Published in *The Oxford Handbook of Philosophy of the Social Sciences* (edited by H. Kincaid), Oxford University Press, 21-45.

Chapter 2: Micro, Macro, and Mechanisms

Petri Ylikoski

2.1 Introduction

This chapter takes a fresh look at micro-macro relations in the social sciences from the point of view of the mechanistic account of explanation. Traditionally, micro-macro issues have been assimilated to the problem of methodological individualism (Udéhn 2001, Zahle 2006). It is not my intention to resurrect this notoriously unfruitful controversy. On the contrary, the main thrust of this chapter is to show that the cul-de-sac of that debate can be avoided if we give up some of its presuppositions. The debate about methodological individualism is based on assumptions about explanation, and once we change those assumptions, the whole argumentative landscape changes.

The idea that social scientific explanations are based on causal mechanisms rather than covering laws has become increasingly popular over the last twenty years or so (Hedström and Ylikoski 2010). Interestingly, a similar mechanistic turn has occurred also in the philosophies of biology and psychology (Wright and Bechtel 2007). Until recently, the connections between these two emerging traditions for thinking about mechanisms have been rare. The aim of this chapter is to employ ideas developed by philosophers of biology to address some issues that the advocates of mechanisms in the social sciences have not yet systematically addressed. I argue that ideas about levels of explanation and

reductive research strategies, which were originally developed in the context of cell biology and neuroscience, can be fruitfully adapted to the social sciences. They can both strengthen the case for mechanism-based explanations in the social sciences and bring the philosophy of social science debates closer to social scientific practice.

The chapter is structured as follows. In the first section, I will take a look at recent work on mechanism-based explanation. While I suggest that the mechanistic account of explanation presupposes some more fundamental ideas about explanatory relevance and causation, I also argue that it provides a fruitful tool for thinking about micro-macro relations in the social sciences. In the second section, I will criticize a common philosophical way of formulating the micro-macro issue and provide my own characterization that is not dependent on the assumption that there is a unique or comprehensive micro level. The third section introduces the distinction between causal and constitutive explanation, and argues that this distinction helps to make sense of the call for microfoundations in the social sciences. The final section will take on a doctrine that I call intentional fundamentalism, and it challenges the idea that intentional explanations have a privileged position in the social sciences.

2.2 Mechanism-based explanation

The idea of mechanism-based explanation has been developed independently among social scientists (Harré 1970; Elster 1989, 2007; Little 1991; Hedström and Swedberg 1998; Hedström 2005; for a review see Hedström and Ylikoski 2010) and philosophers of biology (Bechtel 2006, 2008; Craver 2007; Darden 2006; Wimsatt 2007). In the social sciences, the idea of causal mechanism has been used mainly as a tool for methodological

criticism, while in the philosophy of biology the motivation has been that of finding a descriptively adequate account of biological explanation. Despite these separate origins and motivations, both traditions are clearly building on similar ideas about scientific explanation. For example, both share the same dissatisfaction with the covering law account of explanation (Hedström 2005; Craver 2007).

There is no consensus on the right definition of a causal mechanism. Although some theorists find such a situation frustrating, I do not think this constitutes a real problem. The entities and processes studied by different sciences are quite heterogeneous, and it is probably impossible to propose a mechanism definition that would be both informative and cover all the prominent examples of mechanisms. Some disciplines, such as cell biology (Bechtel 2006) and the neurosciences (Craver 2007), study highly integrated systems, whereas others, such as evolutionary biology and the social sciences, study more dispersed phenomena (Kuorikoski 2009), so it is more plausible to think that informative characterizations of mechanisms are field specific. The task of a philosophical account is to show how these exemplars are related to general ideas about explanation, evidence, and causation, not to engage in verbal sophistry. However, it is possible to give some general characteristics of mechanisms.

First, a mechanism is always a *mechanism for something*; they are identified by the kind of effect or phenomenon they produce. Second, a mechanism is an *irreducibly causal notion*. It refers to the entities of a causal process that produces the effect of interest. Third, a mechanism has a *structure*. When a mechanism-based explanation opens the black box, it makes visible how the participating entities and their properties, activities, and relations produce the effect of interest. 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? Finally, there is *a hierarchy of mechanisms*. 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 will explain them. In other words, the explanations employed by one field always *bottom out* somewhere. However, this fundamental status of certain entities, properties, and activities for a given mechanism is only relative, as they are legitimate targets of mechanistic explanation in another field. Of course, this chain of explanations ends somewhere—there are no mechanism-based explanations for fundamental (physical) processes (Hedström and Ylikoski 2010).

