Published in *The Routledge Handbook of Mechanisms and Mechanical Philosophy*, pp. 401-412, edited by Stuart Glennan and Phyllis Illari.

Social Mechanisms

Petri Ylikoski

Institute for Analytical Sociology, Linköping University (Sweden)

Department of Social Research, University of Helsinki (Finland)

Social mechanisms and mechanism-based explanation have attracted considerable attention in the social sciences and the philosophy of science during the past two decades. The idea of mechanistic explanation has proved to be a useful tool for criticizing existing research practices and meta-theoretical views on the nature of the social-scientific enterprise. Many definitions of social mechanisms have been articulated, and have been used to support a wide variety of methodological and theoretical claims. It is impossible to cover all of these in one chapter, so I will merely highlight some of the most prominent and philosophically interesting ideas.

1 Mechanisms in the social sciences

As in other sciences, mechanism as a notion belongs to the everyday causal vocabulary of many social scientists. In this context the word ‘mechanism’ could refer to a cause, a causal pathway, or an explanation without explicit theorizing about the nature of mechanisms. This casual and occasional mechanistic way of talking is probably as old as the social sciences. Equally old are certain negative connotations of the word ‘mechanical’ implying simple, rigid, and reductionist. Being part of everyday causal vocabulary explains much of the intuitive appeal of mechanistic language, although the negative connotations have made some social scientists suspicious of the mechanistic turn.

Theorizing about mechanisms has multiple origins in the social sciences. Among the most prominent early sources is Rom Harré's (1970) philosophy of science. Although his later work has been highly influential in social psychology, the biggest impact of his philosophy of science was through the so-called critical realist movement. The key thinker in this movement is Roy Bhaskar (1978, 1979), in whose philosophy Harré´s ideas about causation and mechanisms have a central role. Despite Bhaskar's transcendental argumentation, his layered account of reality, and his ideas about essences and internal relations that have raised philosophical suspicions and doubts about their relevance to the social sciences, critical realism is still one of the most influential meta-theoretical movements in the social sciences (Lawson 2003, Elder-Vass 2011). Philosophers of mechanisms might be unfamiliar with critical realism, but in the view of many social scientists the idea of causal mechanisms is strongly associated with this movement.

Another influential early advocate of mechanism-based thinking is Jon Elster, his many books containing excellent examples of mechanistic thinking in action. His early definition (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 Elster's mechanisms tend to be psychological rather than social, his work has inspired many social scientists to open up black boxes and show the cogs and wheels of the internal machinery of social processes. Elster has also strongly associated the mechanistic attitude with the intellectual virtues of clarity and precision: he sees mechanism-based theorizing as clearheaded causal thinking about social processes and for a large group of social scientists this is the core of the mechanistic perspective. Note that there are few connections between social scientists inspired by Bhaskar's critical realism and those supporting Elster's approach. The idea of mechanism-based thinking has multiple interpretations in the social sciences.

Mechanistic thinking started to become mainstream in the 1990s. A work of particular significance in this respect was *Social Mechanisms* (1998), edited by Peter Hedström and Richard Swedberg. Among contributors to this volume are many scholars who have been influential in the development

of mechanism-based thinking in the social sciences. Later this approach developed into the analytical sociology movement (Hedström & Bearman 2009; Hedström & Ylikoski 2010) that has most systematically developed the program of mechanism-based social science. At the same time, philosophers of the social sciences such as Daniel Little (1991) and Mario Bunge (1997) started to talk about mechanisms. In political science mechanisms became a focus of attention a little bit later. There causal mechanisms played an important role in debates concerning research methodology and causal inference. Especially noteworthy has been the development of process tracing methodology for case-study research (Bennett & Checkel 2015). In general, the debates in political science and sociology have been quite similar: advocates of mechanisms have criticized the simplistic use of statistical methodology and the downplaying the importance of causal process assumptions in causal inference. However, there are also some important differences as the later discussion will show.

In the following discussion I focus, first, on what is known as Coleman's diagram, which helps to identify the core challenges of mechanism-based theorizing in the social sciences. It also provides a context in which to discuss the appeal of mechanism-based explanation among social scientists. In an attempt to make sense of this debate I will introduce a distinction between causal scenarios and causal mechanism schemes. By way of illustrating social-scientific thinking about mechanisms I discuss the use of agent-based simulation as a tool for mechanism-based theorizing, and introduce the idea of metamechanism. In the concluding discussion I show how mechanistic ideas have been used in social-scientific debates about causal inference.

