandable as Salmon was one of the first philosophers who talked about the importance of a 'causal-mechanical' account of explanation. However, Salmon's theory has turned out to be a disappointment. In addition to many technical and philosophical problems associated with the theory (Woodward 2003, pp. 350-358), it seems quite ill-suited to provide a foundational account of mechanistic explanation. First, it is very difficult to see how the theory could be applied at all in biological and social sciences (or even some parts of physics) where explanations do not attempt to track spatiotemporally continuous processes like transfer of energy or momentum. Second, as Christopher Hitchcock (1995) has shown, Salmon's approach is marred by the fact that it does not include any considerations of explanatory relevance. As a consequence,

the same counterexamples that have been raised against the covering-law account can be raised against Salmon's approach (see next section for some examples).

A more natural complement to the mechanistic approach would seem to be the account of causation recently developed by James Woodward (2002, 2003; see Table 1). In Woodward's account, causal claims track relations of counterfactual dependence. They tell us what would have happened to the effect if the cause had been subject to a 'surgical' intervention that would not have affected any other part of the causal structure. One of the novelties of Woodward's theory is its account of causal generalizations in terms of invariances. According to Woodward, explanatory qualities of a generalization are determined by its ability to tell us about the counterfactual consequences of possible interventions, not by any of the properties traditionally associated with laws of nature. Woodward's account supplements the mechanistic account by providing an account of explanatory relevance, and by making sense of 'causal production'. However, it does not preempt the importance of the mechanistic perspective. The mechanism approach differs from Woodward's approach by its emphasis on the importance of opening up black boxes and making explicit the causal cogs and wheels through which effects are brought about.

The relevance of mechanisms is not limited to explanation. Especially in non-experimental contexts, mechanisms have often a crucial role in distinguishing true causal relations from spurious correlations. Mechanisms help in causal inference in two ways. The knowledge that there is a mechanism through which X influences Y supports the inference that X is a cause of Y. In addition, the absence of a plausible mechanism linking X to Y gives us a good reason to be suspicious of the relation being a causal one (Hedström 2005). The knowledge

of mechanisms also has an important role in extrapolation of causal findings from one setting to another. As Steel (2008) discusses, the assumption about the similarity of causal mechanisms is crucial for making inferences from one setting or population to another.

It is important to emphasize that mechanisms are not some sort of magic bullet for causal inference, however. The problem often is not the absence of possible mechanisms, but how to discriminate between a number of potential mechanisms. To avoid lazy mechanistic storytelling, the mechanism scheme must be made both explicit and detailed, and its assumptions must be supported by relevant empirical evidence.

While it would be too strong to say that mechanisms are necessary for causal inference (Steel 2008), a fully satisfactory social scientific *explanation* requires that the causal mechanisms are specified. The primary epistemic, in contrast to practical aim of science is to understand phenomena and this is precisely what mechanisms provide. If we take understanding to be an ability to answer what-if questions (Woodward 2003, Ylikoski 2009, Ylikoski & Kuorikoski 2009), the contribution of the mechanistic information becomes apparent. A simple causal claim tells us about counterfactual dependence: it tells us what would have happened if the cause had been different. The mechanism tells us why the counterfactual dependence holds and ties the *relata* of the counterfactual to the knowledge about entities and relations underlying it. The account of causal mechanism integrates an isolated piece of causal knowledge to a much larger body of knowledge and helps us to answer many natural follow-up questions about the conditions under which the causal dependence holds: For example, what are the necessary background conditions and what are the possible intervening factors that has to be absent for the effect to be present? In this way the mechanism expands our

ability to answer what-if questions, i.e., it deepens our understanding (Ylikoski & Kuorikoski 2009).

A comparison with the covering-law account

The mechanistic account of explanation in part has been developed as an alternative to the once dominant covering-law account of explanation of Hempel (1965) and others. The basic idea of the covering-law account is very simple: an explanation is a deductive (or statistical) argument that has a description of the *explanandum* phenomenon as a conclusion and one or more empirically validated general law statements and a set of statements describing particular facts (the initial conditions) as its premise. The core underlying idea is that explanations make the *explanandum* expected. This means that explanation and prediction are more or less the same thing; the only difference is that in the case of explanation we already know the outcome.

On the basis of the extensive debate which has taken place in philosophy of science during the last few decades (Salmon 1989, Woodward 2003), and which has not yet been fully appreciated by some social scientists (see e.g. Opp 2005), it seems fairly safe to conclude that the covering-law approach is a failure as a theory of explanation. Some of its failures are instructive, however, and it is useful to see how the mechanistic alternative avoids them.

The first set of problems concern explanatory relevance. The covering-law account has been unable to make sense of the asymmetry of explanatory relations. Although effects do not explain their causes, nothing in the covering-law account rules this out. As a consequence,

according to the covering-law account, not only can the length of a flagpole explain the length of its shadow, the length of the shadow also can explain the length of the flagpole (Salmon 1989, p. 103). In addition, the philosophical literature is full of examples that satisfy all the requirements of the covering-law account but which nevertheless are clearly non-explanatory such as that the fact that a man eats contraceptive pills is not what explains why he does not get pregnant (Salmon 1989, p. 50, Woodward 2003).

Another set of problems concerns the notion of law. The first covering-law account was developed with the expectation that the notion of a law could be given a satisfactory philosophical account, but this has not yet happened (Woodward 2003). In addition, there is an embarrassing scarcity of covering laws; in sciences like biology, psychology, and the social sciences there are very few observable empirical regularities that could be considered explanatory (Cummins 2000).

A final set of problems concerns the underlying conception of understanding. The symmetry of explanation and prediction does not seem to hold. Diagnostic effects might allow prediction, but they are not explanatory, and the sciences are full of examples of explanatory theories that do not allow precise predictions (Salmon 1989). In addition, the theory does not seem able to account for what provides understanding in explanations that do not satisfy its strict requirements. For example, the crucial explanatory work is assumed to be done by the relevant laws, but they are often unknown to people providing the explanation. While these explanations might be called elliptical, sketchy, or partial versions of the full-fledged explanation, they still seem to provide some understanding. As most explanations in science

and everyday life are like this, the covering law account seems to fail to make sense of a very fundamental aspect of explanatory cognition (Woodward 2003, pp. 159-161).