Although the mechanism-based account is often presented simply as an idea about scientific explanation, the notion of mechanism is associated with a wider set of ideas about scientific knowledge. For example, there are ideas about the justification of causal claims, the heuristics of causal discovery, the presentation of explanatory information, and the organization of scientific knowledge (Ylikoski 2011). There is no doubt that these not yet clearly articulated ideas partly explain the appeal of the approach. For example, as I will show later in this chapter, claims about the explanatory role of mechanisms are often confused with claims about their relevance to the justification of causal claims (see also Kincaid, this volume).

While I think all the above ideas are important advances in understanding explanatory reasoning in science, it is not necessary to assume that the notion of

mechanism is the ultimate solution to all problems in the theory of explanation. On the contrary, the mechanistic theory presupposes accounts of explanatory relevance, causation, and the nature of generalizations that provide the basis for mechanisms. The notion of mechanism should not be treated like a black box. I have argued elsewhere (Hedström and Ylikoski 2010; Ylikoski 2011) that if the mechanistic ideas are combined with the theory of explanation developed by James Woodward (2002, 2003), we can get quite far in solving these problems. While for the present purposes we do not have to consider in detail the relation between mechanisms and generalizations, some comments on explanatory relevance are in order as later arguments depend on it.

A mechanism-based 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 of 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 mechanism-based explanations involve counterfactual reasoning about possible changes and their consequences (Ylikoski 2011). A natural way to understand these causal counterfactuals is to understand them as claims about the consequences of ideal causal interventions (Woodward 2003, 2008). The causal counterfactual tells us what would have happened to the effect if the cause had been subject to a surgical intervention that would not have affected anything else in the causal configuration. An advantage of the interventionist account of causation is that it allows talking about causal dependencies in every context

where the notion of intervention makes sense. Unlike some other theories of causation, such as various process theories, it is level-blind and applicable to special sciences such as cell biology or sociology.

2.2.1 Mechanisms and Reductive Explanation

One of the distinctive features of the mechanistic approach to explanation is that it reorients the issues related to reductionism and reductive explanation. In one sense the mechanistic way of thinking is thoroughly reductionist: It attempts to explain activities of mechanisms in terms of their component parts and their activities, and then subjects the component mechanisms to the same treatment. In this sense, the reductive research strategy has probably been the single most effective research strategy in the history of modern science. However, there is another sense in which mechanism-based explanations are clearly nonreductionist: Although they do refer to the micro level, they do not replace or eliminate the higher-level facts nor the explanations citing them. Rather than serving to reduce one level to another, mechanisms bridge levels (Darden 2006; Craver 2007; Wright and Bechtel 2007; Richardson 2007; McCauley 2007; Wimsatt 2007).

The mechanistic account of reductive explanation differs significantly from the traditional philosophical accounts of intertheoretical reduction that conceive reduction as a derivation of one theory from another (Richardson 2007; McCauley 2007). The mechanical account of reductive explanation does not start with a strongly idealized picture of a discipline-wide theory that contains all knowledge about its level. Nor does it conceive reduction as a deductive relation between such theories (or their corrected versions). Rather, reductive mechanistic explanations are constructed piecemeal with a

focus on particular explanatory targets. While there is an assumption that everything is mechanistically explainable and a presumption that ultimately all mechanistic accounts are mutually compatible, there is no overarching effort to combine them into one grand theory that would cover all the phenomena that the scientific field studies. Also, contrary to the traditional accounts that conceive reduction as elimination or replacement, the mechanisms are inherently multilevel. The components and their operations occur and are investigated at one level, whereas the mechanism itself and its activities occur and are investigated at a higher level. In this sense accounts of mechanisms often have the character of interfield theory (Darden 2006). This makes it difficult to characterize the reductive understanding provided by mechanical explanations as deductive relations between independent theories.

The mechanistic stance also gives reasons for rethinking the notion of levels. According to the traditional layer-cake conception, there is a neat hierarchical layering of entities into levels across phenomena, and the scientific disciplines (e.g., physics, chemistry, biology, psychology, sociology) are distinguished from each other by the level of the phenomena that they are studying (see Oppenheim and Putnam 1958). From the mechanistic point of view, this way of thinking unnecessarily drives together levels of nature and science, and misleadingly suggests that the levels are both comprehensive and the same independently of the investigative context (Craver 2007). The actual scientific disciplines do not match neatly with the metaphysical picture of levels of organization or reality. And while there are many problems in a serious characterization of the metaphysical picture of levels, there do not seem to be any particularly good reasons to accept such a metaphysical constraint for an account of scientific explanation.