2 Coleman's diagram

A useful starting point for discussing social mechanisms is a diagram known as Coleman's boat (Coleman 1987, 1990). It is commonly used in discussions of social mechanisms, and one could say that it has become an emblem of mechanism-based thinking in the social sciences. Coleman himself

did not employ mechanistic vocabulary, but in many ways the diagram exemplifies mechanistic thinking in the social sciences (Ylikoski 2016).

<FIGURE 1 HERE>

Figure 1. Coleman's diagram

The recommended starting point for unpacking the diagram is node D, which represents the macro-social *explanandum* that the sociologist finds interesting. Node A represents some macro-social variable that is associated with D. For example, let us assume that A is the implementation of a government job-training program and D is the decrease in the level of youth unemployment. Did the job-training program cause the decline in unemployment? In other words, can arrow 4 be interpreted as causal?

Coleman's main point is the following. In order to justify causal claims like this it is necessary to understand *how* the suggested cause brings about the effect in question. This idea has two components. The first of these concerns the justification of causal claims: the statistical data on the relevant macro variables is usually so sparse as to be insufficient for establishing such a claim. Coleman's suggestion is to test the causal claim by finding out whether there is an empirically supported mechanism by which A brings about D: if there is, the claim is supported, and if not, then it should be scrapped. The second component concerns explanation. Coleman and other supporters of mechanism-based thinking stress the importance of understanding how the effect was brought about. Thus, even if it were possible to show that variable A is a causal difference-maker for variable D, it would not suffice for a theoretically satisfactory explanation: the explanation has to spell out the mechanism. These two ideas are at the core of social-scientific debates about causal mechanisms. I will discuss both in detail later, but first I will consider other elements of the diagram.

The key point in the diagram is that macro variables have to be connected to activities by agents. In most cases these agents are human persons, but Coleman also allows for various sorts of corporate

agents to take this role. Thus arrow 1 describes the way in which changes in macro conditions influence the relevant agents. The change in A may bring about changes in the beliefs, desires, or other mental attributes of these agents, or it might change the opportunities or incentives they are facing. Hedström and Swedberg (1998) call these influences *situational mechanisms*. They cover the ways in which social structures constrain and enable individuals' opportunities for action, and how the cultural and social contexts influence individuals' goals, beliefs, habits, or cognitive frames.

Arrow 2 in the diagram covers the role of the theory of action in sociological explanation. The purpose of theory of action is to connect changes in agents' opportunities and mental states to changes in their behaviors or actions. Coleman used rational choice theory for this purpose, but it is also possible to utilize other theories of action (Hedström & Ylikoski 2014). The important point here is that mechanical explanations in the social sciences bottom out (Craver 2007) at the level of individual action (Coleman 1990: 4). The behaviors of individual persons are the basic components of social mechanisms and social scientists do not look for explanations for these *action-formation mechanisms*: this is a job for cognitive scientists and psychologists.

Although both situational and action-formation mechanisms certainly pose their own challenges, Coleman argues that the transformational mechanisms (arrow 3) are the biggest bottleneck in sociological theory. Social scientists know a great deal about how individuals' desires, beliefs, and opportunities, for example, are influenced by the social contexts in which they are embedded (situational mechanisms), and about how these desires, beliefs, and opportunities influence actions (action-transformation mechanisms), but when it comes to the link between individual actions and social outcomes they are often forced to resort to hand-waving. This is something Coleman aimed to highlight with his diagram: macro-level patterns are often difficult to predict from individual-level descriptions, and the way in which individual actions produce social patterns is rarely a simple process of aggregation. This is not a problem that is specific to individualistic theories: although holistic theories tend to highlight the contextuality and complexity of everything, in practice their

micro-macro assumptions tend to be rather simplistic. It is frequently assumed that macro facts simply reflect the relevant micro facts, and vice versa. Schelling's (1978) well-known checkerboard model shows how wrong this assumption is even in very simple settings: it shows how residents who do not favor segregation may still end up in a highly segregated neighborhood. It is clear even from this very simple model that one cannot assume that macro facts will reflect individual (or average individual) preferences. It is necessary to understand how the macro outcome is brought about by the interdependent actions of individuals, in other words one has to understand "the rules of the game" (Coleman 1990: 19).