The mechanistic account is not affected by these problems. First, causality takes care of the problems of explanatory relevance: the direction of the explanation derives from the direction of causation (Salmon 1998), and the counterfactual idea of difference-making helps to solve the problems related to irrelevant explanatory factors (Woodward 2003). Similarly, the problems related to laws can be ignored. The absence of covering laws is not a problem, as they are not doing any of the explanatory work. On the contrary, in mechanistic accounts observable regularities (often called effects) are the things to be explained (Cummins 2000). Of course, mechanistic explanations still rely on causal generalizations about the properties, activities, and relations of underlying entities, but they do not have to satisfy the traditional criteria for laws. Rather, it is sufficient that they fulfill Woodward's (2003) invariance requirements. The mechanistic account also severs the close connection between explanation and prediction: it emphasizes the difference between diagnostic and explanatory reasoning and although explanatory understanding is constituted by an ability to make correct what-if inferences, this does not imply that it always provides a basis for empirical predictions. Furthermore, in a mechanistic account, the requirement that the explanation should show the cogs and wheels of the causal processes guarantees that the crucial explanatory factors are epistemically accessible for the subject. Finally, the mechanistic theory of explanation is not wedded to the idea that an explanation is an argument (Ylikoski 2005). Rather than being a relation between sentences, explanation is more properly regarded as a relation between facts (Mackie 1974, Ch. 10; Ruben 1990, Ch. 5). The deductive reconstruction of the explanation

is an advisable practice for checking it for consistency, sufficiency, and implicit premises, but explanation is not a relation between linguistic entities (Ylikoski 2007).

Social mechanisms

As mentioned in the introduction, over the years, the idea of mechanistic explanation has become increasingly influential in the social sciences. It has proved to be a useful tool for criticizing existing research practices and views about the nature of the social scientific enterprise. Many different definitions of social mechanism have been articulated and quite different methodological and theoretical conclusions have been drawn (see Table 1 for some of the most influential definitions).

Table 1. Alternative mechanism definitions.

Author	*Definition*	*References*
Bechtel & Abrahamsen	A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena	Bechtel & Abrahamsen 2005, Bechtel 2006, 2008
Bunge	A mechanism is a process in a concrete system which is capable of bringing about or preventing some change in the system.	Bunge 1997, 2004
Glennan	A mechanism for a behavior is a complex system that produces that behavior by the interaction of a number of parts, where the interactions between parts can be characterized by direct, invariant, change-relating generalizations.	Glennan 2002
Machamer, Darden & Craver	Mechanisms are entities and activities organized such that they are productive of regular changes from start to finish.	Machamer, Darden and Craver 2000, Darden 2006, Craver 2007
Elster I	Mechanisms are frequently occurring and easily recognizable causal patterns that are triggered under generally unknown conditions.	Elster 1989

Elster II	Mechanisms are frequently occurring and easily recognizable causal patterns that are triggered under generally unknown conditions.	Elster 1999
Hedström	Mechanisms consist of entities (with their properties) and the activities that these entities engage in, either by themselves or in concert with other entities. These activities bring about change, and the type of change brought about depends upon the properties of the entities and the way in which the entities are organized spatially and temporally.	Hedström 2005
Little	A causal mechanism is a series of events governed by law-like regularities that lead from the *explanans* to the *explanandum*.	Little 1991
Woodward	A model of a mechanism (i) describes an organized or structured set of parts or components, where (ii) the behavior of each component is described by a generalization that is invariant under interventions, and where (iii) the generalizations governing each component are also independently changeable, and where (iv) the representation allows us to see how, in virtue of (i), (ii) and (iii), the overall output of the mechanism will vary under manipulation of the input to each component and changes in the components themselves.	Woodward 2002

Raymond Boudon was an important early contributor to the mechanism approach. In some of his publications from the 1970s he argued for the importance of 'generative models' for explaining social outcomes (Boudon 1979), and he used computer simulations for this purpose (Boudon 1974). In a famous exchange with Robert Hauser, one of the most vocal proponents of a strict statistical approach to sociological research during that era, Boudon gave expression for one of the core ideas of the mechanism approach. He noted that statistical models of the sort advocated by Hauser are useful for many purposes, but that their usefulness as causal explanations is considerably more restricted than Hauser believed. Boudon emphasized that explanations are not achieved by simply estimating parameters of generic statistical models, but by developing generative models that explicate the mechanisms at work (see Boudon 1976, Hauser 1976).

Jon Elster has probably been the most influential advocate of mechanisms in the social sciences and his many books are full of excellent examples of mechanistic thinking in action. His idea that mechanistic explanations open up black boxes and show the cogs and wheels of the internal machinery captures quite well the mechanistic spirit. However, the various definitions of mechanisms he has provided have been a source of some confusion (see Table 1). His early view (Elster 1989), according to which a mechanism explains by providing a continuous and contiguous chain of causal or intentional links between the *explanans* and the *explanandum*, is quite in line with the general mechanistic perspective, although his choice of wording wrongly suggests that intentional explanations are not causal explanations. In contrast, his later definition: "… mechanisms are frequently occurring and easily recognizable causal patterns that are triggered under generally unknown conditions or with indeterminate consequences" (Elster 1999, p. 1) cannot serve as a general definition of mechanism. We should not require that the instances of mechanisms occur frequently or that they are easily recognizable, as sometimes finding underlying mechanisms is the hardest part of the scientific work. Similarly, although there certainly are cases where we do not know the conditions that trigger the mechanism or other factors affecting the process, it would be a mistake to incorporate these conditions in the definition of a mechanism.

In historical sociology and political science the late Charles Tilly was a prominent advocate of the mechanistic perspective. Despite his inspiring empirical work, his general discussions of mechanistic explanation (e.g., Tilly 2001) left some to be desired. One gets a strong impression that he used the notion of a mechanism as a label to refer to the kind of processes which he for other reasons was interested in. Nevertheless, he had a clear idea of how the

mechanistic perspective changes the research agenda in the historical study of politics. Contentious episodes like revolutions, nationalist mobilizations, and democratizations do not have essences. They are mostly the result of similar kinds of social processes and thereby they illustrate the fact that the same mechanisms can produce different outcomes in different circumstances. For this reason comparative studies should focus on these basic processes and mechanisms rather than large-scale episodes whose classification is mostly a product of retrospective categorization (Tilly 2001, p. 36-38).