The notion of the levels of mechanism plays an important role in the mechanistic account but is free from many of the traditional assumptions about levels. The levels of mechanisms are perspectival in the sense that the levels are dependent on the explanatory target. Macro-level facts are explained by appealing to micro-level processes, entities, and relations, but these items belong to the micro level just because they are required for the full explanation of the macro fact, not because they belong to some predetermined micro level. Whatever is needed for explaining the macro fact is regarded as belonging to the same level. However, there is no guarantee that these components would always be at the same level in all possible explanatory contexts. Nor it is obvious that the micro-level entities and processes that account for these components would be in any simple sense from the same level. For every hierarchy of mechanisms, there is a clear hierarchy of levels of mechanisms, but these levels are local. There is no reason to assume that separate hierarchies of mechanism levels would together produce the neatly delineated and comprehensive levels of nature assumed in the traditional layer-cake model (Craver 2007).

These views have a number of interesting consequences for traditional ways of thinking about reductive explanation and the explanatory role of microfoundations in the social sciences. For example, once we give up the outdated deductive model of theory reduction, many of the traditional fixations of the methodological individualism debate simply become meaningless. For example, there is no need to provide individualistically acceptable redefinitions of macro-social notions because the explanation of macro facts is no longer conceived as a logical derivation. Similarly, the search for any bridge laws between theories becomes pointless. This has the consequence that the key anti-

reductionist argument about multiple realization loses much of its significance. From the point of view of mechanistic explanation, multiple realization is simply an interesting empirical observation that does not pose any serious threat to the possibility of explaining macro properties in terms of micro properties and relations. Just as the sciences have learned to live with the fact of alternative causes, they can learn to live with the phenomenon of multiple realization.

The advocates of the mechanism-based approach in the social sciences have noticed some of these consequences. For example, they have largely given up the old ideas about reductive explanation and have instead emphasized the importance of microfoundations (Elster 1989; Little 1991). However, I do not think that all the implications of the mechanistic perspective have been taken into account in the philosophy of social sciences. This is visible, for example, in the fact that quite often the mechanistic approach is associated with methodological individualism (Elster 1989). Similarly, much of the debate about micro-macro relations is still focused on arguments that are based on a premechanistic understanding of reductive explanation (Sawyer 2005; Zahle 2006).

The aim of this chapter is to sketch what a consistently mechanistic way to think about micro-macro relations would look like and to show that some of the key presuppositions of the traditional debate about methodological individualism should be given up. One of these is the assumption of a comprehensive, unique, and privileged individual level. The notion of *comprehensiveness* refers to the idea that there is a consistent and well-defined individual level that is sufficient to cover all social phenomena and that would serve as a reduction basis for all nonindividual social notions.

Uniqueness refers to the assumption that in all social explanations, the micro level would always be the same level, for example, the level of intentional rational action. Finally, the notion of *privileged* refers to the presumption that explanations in terms of this special level have some special explanatory qualities that set them apart from explanations from other levels. In the following, I will challenge all three assumptions and argue that once they are given up, we can approach the micro-macro issues in the social sciences in a more clear-headed manner.

2.3 Rethinking the macro

A popular argumentative strategy among anti-individualists has been to borrow ideas from the philosophy of mind. They are inspired by the arguments for nonreductive materialism, so they build their argument based on an analogy with the mind-brain relation. Given that these arguments are not very mind specific—it is a general practice just to talk about M- and P-predicates—their appeal is understandable. The ideas of supervenience and multiple realization seem to provide a neat way to argue against reductionism, at least if one accepts the traditional idea of derivational reduction. While there are reasons to suspect that the notion of supervenience is less illuminating than is often assumed (Horgan 1993; Kim 1993) and that the traditional view of reduction does not completely collapse under multiple realization (Kim 1998), we can set these issues aside as their relevance presupposes a premechanistic account of reductive explanation. Here I want to focus on the mind-brain analogy as I think it is misleading.