Coleman was never able to fully unpack the metaphor of the rules of the game, but his discussion and examples were enough to interest sociologists in the theoretical challenge of the micro-macro link. Consequently, whereas others might talk loosely about emergent macro-scale properties, sociologists following Coleman set themselves the much harder task of explaining how that emergence comes about. The key mechanistic idea captured in the diagram is that only when we have understood the whole chain of situational, action-formation and transformational mechanisms have we understood the relation between the macro-scale social facts. Underlying each A-D association or causal relation is a combination of these mechanisms. All three are required otherwise the explanation would not cover all elements of the micro-macro link.

Users of the diagram have been criticized (Jepperson & Meyer 2011; Little 2012) for being committed to a reductive ideal of methodological individualism. I believe this criticism is misplaced, at least to some extent. The diagram does not describe the reduction of macro facts to facts about individuals and their relations. The mistake is the assimilation of the Coleman diagram into the supervenience/realization diagrams used in the philosophical debate on mental causation. Coleman's arrows represent explanatory rather than reductive relations, and if one considers the details it is evident that various sorts of structural assumptions play a central role in situational and transformational mechanisms. Coleman's main point is that the structural facts are not explanatory in themselves. It is necessary to understand *how* they bring about their effects via the activities and

cognitions of individuals, and a full understanding of how the social whole works requires an understanding of how its behavior is *generated* by the activities of its members. In Coleman's view, individual agents are the basic building blocks of social mechanisms, and therefore have to be *included* in any mechanistic social explanation. However, this does not imply that the explanation entails the reduction of macro-social facts to facts about individuals.

3 Mechanism-based explanations

In the social sciences, as in the philosophy of science, the mechanism-based accounts of explanations have been developed as alternatives to the once dominant covering-law account of explanation (Hempel 1965). While social scientists have been familiar with famous counterexamples to that theory and philosophical problems raised by them (such as problems of the explanatory relevance, the symmetry between explanation and prediction, the asymmetry of explanation, and many problems in analyzing the notion of law), social scientists have mainly been concerned with the apparent implausibility of the theory as a model for the social sciences. There are very few laws in the social sciences, and even those are better described as *explananda* rather than *explanatia*. Most social-scientific generalizations that are not truisms are quite limited in domain, and include exceptions that neat *ceteris paribus* conditions cannot cover. Furthermore, the strategy of formulating social-scientific theories in terms of axioms and laws has turned out to be very unproductive, resulting in incomplete collections of sterile generalizations marred with unclear domains of application and high levels of conceptual indeterminacy. Finally, the basic idea of a covering-law theory is counterintuitive for social scientists. The best social-scientific explanations seem to do more than simply subsume the phenomenon under more general empirical regularities: a good explanation shows *how* the suggested cause brings about the effect to be explained. The initial appeal of mechanistic ideas lies in this generative notion of explanation.

The dissatisfaction with covering-law theory has not led to general agreement on the definition of a mechanism, however. The intuitive idea can be developed in multiple directions, especially when

the people have different applications in mind. As a result, the literature on social mechanisms notoriously abounds with apparently incompatible definitions of mechanisms. Mahoney (2001), for example, lists 24, and subsequent contributions have added to the number. Some critics (e.g., Norkus 2005) regard this multiplicity as a serious problem for mechanism-based explanation. This could be an overstatement in that, although the definitions are formally incompatible, most of them could be considered attempts to capture the same basic ideas. The absence of a generally agreed definition for basic concepts such as gene and species has not stalled the development of the biological sciences. As long as mechanism-oriented social scientists agree on the central exemplars of mechanism-based explanations and share a similar understanding of their general characteristics (Hedström & Ylikoski 2010), they should manage well without a general definition. Thus the main challenge is not to provide a general definition of mechanism-based explanations, it is rather to arrive at a consensus about prototypical examples of mechanistic explanation.