The critical realism movement (e.g., Sayer 1992, Archer et al. 1998) has convinced many of the importance of mechanisms in social scientific explanations. The original source of their ideas about causation and scientific explanation is Rom Hárre's (1970) pioneering work, but the principal inspiration for critical realists has been Roy Bhaskar (1978, 1979). His critical realism is a tight philosophical package of ontological and epistemological views in which the idea of mechanistic explanation is only a small – but the most intuitive – part. Many social scientists rightly have been attracted by his criticism of empiricist views of causation, but other parts of the package are problematic. For example, his transcendental argumentation, his layered account of reality, and his ideas about essences and internal relations have raised philosophical suspicions and doubts about their relevance for the social sciences. It also seems that the development of critical realism has stalled: although it is advocated in numerous books and articles, for the most part its supporters have tended to repeat or reformulate often quite cryptic original statements by Bhaskar instead of engaging with recent developments in philosophy of science, or by making use of the ideas to explain important social facts.

Within philosophy of the social sciences both Daniel Little and Mario Bunge have been important spokesmen for mechanistic thinking, and more generally, for scientific realism. Little's textbook (Little 1991) and articles (Little 1998) shows in an accessible way how thinking in terms of causal mechanisms helps to resolve various methodological problems which characterize other approaches to the social sciences. His definition of mechanism as "a series of events governed by law-like regularities that lead from the *explanans* to the *explanandum*" (Little 1991, p. 15; Table 1) is somewhat problematic because of its reference to law-like regularities, but it recognizes the importance of causal generalizations in mechanisms and Little strongly emphasizes the important role of microfoundations in mechanistic explanations.

Mario Bunge's (1997, 2004) account of mechanistic explanation is part of his more general systemistic philosophy of science. For him a system is "*a complex object whose parts or components are held* together by bonds of some kind." (Bunge 2004, p. 188) and a mechanism is a set of processes that are typical for material systems of certain kinds. These processes characterize the relations and interactions between the system's parts, its structure and its environment. Just like in case of Bhaskar, Bunge's basic ideas about mechanistic explanations are pretty much in line with the account outlined above but it carries an extra philosophical baggage that we may want to avoid. For example, it is difficult to say whether Bunge is saying something profound or something very trivial when he states that according to systemism "everything in the universe is, was, or will be a system or a component of one" (Bunge 2005, p. 190). It might be that all valid social scientific findings could be translated to Bunge's abstract systemic language, but it is questionable whether it provides much guidance for the construction and evaluation of mechanistic explanations.

In political science, causal mechanisms have had an important role in debates concerning research methodologies and causal inferences (e.g., Mahoney 2001, George & Bennett 2005, Gerring 2008, Box-Steffensmeier et al. 2008). Mechanistic thinking often has been presented there as an alternative to statistical methodologies and as providing foundations for case studies and other small-N studies. As George and Bennett (2005) show, the idea of process tracing is useful for both development and testing of mechanistic explanations, and it is not the case that only statistical evidence is relevant for the assessment of causal claims. In general, the debate in political science has been quite similar to that in sociology (see Sørensen 1998): the advocates of mechanisms have criticized the simplistic empiricist uses of statistical methodology for ignoring the importance of causal process assumptions in causal inference.

In criminology, the mechanism approach is particularly associated with the work of P-O Wikström. In opposition to the dominant risk-factor paradigm within quantitative criminology, Wikström's so-called situational action theory is a general theory of criminal behavior that specifies various mechanisms through which individual and situational factors jointly influence the probability of individuals' engaging in criminal activities (see Wikström, 2006, Wikström & Treiber, 2009). The situational action theory is particularly concerned with moral rules and the mechanisms that explain why some people follow and others breach moral rules as defined in law.

In addition to the above mentioned scholars, the volume *Social Mechanisms* (1998) edited by Peter Hedström and Richard Swedberg collected many social scientists who have been important in making the mechanistic perspective visible. The chapters by Raymond, Boudon, Jon Elster, Diego Gambetta, Peter Hedström, Gudmund Hernes, Thomas Schelling, and Aage Sørensen represent what has become known as the analytical approach to social mechanisms, and we will discuss this approach in more detail below. It would be a mistake to take this collection of essays as a definite formulation of this approach, however. There are important differences between the authors and the approach is still developing (see in particular Hedström 2005 and the various chapters in Hedström and Bearman 2009). In the following we will comment on some of these developments as we try to articulate a more systematic account of the implications of mechanistic ideas for the social sciences.

From mechanisms to analytical sociology

While the idea of mechanistic explanation helps social scientists to avoid some philosophical pitfalls, the mere adoption of mechanism talk will not fulfill its promise. Much depends on the way in which mechanistic ideas are put to use; otherwise we will end up with mere mechanistic storytelling that lacks both theoretical rigor and empirical relevance. What is needed, is a broader vision of sociology and the so-called analytical sociology movement has attempted to articulate such a vision (see Hedström & Bearman 2009). The identity of this reform movement in sociology is not based on a common object of study, a shared historical tradition in sociological theory, or an empirical research method, but on some general ideas about what good social science is all about. At the most fundamental level, analytical sociology emphasizes the importance of the intellectual virtues of precision and clarity. As

Elster (2007 p. 455) expressed it: "What one might call the 'analytical turn' in social science … rest … on a near-obsessive concern with clarity and explicitness." What he means is that social scientific theories should be presented with such precision and clarity that it is possible to clearly distinguish the causal mechanisms upon which they are founded. A fancy theoretical jargon never can be a substitute for precise and explicit theorizing.

The importance of the intellectual virtues of clarity and explicitness derive from our susceptibility to the illusion of depth of understanding: we tend to overestimate the detail, coherence, and depth of our understanding (Keil 2003, Ylikoski 2009). With its implicit and vague explanatory standards, its largely verbal mode of theorizing, and its highly complex object of study, social theory is probably a primary example of this failure of metacognition. One of the key ways to fight this illusion is by making explanations explicit: clearly and precisely articulated accounts of mechanisms can be subjected to a piecemeal scrutiny and their implications can be assessed more accurately.