The mind-brain analogy is inappropriate because it mischaracterizes the nature of the social scientific micro-macro problem. The central problem in the philosophy of mind

is to figure out how the explanations provided by psychological theories that employ mental concepts are related to the accounts of the brain's working provided by the neurosciences. The challenge is to relate two levels of description that are fundamentally talking about the same thing. The (nondualist) antireductionist position does not typically challenge the causal sufficiency of the neural-level facts. The setup is quite different in the social scientific micro-macro debates.

The problem in the social sciences is not that of bridging a comprehensive and exhaustive individual-level understanding of social processes (the analogue to the idealized knowledge of the brain) to a more social or holistic description (that would be analogue to the idealized psychological theories employing the mental vocabulary). It is typical for anti-individualists to challenge the causal sufficiency of individual facts. They often claim that the facts about individuals allowed by the individualist are either not sufficient to account for all social facts or the individualists are cheating by accepting facts that are not properly individualistic. This is because the issue is not really that of relations between two comprehensive (and potentially competing) levels of description, but that of seeing how local facts about individuals and their social interactions are related to large-scale facts about groups, organizations, and societies. So, the relation is really more like the one between the whole brain and its parts than the mind and the brain. While this contrast is useful for highlighting the inappropriateness of the mind-brain analogy, I do not want to develop it further as there are many problems with the organ-society analogy. It is better to skip all the brainy analogies and to take a fresh look at the micro-macro problem as the social scientists face it.

A useful starting point is the observation that macro social facts are typically *supra-individual*: They are attributed to groups, communities, populations, and organizations, but not to individuals. There might be some attributes that apply both to individuals and collectives, but typically macro social properties, relations, and events are such that they are not about individuals.

Another salient feature of many social micro-macro relations is the part-whole relationship. One way or another, the macro social entities *are made of* their constituting parts. Usually this relation of constitution is more than mere mereological aggregation or simple material constitution. First, many social wholes are composed of a heterogeneous set of entities; there are intentional agents, their ideas, and material artifacts. Second, in all interesting examples of social wholes, the *relations* between the components play an important role. (Similarly, often the relations between social wholes and between the social whole and its environment are also important.) However, the important thing is that the part-whole relationship makes it possible to see the micro-macro relation as a question of scale: The difference between micro and macro is the difference between small- and large-scale social phenomena.

I do not propose that we can simply define the micro-macro contrast as an issue of scale. All differences in scale do not constitute a meaningful micro-macro relation, and the heterogeneous nature of macro social facts makes it difficult to characterize the additional requirements for their defining features. However, I do want to suggest that it provides a fruitful way to think about micro-macro relations and an antidote for the tendency to see parallels in the philosophy of mind.

Thinking of the micro-macro issue as an issue of scale makes it possible to conceive of it as being without a unique micro level. Whereas the contrast between 'individual' and 'social' levels is categorical, the contrast between small and large is relative and allows a continuum of various sizes. Depending on the application, the micro entities could be individuals, families, firms, or groups. This flexibility is in accordance with the way social scientists think. They do not assume that micro is always about one specific set of entities.

Another consequence is that whether an attribute is a macro or micro property depends on what it is contrasted with. A friendship relationship is a macro property from the psychological point of view, but a micro property when considered from the point of view of the social networks within a community. Rather than being set a priori, the contrast between micro and macro depends on one's explanatory interests. For example, international politics and organizational sociology construct the micro-macro contrast quite differently. In the former, states and other organizations are often treated as individuals, whereas in the latter, the organizations and their properties are the macro reality to be explained. Similarly, an economist studying market processes can treat firms and households as the micro level, while for disciplines such as industrial organization and family sociology, they are the macro items that require explanation.

From the point of view of a mechanistic philosophy of science, this flexibility is not surprising. The same dependence of levels on epistemic concerns is also observable in the biological sciences. The cell biologists or neuroscientists do not think in terms of comprehensive or unique micro levels either. The levels of mechanisms found there depend on the explanatory concerns, not on a priori ontological considerations. This is

not worrisome for the mechanistic point of view, as the key assumption is that whatever is found at the micro level can always be turned to a macro-level *explanandum* for another set of enquiries.

The social macro properties do not constitute a unified kind, so it makes sense to characterize them with a sample of examples rather than with a general definition. The following classification of typical sociological macro social properties is not intended to be exhaustive of sociology or the social sciences in general. There are many parts of macro social reality that fall between my four categories. However, I hope the four examples can be used to illustrate the applicability of the scale perspective.