A more serious problem is that much of the mechanism discourse in the social sciences is quite loose. As noted above, as a notion, mechanism belongs to the general causal vocabulary of social scientists. Some advocates of social mechanisms seem to go along with this loose talk, and it is not uncommon to name processes that produce certain kinds of outcome as mechanisms and leave it at that. In such usage, the mechanism is just a label for a black box, a name for an effect, not an explanation. The invisible hand, cumulative advantage, and democratization are not mechanisms in themselves, they are processes that produce specific kinds of outcomes. At most they are names for families of mechanisms. However, if one does not distinguish at least some family members, all one has is a placeholder for something substantial. The key point here is that if one looks into *how* the invisible hand or cumulative advantage work in practice, one finds that there may be multiple mechanisms underlying these effects, which sometimes work separately, and sometimes act together in various combinations. The point here is not about the proper use of the word ‘mechanism’: it is about dangerous ambiguity. Confusion between naming an effect and providing a mechanistic explanation for it may give rise to an illusion of depth of understanding (Ylikoski

2009). Furthermore, the ambiguity makes one blind to the possibility of there being multiple mechanisms that are responsible for similar kinds of effect.

4 Causal scenarios and causal mechanism schemes

Social scientists are interested in explaining particular causal outcomes and in developing general theories about social mechanisms. In both contexts they refer to mechanisms, which sometimes causes unnecessary confusion. To avoid this, it is useful to distinguish between causal scenarios and causal mechanism schemes (Ylikoski & Aydinonat 2014). *Causal scenarios* are (selective) representations of particular causal processes responsible for some concrete event or phenomenon. Used thus, causal mechanism refers to a causal narrative that describes the process that is responsible for the *explanandum*. This narrative may be highly detailed or a mere sketch, but in any case it does more than cite a cause that had an effect on the event (see Chapter 34): it describes the crucial elements in the relevant causal chain. In other words, the causal scenario describes how the *explanandum* event came about. The distinction between *how possibly* and *how actually* explanations applies here. Usually there are many different ways in which the outcome *could* have come about, which competing how-possibly scenarios describe. The challenge for researchers is to find evidence that could discriminate between these alternatives and enable them to make a judgment about which scenario is the true explanation.

When political scientists talk about process tracing they are concerned with causal scenarios. On the other hand, when analytical sociologists discuss causal mechanisms they tend to refer to *causal mechanism schemes* (compare Chapter 19). For example, sociologists talking about self-fulfilling prophecies or vacancy chains are referring to causal mechanism schemes, which are abstract representations of mechanisms that could bring about effects of a certain kind. The *explanandum* of a causal mechanism scheme tends to be quite abstract, or stylized, reflecting the fact that such schemes are not primarily explanations of particular facts, but schemes for constructing them. Thus it is useful to think of them as abstract building blocks that can be adapted and filled in to serve a

role in causal scenarios that explain particular facts. A single causal scenario might be a combination of many different causal mechanism schemes, and might even contain mechanism schemes that have opposite causal effects.

Causal mechanism schemes are at the core of analytical sociology's account of growth of theoretical knowledge, according to which social-scientific knowledge accumulates through the development of middle-range theories (Hedström & Udéhn 2009) about social mechanisms. In this view the core theoretical knowledge comprises a collection of causal mechanism schemes that can be adapted to particular situations and explanatory tasks. According to this toolbox view (Hedström & Ylikoski 2010, Elster 2015), social-scientific knowledge is not integrated into highly abstract general theory, but consists of a growing collection of causal mechanism schemes that are mutually compatible. The understanding of the social world accumulates as the knowledge of the mechanism schemes becomes more detailed and the number of known mechanisms increases. Understanding of more complicated phenomena requires combining different mechanism schemes, hence knowledge also expands through learning how to create these "molecular" mechanisms.

The similarity between this view of the architecture of theoretical knowledge in the social sciences and the mechanistic view of the biological sciences is obvious (see Chapter 19), and offers a fruitful opportunity to compare two domains of knowledge. However, the relevance of the toolbox vision is not limited to meta-theory: it also gives new tools to counter the fragmentation of the social sciences. Causal mechanism schemes can be shared among the different subfields, which would allow for a novel type of integration: various subfields employ and develop the same theoretical toolbox and thereby benefit from each other's work. The toolbox vision for sociological theory shows that in the social sciences mechanistic ideas are not confined to discussions about explanation and causation, but they also play an important role in how social scientists think about the nature of social-scientific knowledge.