Analytical sociology emphasizes the importance of closely integrating theoretical and empirical work, but it is not closely tied to any specific research methodology. The crucial question is what kind of access a certain piece of evidence provides to the causal process of interest, not whether the method being used is quantitative or qualitative, experimental or non-experimental.

While the principles of analytical sociology are still developing, some central characteristics of the approach can be discerned. In the following we will discuss four issues that are closely

related to the idea of mechanistic explanation: (1) the typical sociological *explananda*; (2) the role of a theory of action; (3) mechanistic theories as middle-range theories; and (4) the role of agent-based modeling in analytical sociology.

The sociological *explananda*

For sociology, the most important *explananda* are social phenomena. Although sociology shares with psychology and social psychology the interest in explaining properties and behaviors of individuals, its key challenge is to account for collective phenomena that are not definable by reference to any single member of the collectivity. Among such properties are (Hedström 2005, p. 67):

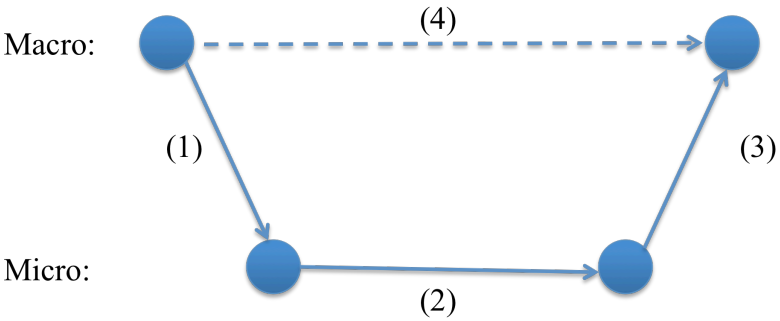
(1) Typical actions, beliefs, or desires among the members of society or a collectivity.

(2) Distributions and aggregate patterns such as spatial distributions and inequalities.

(3) Topologies of networks that describe relationships between members of a collectivity.

(4) Informal rules or social norms that constrain the actions of the members of a collectivity.

From the point of view of mechanistic explanation, the basic entities that explain these kinds of phenomena are intentional human agents and their relations. The so-called Coleman's boat illustrates clearly this aspect of the mechanistic approach (see Figure 1).



(1): Situational mechanisms
(2): Action-formation mechanisms
(3): Transformational mechanisms

Figure 1. A typology of social mechanisms (Hedström & Swedberg 1998)

A basic point of the mechanistic perspective is that explanations that simply relate macro properties to each other (arrow 4) are unsatisfactory. These explanations do not specify the causal mechanisms by which macro properties are related to each other. Deeper explanatory understanding requires opening up the black box and finding the causal mechanisms that have generated the macro-level observation. Rather than analyzing relationships between phenomena exclusively on the macro level, one should identify the *situational mechanisms* by which social structures constrain individuals' action and cultural environments shape their desires and beliefs (arrow 1), describe the *action-formation mechanisms* according to which individuals choose how to act (arrow 2), and single out *transformational mechanisms* by which individuals, through their actions and interactions, generate various intended and unintended

social outcomes (arrow 3). Only by understanding the whole chain of situational, action-formation and transformational mechanisms, have we made sense of the observed macro-level relationship (Coleman 1990, Hedström & Swedberg 1998).

This emphasis on detailing mechanisms implies that explanations should refer to individuals, their relationships, and their actions. It is important to recognize that the basic building blocks of social explanations are not mutually independent actions performed by atomistic individuals, however. Rather, the individuals' actions typically are oriented towards others and their relations to others therefore become central when it comes to explaining why they do what they do. In addition, social relations are central for explaining why, acting as they do, individuals bring about the social outcomes they do. Consider, as an example, the case of vacancy chains as analyzed by White (1970). A retirement, motivated perhaps by a desire for a more leisurely life, creates an opportunity for others, i.e., a job vacancy waiting to be filled by a new occupant. The vacancy generated by the retirement is filled by another individual whose reason for taking the job, perhaps, is to attain more status, or a higher salary, or just a change in venue, but this creates another vacancy in this persons old job, and in this way vacancy chains create social interdependencies which are important for explaining mobility. Individuals' desires – for retirement, for promotion, for status, or for a change in venue – motivate the system. Without such orientations, people may not move. But, explanatory understanding is only achieved by recognizing that actions take place in relational structures that in this case channel mobility opportunities and thereby explain why we observe what we observe (Hedström & Bearman 2009).

The above argument does not imply any commitment to the doctrine of methodological individualism. Most formulations of methodological individualism are much stronger than the mechanistic perspective requires (for a review of the literature on methodological individualism, see Udéhn 2001). Some form of structural individualism (Coleman 1990, Udéhn 2001) is sufficient for the purposes of mechanistic explanations of social phenomena. Structural individualism is a doctrine according to which all social facts, their structure and change, are in principle explicable in terms of individuals, their properties, actions, and relations to one another. Structural individualism differs from most formulations of methodological individualism by emphasizing the explanatory importance of relations and relational structures. It does not require that all explanatory facts are facts about individual agents in the strict sense: facts about topologies of social networks, about distributions of beliefs, resources or opportunities, and about institutional or informal rules and norms can have a significant role in mechanistic explanations. For example, the transformational mechanisms (arrow 3 in Figure 1) are not based on simple aggregation of individual actions, as in many formulations of methodological individualism, but depend on structural arrangements that cannot be defined by reference to atomistic attributes of individual agents. Structural individualism is far from a trivial doctrine, however. It is incompatible with radical holism and structuralism, since it requires that the social pattern to be explained should be understood in terms of interactions of individual agents.

Theories of action

As intentional action plays central role in social mechanisms, it raises the question concerning the role of a theory of action. The idea of social mechanism in itself does not tell

us how to conceptualize human action. Rather than relying on some preconceived ideas about human motivation or cognitive processing, the mechanistic perspective suggests that our account of human agency should be based on findings and theories of psychological and cognitive sciences.