2.3.1 (1) Statistical properties of a population

A major concern for sociology is the various statistical attributes of populations. Among these are *distributions* and *frequencies*. Sociologists are interested in both distributions of attributes to various kinds of individuals and distributions of individuals with certain attributes to social positions and spatial locations. For example, when they are studying the ethnic segregation of cities, comparing societies in terms of inequality, or describing the social stratification of a society, they are attempting to account for distributions. Another relevant property of distributions are frequencies. Sociologists are interested in typical, rare, dominant, or marginal behaviors, beliefs, or attitudes within a specified population. Similarly, they are interested in ratios of attributes such as unemployment or incarceration within the population. So, when sociologists are studying changes in racial prejudices over time, comparing the level of conformism between communities or

tracking the changes in the level of union memberships, they are interested in explaining frequencies.

All these statistical macro social properties are inferred (or estimated) from data about the members of a population. There is no other way to access them. However, it does not make any sense to attribute these properties to individual units. Another important thing about these macro social facts is that the units of these statistics do not have to be individuals; they can as well be families or firms. It is noticeable that statistical macro properties are in no way dependent on the members' beliefs and attitudes about them. The members of the population can have false, or even crazy, beliefs about distributions and frequencies that characterize their own society.

While the statistical properties of populations usually only serve as *explananda* in the social sciences, they do have some legitimate and nonreducible explanatory uses. For example, in the cases of frequency-dependent causation (e.g., cases in which the causal effect of an individual having a certain property depends on the frequency of that property in the population), the statistical facts are the crucial difference makers. Similarly, in many social scientific explanations, the correlations between various variables (for example, wealth, education, taste, and place of residence) play an important role in accounting for individual differences in behavior and attitudes. Both of these cases are quite easily conceived as cases of larger-scale facts influencing smaller-scale phenomena, while other ways to think about levels are not as natural.

2.3.2 (2) Topologies of social networks within a population

Sociologists are also interested in relations and interactions between individuals. When considered together, these relations constitute networks of social relations within the population. A social network can be regarded as a map of all of the relevant ties between the members of a specified population. When sociologists are studying the spread of information within an organization, comparing groups with respect to their level of network clustering or analyzing the brokering opportunities of an individual occupying a structural hole (i.e., a position between two networks that are not otherwise connected), they are examining social networks.

The importance of social networks is increasingly being recognized in the social sciences, and social network analysis is becoming increasingly popular in various social sciences. Social network analysis is based on the observation that networks have many interesting (formal) properties, such as centralization, cohesion, density, and structural cohesion (Scott 2000). While the social network is inferred from knowledge about individual relationships, the properties of the network are prototypical macro properties. It does not make any sense to apply these attributes to individual nodes of the network. Similarly to statistical properties, the units of network analysis are flexible. There is no requirement that the nodes of the network (the members of the population) are persons. They can also be groups, families, organizations, or even states.

The properties of social networks serve both as the *explananda* and the *explanantia* in sociology. As an example of the latter, consider the notion of a structural hole (Burt 1992), which is used to explain the differences in agents' ability to access information and in their opportunities to influence social processes. In these explanations the structure of the network plays an irreducible role, and it is quite natural to think of the

social network as a large-scale social phenomenon influencing local interactions between individuals. In contrast, it is very difficult to think about them in terms of social and individual levels. As social networks are attributes of the population, it would be quite a stretch to call social networks individual properties. But if they are macro-level properties, what would be the individual-level properties that could be regarded as their bases? Collections of relevant individual relations, one might suggest, but that would be just a vague way to talk about networks. Things are simpler if one does not have to bother with such questions. A network is simply a more extensive entity that constituted by more local relations and it can have properties that are not properties of its components.

2.3.3 (3) Communal properties

By communal properties I refer to social scientific notions that apply to specific communities, but not to isolated individuals. Among these notions are such things as culture, customs, social norms, and so on. For example, cultural differences are primarily between groups, not between individuals. Similarly, social norms and customs are properties of communities—attributing them to solitary individuals does not make sense. Many of these notions do not have precise definitions, and their explanatory uses are often confusing (Turner 1994; Ylikoski 2003), but they do have an important role in the social sciences.