5 Agent-based simulation and generative explanation

One of the main points about mechanism-based explanation is that it should describe how the properties, activities and relations of components bring about the phenomenon to be explained (see Chapters 1 and 19). This concern with generative processes makes *generative sufficiency* a central concern in the evaluation of explanations: the suggested explanation should, at least in principle, be capable of bringing about the outcome in the specified circumstances. However, demonstrating generative sufficiency is not easy. The social outcomes of interest typically result from numerous individuals acting and interacting with one another over extended periods of time. Furthermore, the aggregate behavior of these complex dynamic systems is extremely difficult to understand and to predict without the aid of analytical tools. *Agent-based simulation modeling* (ABM) provides such a tool (Miller & Page 2007, Squazzoni 2012, Ylikoski 2014). There is a natural affinity between the components of mechanism-based explanations and ABM. Like any society, social-scientific ABMs comprise agents with goals and beliefs. These agents possess resources and influence each other. It is easy to see how the macro-patterns they create come to be regarded as analogical to the processes that social scientists study. An ABM sheds light on how the phenomenon to be explained could have been generated, and how changes in agents' attributes or relational structures change the macro outcomes.

Joshua Epstein (2006) offers the strongest ABM-based formulation of the idea of generative explanation. According to Epstein, "*If you didn't grow it, you didn't explain it*." He meant that producing the macro-level outcome by means of ABM is a necessary condition for its explanation. However, it could be argued that simply "growing" the phenomenon of interest is not sufficient to engender a proper understanding. The crucial challenge is to understand *how* the specified micro-configuration produces the phenomenon. Thus, in reformulating Epstein's slogan Macy and Flache (2009: 263) make a crucial point: "*If you don't know how you grew it, you didn't explain it*." In any case, social scientists using ABM are not satisfied with mechanistic explanation as mere storytelling. They do not think it is enough to provide a qualitative narrative about the process and its components. They also consider it important to demonstrate that the suggested mechanism can,

in fact, bring about the effect to be explained. Having a detailed *how possibly* scenario is also a precondition for the real empirical testing of it. Only if competing causal scenarios are clearly articulated is it possible to look for crucial empirical evidence that discriminates between them. Still, ABM has a long way to go to become a mainstream tool in sociological research: it is basically a tool for working with causal mechanism schemes rather than concrete causal scenarios (Ylikoski 2014). However, the connection between ABM and mechanism-based thinking is strong.

6 Metamechanisms

Jeremy Freese and Karen Lutfey's (2011) *metamechanism* is an interesting addition to the conceptual toolbox of mechanical philosophy. Behind this idea is Link and Phelan's (1995) suggestion that socioeconomic status is a *fundamental cause* of health differences. They refer to an extremely robust empirical finding: socio-economic status and health are strongly correlated. The lower-status people are, the sooner they die, and the worse health they have while alive. This association holds for virtually any society and historical period for which there is adequate empirical data. The puzzling thing is that the causes of death and disease have changed a lot over the past hundred years. In other words, the proximate mechanisms of death and ill-health are highly variable. What explains this puzzling pattern?

Freese and Lutfley suggest that the idea of fundamental cause lends itself to a mechanistic interpretation: socioeconomic status (SES) is associated with a metamechanism, in other words a general mechanism that explains the generation of multiple proximate mechanisms that reproduce a particular relationship in different places and at different times. This helps to make sense of Link and Phelan's findings. People with a higher SES have more resources and education, which makes it easier for them to utilize new medical services and health-improving inventions. Even if healthcare is universal and free, the higher-SES people are better placed to make use of it. Other metamechanisms may have a similar influence. There are spillover effects, for example, even among individuals who do not especially care about their health. Higher-SES individuals will have

better health because they tend to gain more benefits from the purposive actions of others in their social networks (that are partly based on SES). Similarly, health-related institutions might be biased towards higher-SES people: they are given better service and better understand the instructions they get. The proximate mechanisms vary, but as long as the metamechanisms remain in place there will be health inequalities. The usefulness of such mechanisms is not limited to the sociology of health, and could extend to some biological contexts, for example.