The explanatory agenda of mechanistic social science has some methodological implications, however. As understanding of complex phenomena is only possible in a piecemeal way, we must abstract away from many details of human mental life. Only those aspects of cognition that are relevant for the explanatory task at hand should be included in the explanation, and the explanatory task thus determines how rich the psychological assumptions must be. For example, it is natural to employ a richer psychological apparatus when articulating action formation mechanisms than when modeling transformational mechanisms.

So although the mechanistic approach emphasizes the importance of intentional action in the explanation of social phenomena, it does not subscribe to an axiomatic vision according to which a specific action theory should be used for all purposes. For many social scientific purposes, a relative simple desire-belief-opportunity model (DBO-theory) will be sufficient (Hedström 2005). For other purposes a pragmatist theory of action (Gross 2009) could be fruitful. Habits, routines and various unconscious cognitive processes are important parts of modern naturalistic accounts of human cognition (Wilson 2002, Hassin et al. 2005) and sociology needs to take such factors into account. Their incorporation into detailed sociological models requires much additional work, however.

The idea of social mechanisms is quite often associated with rational choice theory (Abbott 2007, Gross 2009). However, from a philosophical and general sociological point of view the connection between the two is quite weak. There is nothing in the idea of a mechanistic explanation that would require the explanation to be articulated in terms of rational choice theory. On the contrary, the requirement that mechanistic explanations cite actual causes of the phenomenon to be explained often makes rational choice explanations unacceptable as they are built upon implausible psychological and sociological assumptions. Empirically false assumptions about human motivation, cognitive processes, access to information, or social relations cannot bear the explanatory burden in a mechanistic explanation. It is not enough that the model ”saves the phenomena”, it should represent the essential features of the actual causal structure that produces the observed phenomena. Rather than accepting the instrumentalist ‘as if’ attitude displayed by many economists, the mechanistic approach requires that one should strive for theoretical assumptions that are both empirically valid and compatible with the results of other disciplines. In order for unrealistic psychological and social assumptions to be acceptable, they must be simplifying idealizations that help the modeling but do not affect the central explanatory relationships in any crucial manner. This rarely is the case of rational-choice theory.

Mechanisms and middle-range theories

The idea of causal mechanisms is related to broader ideas about the growth and organization of scientific knowledge. According to an old, but still influential empiricist view, general scientific knowledge consists of empirical generalizations and more abstract theoretical principles from which these generalizations can (ideally) be deduced. This vision is

challenged by the mechanistic account of knowledge with its emphasis on scientific knowledge as embedded in mechanism schemes and not in empirical generalizations.

According to this view, scientific knowledge expands by adding items to or improving upon items already present in the toolbox of possible causal mechanisms. Understanding accumulates as the knowledge of mechanisms gets more detailed and the number of known mechanisms increases. This vision of knowledge does not require that mechanisms are ultimately organized into a grand unified theory. It is only required that the accounts of mechanisms provided by different disciplines are mutually compatible and that they form an integrated web in which mechanisms at lower levels of organization explain the mechanisms that higher-level disciplines take for granted. For example, psychology explains (and corrects) the assumptions that social scientists make about human cognitive processes.

The mechanism idea is important in a highly specialized and fragmented discipline like sociology. Although empirical data, research methods, and substantial theories differ from one subfield of sociology to another, the general ideas about possible causal mechanisms are something these fields could share and thereby benefit from each other's work. In this vision, sociological theory provides a set of explanatory tools that can be employed and adapted to particular situations and explanatory tasks. The mechanisms are (semi-) general in the sense that most of them are not limited to any particular application. For example, the same type of mechanism can be used for (partially) explaining residential segregation (Bruch & Mare 2006) and success in cultural markets (Salganik & Watts 2009).

This mechanistic vision of knowledge has much in common with Robert K. Merton's idea of sociological theories of the middle range (Merton 1968, Hedström & Udehn 2009). According to Merton, theories of the middle range are

> … theories that lie between the minor but necessary working hypotheses that evolve in abundance during day-to-day research and the all-inclusive systematic efforts to develop a unified theory that will explain all the observed uniformities of social behavior, social organization and social change. (Merton 1968, p. 39).

An ideal theory of the middle range is clear, precise, and simple. It does not bore the reader by attempting to describe the causal process in all its detail; instead it seeks to highlight the heart of the story by isolating a few explanatory factors that explain important but delimited aspects of the outcomes to be explained. A theory of the middle range can be used for partially explaining a range of different phenomena, but it makes no pretense of being able to explain all social phenomena, and it is not founded upon any form of extreme reductionism.

Merton's account of self-fulfilling prophecies provides a clear example of a social mechanism. The focus of the analysis is on the process through which an initially false belief evokes behavior that eventually makes the false belief come true and its basic idea is captured by the so-called Thomas Theorem: "If men define situations as real, they are real in their consequences" (Merton 1968, p. 475). Merton shows how an endogenous and self-reinforcing process can bring about a collective outcome that is unintended by all the individuals involved, and the canonical example he uses is a run on a bank. Once a rumor of insolvency gets started, some depositors are likely to withdraw their savings, acting on the

principle that it is better to be safe than sorry. Their withdrawals strengthen the beliefs of others that the bank is in financial difficulties, partly because the withdrawals may actually hurt the financial standing of the bank, but more importantly because the act of withdrawal in itself signals to others that something might be wrong with the bank. This produces even more withdrawals, which further strengthens the belief, and so on. By this mechanism, even an initially sound bank may go bankrupt if enough depositors withdraw their money in the (initially) false belief that the bank is insolvent.

The mechanism sketched by Merton is an important part of the theoretical toolbox of sociology, and it is an ideal example of what the analytically oriented mechanism approach is all about. The basic structure of Merton's mechanism is as depicted in Figure 2 (Hedström & Udéhn 2009).

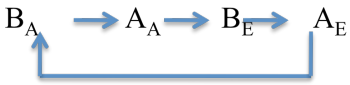


Figure 2. The logic of self-fulfilling prophecies. The subscripts identify the individuals in question, Ego and Alter(s), and B and A represent the beliefs and actions of the relevant individuals.

The self-fulfilling prophecy has the following general argument pattern: The beliefs of one or

several Alters are such that they decide to perform a certain action; their actions influence Ego's belief about the value of performing the act is such a way that Ego also decides to act; and Ego's action in turn strengthen the beliefs of others in the value of performing the act. This type of self-reinforcing and belief-centered cycle then is repeated and can make an initially false belief come true. This mechanism is in principle quite general and it can be used to explain a range of different types of phenomena (Biggs 2009). The development of this mechanism scheme also provides a good example of growth of mechanistic knowledge. The original sketch of a mechanism for a limited number of phenomena has grown to a whole family of models with a wide range of applications.

Agent-based modeling

In their effort to develop clear and precise explanations of social phenomena, agent-based computer simulation has emerged as an essential tool for analytical sociology (see Macy & Flache 2009). Much of the development of mechanistic knowledge consists of developing how-possibly explanation schemes. These schemes are not intended to directly explain any particular empirical facts, but to provide general understanding of how things could work. Given the limitations of experimental methods and the complexity of social phenomena, computer simulations are important for this kind of endeavor. They allow systematic exploration of consequences of modeling assumptions and they make it possible to model much more complex phenomena than was possible earlier.

The promise of agent-based simulation is based on the fact that the dynamics observed at the social level typically are complex and hard to understand, but often its possible to describe the basic cogs and wheels of these social processes with rather simple models. Thomas Schelling's (1978) famous segregation model is a good illustration of this. Schelling demonstrated that even if individuals do not have strong preferences against mixed neighborhoods, segregation often arises as an unintended consequence. He also showed that small and seemingly trivial differences at the micro-level can make huge differences at the macro-level. More generally, Schelling's simple and stylized example demonstrates that macro-level outcomes and relationships tell us very little about the macro-level outcomes and relationships we observe. Only by explicitly considering the micro-level actions and relations and how they unfold over time, can macro-level outcomes be explained. This basic insight is at the heart of analytical sociology: to understand collective dynamics, we must study the collectivity as a whole, but we must not study it as a collective entity.

One important feature of agent-based simulations is that they do not impose any a priori constraints on the mechanisms assumed to be operating. Unlike rational-choice theory, agent-based modeling is not based on any specific theory of action or interaction. It is a methodology for deriving the social outcomes that groups of interacting actors are likely to bring about whatever the action logics or interaction structures may be. The simulation allows the researcher to see how the phenomenon to be explained is generated and how changes in action logics or relational structures are likely to change the social outcome.

As emphasized by Manzo (2007), agent-based modeling is not only a useful tool for analytical sociology, there is a natural affinity between the components of mechanism-based explanations and agent-based simulation models:

> When we *write* a set of computational algorithms (the program), formalising the generative hypotheses […], what we are doing is hypothesising a series of generative mechanisms. When we *execute* the program […] we engender the process deriving from the set of posited generative mechanisms. With the technical distinction between program "writing," "compilation," and "execution" it becomes clear that a "process" is nothing more than the dynamic aspect of one (or several) mechanism(s): it is what the mechanism can trigger. (Manzo 2007, pp. 5-6)

That is to say, agent-based simulation methods provide a technical 'infrastructure' tightly coupled to the theoretical agenda of analytical sociology. The simulation allows us to see how the phenomenon to be explained is generated and how changes in the assumptions change the outcome. If explanatory understanding consists of knowledge of dependencies and ability to make correct what-if inferences on the basis of this knowledge (Ylikoski & Kuorikoski 2009), it is clear that computer simulations increase our ability to see the dependencies that follow from different sets of assumptions and that it increases the reliability of our inferences. As so eloquently illustrated by Schelling in his segregation analysis, the latter is especially important as our unaided inferences are quite error-prone.

Agent-based simulations should not only be regarded as a tool for theoretical exploration. Empirically calibrated agent-based models make it possible to integrate theoretical ideas with

the results of empirical research. Simulation studies that aim to account for empirical observations employ a *generative research strategy* (Hedström & Bearman 2009, Epstein 2006). The basic structure of this research strategy is the following:

1. Start with a clearly delineated social fact that is to be explained.
2. Formulate different hypotheses about relevant micro-level mechanisms.
3. Translate the theoretical hypotheses into computational models.
4. Simulate the models to derive the type of social facts that each micro-level mechanism brings about.
5. Compare the social facts generated by each model with the actually observed outcomes.

One recent study that nicely illustrates this approach is the study of the sexual and romantic networks of adolescents undertaken by Bearman, Moody, and Stovel (2004). The context of their study was a single high school of roughly thousand students in the Midwestern United States. The macrostructure they sought to explain was the surprising discovery that the sexual and romantic network structure of the students resembled a spanning tree. They identified different micro-level processes that potentially could explain this macro-level pattern and used simulations to derive what the macro structure would look like if a particular micro process were at work. By performing different simulations they came to the conclusion that the spanning-tree structure most likely was the result of a social norm that prohibited cycles of length four. From a boy's point of view, this implies that he should not form a partnership with his prior girlfriend's current boyfriend's prior girlfriend. Because of lack of data, they could not examine why students obeyed this norm, but Bearman et al. convincingly argue that the status loss associated

with such relationships is a major factor as is the so-called yuck factor that arises from too much closeness.

Conclusions

During the last decade, mechanisms and mechanism-based explanations have attracted a great deal of attention in the social sciences as well as in the philosophy of science. In this article we have reviewed the most important mechanism-related literature and discussed how ideas about mechanistic explanations fit into a broader view of the aims and standards of sociology.

Much has been accomplished during these years and the literature on mechanisms continues to be intellectually exciting and to focus on issues of considerable relevance for the future development of the discipline. In these concluding remarks we reconnect to a theme that we have touched upon above: the dangers of lazy mechanistic story telling.

If the mechanism approach is to deliver on its promise, it must be taken seriously. Continuing along the same paths as before and simply interpreting one's research findings in mechanism terms will not suffice. A proper mechanistic explanation is not an *ad hoc* addition that can be added to any set of empirical observations at will. Underlying the mechanistic approach is a commitment to realism and an opposition to any form of instrumentalism: the explanations should reflect the causal processes actually responsible for the observations.

The mechanistic approach requires both stringency in theoretical practice and imagination in the design of empirical research. There is still much work to be done in the translation of the general principles of the mechanism approach into concrete research practices, but we find the recent developments in analytical sociology promising in this regard. The key to future progress is the development of good exemplars of analytical sociology rather than, say, engaging in debates about the proper definition of mechanism.