7 Mechanisms and statistical methodology

Apart from explanation, mechanisms are also important in the context of justification of causal claims. Especially in non-experimental contexts that are common in the social sciences, they are said to play a crucial role in distinguishing true causal relations from spurious correlations. Knowing that there is a mechanism through which X could influence Y supports the inference that X is a cause of Y. In addition, the absence of a plausible mechanism linking X to Y gives good reason to be suspicious of any straightforward causal interpretation of the association. Knowledge of mechanisms is also applied in extrapolating causal findings. The assumption of similarity among causal mechanisms is a crucial element in making inferences from one setting or population to another (Steel 2008, see also Chapter 32). However, there is much ambiguity in these mechanistic slogans. It is impossible to ascertain that knowledge of mechanisms is *necessary* for justifying causal claims without a clear idea of what such knowledge consists of, and how much of it is needed. All causal inference presupposes some causal background assumptions, but do all such assumptions concern causal mechanisms? It should also be recognized that mechanisms are not a magic wand for causal inference in the social sciences. The problem in many cases is not the absence of a possible mechanism, but insufficient evidence to discriminate between competing mechanistic hypotheses. Similarly, lazy mechanism-based storytelling is a constant threat: having a good story is no substitute for real statistical evidence. It is not rare for a good story about a (possible) mechanism to make people forget how important it is to test whether such a mechanism really is in place and whether it can really account for the intended *explanandum*.

Thinking in terms of mechanisms is often set against statistical methodology in the social-scientific debates on causal inference. This opposition takes different forms. Many critical realists rely on theorizing about mechanisms as an alternative to using statistics and causal-modeling techniques. Causal modeling is said to embody problematic Humean ideas about causation that make it suspect and of limited value, which is why critical realists tend to use statistics for descriptive purposes only and prefer qualitative evidence and theoretical argumentation.

An alternative is not to give up statistical and causal modeling, but to object to their use without consideration of the relevant causal mechanisms. Here the claim is that statistical techniques have replaced substantial theorizing. Peter Hedström conveys a fairly common sentiment in the following: "Although most causal modelers refer to sociological theories in their work, they rarely pay it any serious attention. More often than not, they simply use theories to justify the inclusion of certain variables taken from a data set that has often been collected for entirely different purposes than the one to hand. Theoretical statements have become synonymous with hypotheses about relationships between variables, and variables have replaced actors as the active subjects with causal powers." (Hedström 2005: 105). What he is calling for is a fuller incorporation of theory (about causal mechanisms) into the research design and the interpretation of statistical data. According to this view, sociological research should not be limited to the measurement of causal effects among conventional variables, but should focus on how the social world works. Given this background, it is understandable that advocates of mechanisms have been resistant to the common assumption that mechanisms are just intervening variables. Although the existence of a causal process or mechanism implies that there are intervening variables, not all of them are necessarily of the right sort. They may tell more about other effects of the mechanism than about the mechanism itself, for example. Thus, mere intervening variables do not guarantee the explanatory depth that is the main concern among mechanists.

The promise of mechanisms is to provide something more to causal inference, but what is that additional element? It is now generally recognized that the uses of statistical tools such as

regression analysis presupposes substantial assumptions about the causal relations that are modeled (Kincaid 2012). A substantial proportion of these background assumptions concern possible causal mechanisms. There have also been attempts to incorporate mechanistic thinking into causal modeling (e.g., Knight and Winship 2014; see also Imai *et al.* 2011), that define mechanisms as *modular sets of entities connected by relations of counterfactual dependence*. (Philosophers will easily recognize the influence of Woodward 2002 here.) According to Knight and Winship, as long as the mechanisms studied satisfy the requirement of modularity, Judea Pearl's DAG (Directed Acyclic Graph) calculus is a powerful tool in terms of facilitating the rigorous consideration of mechanisms in causal analysis. Their main argument is that mechanisms and causal analysis can be combined fruitfully in a way that could help in identifying causal effects even when traditional techniques fail. It is to be expected that attempts like this to combine mechanistic thinking with causal modeling will become more frequent. It appears that the issue is not really about statistical methods, but concerns the way they are used.

Causal graphs are useful when it comes to thinking about social mechanisms, and mechanists should welcome them. However, it should be borne in mind that there are serious limitations in terms of what can be represented as DAGs. Coleman's diagram discussed above cannot be interpreted as a DAG, for example. One problematic spot is arrow 1 between A and B. Consider how the demographic change caused by war affects the number of potential marriage partners available to women, how laws allowing same-sex marriage change the opportunities of same-sex couples to arrange their legal relationship, and how the improved educational level of society is related to the education of individuals. It cannot be said in any of these cases that the relations are strictly causal, whereas it is plausible to say in these cases the cited A-facts partially consist of the mentioned B-facts. However, some of the relevant consequences of changes in A might well be causal at the same time. Thus it seems that A-B explanatory dependency is based on (various) mixtures of causal and constitutive relations (Ylikoski 2013, 2016). This makes sense on the

theoretical level, but also makes it impossible to interpret arrow 1 as causal or the diagram as a DAG.

In political science debates, especially in international relations, mechanistic thinking has also been set against statistical methods. However, the context has been that of qualitative case studies, the point being that "the standard quantitative template" is ill adapted for such research. The alternative methodology for causal inference is called *process tracing*, which Bennett & Checkel define as "… the analysis of evidence on processes, sequences, and conjunctures of events within a case for the purposes of either developing or testing hypotheses about causal mechanisms that might causally explain the case." (Bennett & Checkel 2015: 7) As this definition implies, mechanisms – causal scenarios – play a central role in process tracing. Causation is understood here as a continuous process and the task is to explain a singular event. A central concern in process tracing is with the sequence of events and mechanisms involved in the unfolding of the process. The researcher looks for diagnostic evidence that can be used to discriminate between alternative causal scenarios that could explain the event.

Process tracing is often presented as a method for "within-a-case causal inference", but its functions remain somewhat unclear: it has been presented as a tool for theory testing, theory development, and for explaining singular outcomes (Beach & Pedersen 2013). These roles are naturally interlocking and not so easy to distinguish. It is rare to begin a case study with the goal of developing a theory, but the search for explanation might generate novel theoretical ideas. The competing hypotheses are competing explanations, and hence it does not make much sense to distinguish hypothesis testing and explanation as separate activities. Furthermore, it is doubtful that process tracing captures something that is unique to qualitative case-study research given that similar case-based *causal process observations* play an important role in the evaluation of evidence in experimental, comparative, and statistical studies. Consideration of the causal processes that produce the data to be analyzed is a major concern in all research. Thus there is much room for building new bridges between different research methodologies based on mechanistic ideas.

An interesting contribution in the literature on process tracing is the taxonomy of tests for causal hypotheses. Originally presented by Van Evera (1997), but later adapted by others (Beach & Pedersen 2013, Bennett & Checkel 2015), this taxonomy describes different kinds of tests that people struggling with the problem of multiple competing causal scenarios could look for. Passing a *Smoking-Gun test* gives strong support to the hypothesis and substantially weakens its rivals, but failure does not imply that the hypothesis is eliminated: it is only weakened and the rivals gain some additional support. The very name of the test is illustrative: finding a person with a smoking gun straight after a shooting makes him or her a strong suspect, but the lack of such evidence does not eliminate this person from the list of suspects. In the case of the *Hoop test* the implications are the opposite: passing the test affirms the relevance of the hypothesis, but does not confirm it. However, if the hypothesis fails the test, it is eliminated. Here the illustration is the familiar idea of an alibi: giving a speech to an audience of dozens of people at the time of the crime provides strong grounds for elimination from the list of suspects. However, the mere lack of an alibi does not yet provide positive evidence of guilt. *Doubly Decisive tests* are rare, but they are the strongest. A hypothesis that passes the test is confirmed and competing hypotheses are eliminated. The consequences of failure are also drastic, but only for the failed hypothesis: it is eliminated. Having a clear video recording of the crime with the shooter's face clearly visible is an example of a doubly decisive test: the video evidence demonstrates that the particular suspect is responsible, and also shows that other suspects took no part in the shooting. Finally, there are *Straw-in-the-Wind tests*. Passing such a test gives the hypothesis some support, but failure does not mean that it is eliminated. The evidence is weak or circumstantial and cannot in itself prove or disprove the suspect's guilt. However, in favorable circumstances enough accumulated evidence of this type may convince the jury. In the context of the social sciences most evidence is inherently of the *Straw-in-the-Wind* kind, and social scientists rarely encounter evidence that could be considered *Doubly Decisive,* or that would constitute a *Smoking-Gun* test. However, the taxonomy is useful in

highlighting the fact that the value of evidence depends on the set of alternative hypotheses, not on some intrinsic relationship between a single hypothesis and the empirical material.

8 Conclusion

In the above review I have covered some prominent and interesting themes in the social-scientific debate on mechanisms. I have left many things out, some of which are discussed in other chapters of this book. I have attempted to show that mechanism-based thinking is a strong and expanding meta-theoretical idea in the social sciences, and that some of the ideas, such as the distinction between causal scenarios and causal mechanism schemes, and the notion of metamechanisms, might also be of interest in other disciplines.

References

Beach D & Pedersen RB (2013). *Process-Tracing Methods. Foundations and Guidelines*. The University of Michigan Press.