One of the basic premises of analytical sociology is that proper understanding of collective processes requires that the researchers pay attention to entities that mechanisms are made of (the agents, their properties, actions, and relations) rather than treating them as opaque entities. This 'dogma' of analytical sociology means that theoretical models as well as empirical research must focus on actions and relations and how they unfold over time. This basic principle has far-reaching consequences for the type of empirical data that needs to be used as well as for the type of theoretical work that is needed.

As far as data and empirical research is concerned, longitudinal data with relational information is essential since the basic entities of social mechanisms are actors, their actions and relations. Similarly the mechanism approach has considerable consequences for how theoretical work ought to be conducted. As discussed above, in order to explain macro-level change rigorous theorizing is needed that explicitly considers the micro-level mechanisms at work and the dynamic processes that they give rise to. We are convinced that agent-based computer simulations will become a basic tool of social theory in general and mechanism-based theory in particular. The flexibility of simulation models allows theoretical ideas to be translated into generative models without distorting the original ideas. Although much work

remains to be done, it also seems clear to us that a further development of the mechanism approach calls for empirically calibrated simulation models of the sort discussed above. Empirically calibrated simulation models seem to be the natural way of bridging the unfortunate gap between theoretical and empirical work that has hampered the discipline for so many years.

Literature cited

Abbott A 2007. Mechanisms and Relations. *Sociologica* 2: 1-22

Archer, MS, Bhaskar R, Collier A, Lawson T, Norrie A. eds. 1998. *Critical Realism: Essential Readings*. London: Routledge

Bearman P, Moody J, Stovel K 2004. Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks. *American Journal of Sociology* 110: 44–91

Bechtel W 2006. *Discovering Cell Mechanisms. The Creation of Modern Cell Biology*. New York: Cambridge Univ. Press

Bechtel W 2008. *Mental Mechanism. Philosophical Perspectives on Cognitive Neuroscience*. London: Routledge

Bechtel W, Abrahamsen A 2005. Explanation: A Mechanist Alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: 421–441

Bechtel W, Richardson RC 1993. *Discovering Complexity. Decomposition and Localization as Strategies in Scientific Research*. Princeton: Princeton Univ. Press

Bhaskar R 1978. *A Realist Theory of Science.* 2nd ed. Brighton: The Harvester Press

Bhaskar R 1979. *The Possibility of Naturalism*. Brighton: The Harvester Press

Biggs M 2009. Self-Fulfilling Prophecies. See Hedström & Bearman eds. 2009, pp. 294-314

Boudon R 1974. *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. New York: Wiley

Boudon R 1976. Comment on Hauser's "Review of Education, Opportunity, and Social Inequality". *American Journal of Sociology* 81: 1175-87

Boudon R 1979. Generating Models as a Research Strategy. In *Qualitative and Quantitative Social Research: Papers in Honor of Paul F. Lazarsfeld*, ed. Rossi PH, pp. 51-64. New York: The Free Press

Box-Steffensmeier JM, Brady HE, Collier D eds. 2008. *The Oxford Handbook of Political Methodology*. Oxford: Oxford Univ. Press

Bruch EE, Mare RD 2006. Neighborhood Choice and Neighborhood Change. *American Journal of Sociology* 112: 667–709

Bunge M 1997. Mechanism and Explanation. *Philosophy of the Social Sciences* 27: 410-65.

Bunge M 2004. How Does It Work? The Search for Explanatory Mechanisms. *Philosophy of the Social Sciences* 34: 182-210

Coleman JS 1990. *Foundations of Social Theory*. Cambridge, MA: The Belknap Press

Craver C 2007. *Explaining the Brain. Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Oxford Univ. Press

Cummins R 2000. ”How Does It Work?” versus “What Are the Laws?”: Two Conceptions of Psychological Explanation. In *Explanation and Cognition,* eds. Keil & Wilson, pp. 117-44. Cambridge, MA: MIT Press

Darden L 2006. *Reasoning in Biological Discoveries. Essays on Mechanisms, Interfield Relations, and Anomaly Resolution*. Cambridge, UK: Cambridge Univ. Press

Elster J 1989. *Nuts and Bolts for the Social Sciences*. Cambridge, UK: Cambridge Univ. Press

Elster J 1999. *Alchemies of the Mind: Rationality and the Emotions*. Cambridge, UK: Cambridge Univ. Press

Elster J 2007. *Explaining Social Behavior. More Nuts and Bolts for the Social Science*. Cambridge, UK: Cambridge Univ. Press

Epstein J 2006. *Generative Social Science. Studies in Agent-Based Computational Modeling*. Princeton, NJ: Princeton Univ. Press

Gambetta, D. 1998. Concatenations of Mechanisms. See Hedström & Swedberg 1998, pp. 102-24

George AL, Bennett A 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press

Gerring J 2008. The Mechanismic Worldview: Thinking Inside the Box. *British Journal of Political Science* 38: 161-79

Glennan S 1996. Mechanisms and The Nature of Causation. *Erkenntnis* 44: 49-71

Glennan S 2002. Rethinking Mechanistic Explanation. *Philosophy of Science* 69: S342-S353

Gross N 2009. A Pragmatist Theory of Social Mechanisms. *American Sociological Review* 74: 358-379

Harré R 1970. *The Principles of Scientific Thinking*. London: Macmillan

Hassim RR, Uleman JS & Bargh JA 2005. *The New Unconscious*. Oxford: Oxford Univ. Press

Hauser RM 1976. Review Essay: On Boudon's Model of Social Mobility. *American Journal of Sociology* 81: 911-28

Hedström P 2005. *Dissecting the Social. On the Principles of Analytical Sociology*. Cambridge, UK: Cambridge Univ. Press

Hedström P, Bearman P eds. 2009. *The Oxford Handbook of Analytical Sociology*. Oxford: Oxford Univ. Press

Hedström P, Bearman P 2009. What is analytical sociology all about? An introductory essay. See Hedström & Bearman eds. 2009, pp. 3-24

Hedström P, Swedberg R 1998. Social mechanisms: An introductory essay. See Hedström & Swedberg eds. 1998, pp. 1-31