Bennett A & Checkel JT 2014: *Process Tracing. From Metaphor to Analytic Tool*. Cambridge University Press.

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

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

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

Coleman, JS. (1987). Microfoundations and Macrosocial Behavior, in Alexander, Giesen, Münch & Smelser (eds.): *The Micro-Macro Link*. University of California Press: 153-173.

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

Elder-Vass, D. (2011). The Causal Power of Social Structures. Emergence, Structure and Agency. Cambridge, UK: Cambridge Univ. Press.

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

Elster J. (2015). *Explaining Social Behavior. More Nuts and Bolts for the Social Sciences*. 2nd. ed. Cambridge, UK: Cambridge Univ. Press.

Freese, J. & Lutfey, K. (2011). Fundamental Causality: Challenges of an Animating Concept for Medical Sociology, in Pescosolido et. al. (eds.): Handbook of the Sociology of Health, Illness, and Healing. Springer: 67-81.

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

Hedström, Peter (2005). *Dissecting the social. On the principles of analytical sociology*. Cambridge: Cambridge University Press.

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

Hedström, P. & Swedberg R. (eds.) (1998). *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press.

Hedström, P. & Udeʹhn, L. (2009). Analytical sociology and theories of the middle range. in Hedström & Bearman (eds.) (2009). *The Oxford Handbook of Analytical Sociology*. Oxford: Oxford Univ. Press: 25-49.

Hedström, P. & Ylikoski, P. (2010). Causal mechanisms in the social sciences. *Annual Review of Sociology* 36, 49-67.

Hedström, P. & Ylikoski, P. (2014). Analytical Sociology and Rational Choice Theory', in *Analytical Sociology: Norms, Actions and Networks* (edited by Gianluca Manzo), Wiley: 57-70.

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

Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2011). Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies. *American Political Science Review 105:* 765-789.

Jepperson, R. & Meyer JW (2011). Multiple levels of analysis and the limitations of methodological individualisms. *Sociological Theory* 29, 54-73.

Kincaid, H. (2012). Mechanisms, Causal Modeling, and the Limitations of Traditional Multiple Regression, in Kincaid (ed.): *The Oxford Handbook of Philosophy of Social Science*. Oxford: Oxford University Press: 46-64.

Knight, CR and Winship, C. (2014). The Causal Implications of Mechanistic Thinking: Identification Using Directed Acyclic Graphs (DAGs), in Morgan S. (ed.) *Handbook of Causal Analysis for Social Research*, Springer: 275-300.

Lawson, T. (2003). *Reorienting Economics*, London and New York: Routledge.

Link, B. G. and Phelan, J. (1995). Social Conditions as Fundamental Causes of Disease. *Journal of Health and Social Behavior* 35 (Extra Issue):80–94.

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

Little, D. (2012). Explanatory autonomy and Coleman's boat, *Theoria* 74, 137-151.

Mahoney, J. (2001). Review essay: Beyond correlational analysis: recent innovations in theory and method. *Sociological Forum* 16, 575-593.

Miller, JH. & Page, SE (2007). *Complex Adaptive Systems. An Introduction to Computational Models of Social Life*. Princeton: Princeton University 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: 348–72.

Squazzoni, F. (2012). *Agent-Based Computational Sociology*. Chichester: John Wiley & Sons.

Woodward J. (2002). What is a mechanism? A counterfactual account. *Philosophy of Science* 69: S366–S377.

Ylikoski, P. (2009). The illusion of depth of understanding in science. In H. De Regt, S. Leonelli, & K. Eigner (Eds.), *Scientific understanding: Philosophical perspectives* (pp. 100-119). Pittsburgh: Pittsburgh University Press.

Ylikoski, P. (2013). 'Causal and constitutive explanation compared', *Erkenntnis* 78, 277-297.

Ylikoski, P. (2014). Agent-based simulation and sociological understanding. *Perspectives on Science* 22, 318-335.

Ylikoski, P. (2016). Thinking with the Coleman Boat, forthcoming.

Ylikoski, P. & Aydinonat, E. (2014). Understanding with theoretical models. *Journal of Economic Methodology* 21, 19-36.

Acknowledgements

This research has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 324233, Riksbankens Jubileumsfond (DNR M12-0301:1), and the Swedish Research Council (DNR 445-2013-7681 and DNR 340-2013-5460).