Hedström P, Swedberg R eds. 1998. *Social Mechanisms. An Analytical Approach to Social Theory*. Cambridge, UK: Cambridge Univ. Press

Hedström P, Udehn L 2009. Analytical sociology and theories of the middle range. See Hedström & Bearman eds. 2009, pp. 25-49

Hempel C 1965. *Aspects of Scientific Explanation*. New York: The Free Press

Hitchcock C 1995. Discussion: Salmon on Explanatory Relevance. *Philosophy of Science* 62: 304-20

Keil FC 2003. Folkscience: coarse interpretations of a complex reality. *Trends in Cognitive Sciences* 7: 368-73

Kincaid H 1996. *Philosophical Foundations of the Social Sciences. Analyzing Controversies in Social Research*. Cambridge, UK: Cambridge Univ. Press

Kuorikoski J 2009. Two Concepts of Mechanism: Componential Causal System and Abstract Form of Interaction. *International Studies in the Philosophy of Science* 23: 143-60

Little D 1991. *Varieties of Social Explanation: An Introduction to the Philosophy of Social Science*. Boulder, CO: Westview Press

Little D 1998. *Microfoundations, Method, and Causation. On the Philosophy of the Social Science*. New Brunswick: Transaction Publishers

Machamer P, Darden L, Craver C 2000. Thinking About Mechanisms. *Philosophy of Science* 67: 1–25

Mackie JL 1974. *The Cement of the Universe*. Oxford: Oxford Univ. Press

Macy M, Flache A 2009. Social dynamics from the bottom up: Agent-based models of social interaction. See Hedström & Bearman eds. 2009, pp. 245-68

Mahoney J 2001. Beyond Correlational Analysis: Recent Innovations in Theory and Method. *Sociological Forum* 16 (3): 575-93

Manicas P 2006. *A Realist Philosophy of Social Science. Explanation and Understanding*. Cambridge, UK: Cambridge Univ. Press

Manzo G 2007. Comment on Andrew Abbott. *Sociologica* 2: 1-8

Mayntz R 2004. Mechanisms in the Analysis of Social Macro-Phenomena. *Philosophy of the Social Sciences* 34: 237-59

Merton RK 1968. *Social Theory and Social Structure*. New York: The Free Press

Milgram S 1967. The Small World Problem. *Psychology Today* 2: 60-7

Morgan SL, Winship C 2007. *Counterfactuals and Causal Inference. Methods and Principles for Social Research*. New York: Cambridge Univ. Press

Norkus Z 2005. Mechanisms as miracle makers? The rise and inconsistencies of the "mechanismic approach" in social science and history. *History and Theory* 44 (3): 348-72

Opp K-D 2005. Explanations by mechanisms in the social sciences. Problems, advantages and alternatives. *Mind & Society* 4: 163-78

Ruben D-H 1990. *Explaining Explanation*. London: Routledge

Salganik MJ, Watts DJ 2009. Social influence: The puzzling nature of success in cultural markets. See Hedström & Bearman eds. 2009, pp. 315-41

Salmon W 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton Univ. Press

Salmon W 1989. *Four Decades of Scientific Explanation*. Minneapolis: Univ. of Minnesota Press

Salmon W 1998. *Causality and Explanation*. Oxford: Oxford Univ. Press

Sayer A 1992. *Method in Social Science. A Realist Approach.* 2nd ed. London: Routledge

Schelling TC 1978. *Micromotives and Macrobehavior*. New York: W.W. Norton

Schmidt M 2006. *Die Logik mechanismischer Erklärungen*. Wiesbaden: VS Verlag

Skipper RA, Millstein RL 2005. Thinking about evolutionary mechanisms: natural selection. *Studies in History and Philosophy of Biology and Biomedical Sciences* 36: 327-47

Sørensen AB 1998. Theoretical mechanisms and the empirical study of social processes. See Hedström & Swedberg eds. 1998, pp. 238-66

Steel D 2008. *Across the Boundaries. Extrapolation in Biology and Social Sciences*. Oxford: Oxford Univ. Press

Thagard P 1999. *How Scientistist Explain Disease*. Princeton, NJ: Princeton Univ. Press

Tilly C 2001. Mechanisms in Political Processes. *Annual Review of Political Science* 4: 21-41

Udéhn L 2001. *Methodological Individualism: Background, History and Meaning*. London: Routledge

Watts DJ, Strogatz SH 1998. Collective Dynamics of 'Small World' Networks. *Nature* 393: 440-42

White HC 1970. *Chains of Oppurtunity: System Models of Mobility in Organizations.* Cambridge: Harvard Univ. Press

Wikström P-O 2006. Individuals, settings and acts of crime. Situational mechanisms and the explanation of crime. In *The explanation of crime: Context, mechanisms and development*, eds. Wikström P-O, Sampson RJ, pp. 61-107. Cambridge, UK: Cambridge Univ. Press

Wikström, P-O, Treiber K 2009. Violence as situational action. *International Journal of Conflict and Violence* 3: 75–96

Wilson TD 2002. *Strangers to Ourselves. Discovering the Adaptive Unconscious*. Cambridge, MA: The Belknap Press

Wimsatt W 2007. *Re-Engineering Philosophy for Limited Beings. Piecewise Approximations to Reality*. Cambridge, MA: Harvard Univ. Press

Woodward J 2002. What Is a Mechanism? A Counterfactual Account. *Philosophy of Science* 69: S366-S377

Woodward J 2003. *Making Things Happen. A Theory of Causal Explanation*. Oxford: Oxford Univ. Press

Ylikoski P 2005. The Third Dogma Revisited. *Foundations of Science* 10: 395-419

Ylikoski P 2007. The Idea of Contrastive *Explanandum*. In *Rethinking Explanation,* eds. J Persson, P Ylikoski, pp. 27-42. Dordrecht: Springer

Ylikoski P 2009. Illusions in Scientific Understanding. In *Scientific Understanding: Philosophical Perspectives,* eds. H De Regt, S Leonelli, K Eigner, pp. 100-19. Pittsburgh: Pittsburgh Univ. Press

Ylikoski P, Kuorikoski J 2009. Dissecting Explanatory Power. *Philosophical Studies*, In press