While communal properties are attributed to groups, they are quite straightforwardly based on facts about individuals. Underlying these notions is the idea that the members of a group share certain beliefs, expectations, preferences, and habits. However, it is crucial that the sharing of these individual attributes is not purely

accidental: The members have these individual properties because the other members of the group have them. The sharing of these properties is due to continuing interaction. For example, the existence of a social custom presupposes that the novices learn specific expectations and habits when they become members and that the members of the group stick to these expectations and habits because others also do so. Underlying the (relative) unity of a culture are facts about the shared origins of the ideas, values, and practices of the members and their constant interaction with each other. Similarly, the cohesion of a culture is based on the frequency of interactions with the group and the rarity of interactions with outsiders, not on any kind of higher-level influence on individuals.

Descriptions of customs, social norms, and cultures are always based on idealization and abstraction. Members of a community never have exactly the same ideas, preferences, or routines. That would be a miracle, given what is known about human learning and communication (Sperber 1996). There is always some variation among the members, no matter how comprehensive the socialization processes are. However, these idealized descriptions are still useful. They draw attention to features of the group that are typical and salient when it is contrasted with some other group.

Although communal properties, as I have described them, are tied to a social community defined by frequent interactions, the boundaries of these communities are fluid. This makes it possible to describe culture on various scales—for example, on the levels of a village, a local area, and a nation. However, descriptions on larger scales are bound to be more abstract and less rich in detail as individual variation takes its toll. The same flexibility that characterizes statistical and network properties applies also to

communal properties, which can also be attributed to nonpersonal units. For example, it is possible to describe social norms that govern interactions between organizations.

When we consider communal properties as *idealizing abstractions from shared individual properties*, there is no need to refer to them as any kind of autonomous level of reality. They just describe more extensive facts than descriptions of the individual attitudes, habits, and preferences that constitute them. The scale perspective also appears natural when the explanatory use of communal properties is considered. For example, when we are explaining the behavior of an individual by appealing to social norms, we are referring to larger-scale facts about the group members that are causally relevant to the micro-level behavior. There is no need to postulate a separate realm of norms to understand what is happening. It is just that the expectations and responses of the other group members influence the individual's judgments about appropriate behavior.

2.3.4 (4) Organizations and their properties

Organizations such as states, firms, parties, churches, and sport clubs are important parts of the social reality. While the community that is the basis for communal properties is not often clearly demarcated, a clear demarcation is often the case with organizations. They usually have specified criteria for membership, at least for the operational members. They also have rules that define the rights and duties of members and the roles of various functionaries. These (written or nonwritten) rules make it possible for organizations to have stability and continuity, so that it makes sense to talk about their continuing existence when their functionaries are replaced and the members change. Furthermore,

many organizations exist (and are defined) in the context of other organizations, so one has pay special attention to context when attempting to make sense of organizations.

Organizations as entities can have many properties that are not properties of their members. They can even have goals that are not the personal goals of their members, and some organizations are treated as legal persons. This has convinced many that organizations are real entities that should be treated as a separate ontological category. I do not have strong opinions about issues of ontological bookkeeping, as it is remembered that organizations are human artifacts that are always made of persons, their ideas about the rules, and often, of material artifacts. Whatever the organization does, is done by its members in its name. It is of crucial social importance whether an action, for example, a questionable comment, was made as a representative of an organization or as a private person. But these are facts about the status attributed to the behavior, not about the two completely different entities producing the behavior.

When a person causally interacts with an organization, she interacts with other persons (although this interaction is increasingly mediated via material artifacts such as ATM machines). There is no downward causal influence from a higher level. Everything happens at the same level; it is just that the intentional attitudes and relations of a larger group of people are important to the details of the local situation. Similarly, the influence of the organization on its members happens through other members, no matter how high up some of the members are in the organizational hierarchy. While the rules (and their interpretation by others) are external to any individual person, there is no need to posit them as a separate ontological category. These observations suggest that even in the case of organizations, the layer-cake model of the social world is not very illuminating. What

is interesting about organizations is the habits and mental representations of their members, the resources they control as members of the organization, and their (materially mediated) interactions, not some higher ontological level.

Again it is good to return to real social scientific questions. They concern issues such as: How do large-scale collective enterprises—for example, organizations—manage (or fail) to achieve certain things? What kinds of unintended consequences do these collective activities have? How does a membership in such collective enterprises influence the individual members? The explanatory answers to these questions often refer to organizations and their properties, but there is no problem in conceiving them as large-scale things influencing smaller-scale things or other large-scale things.

These examples of macro social facts suggest a kind of flat view of society in which the difference between micro and macro is one of scale, not of different levels. The large-scale facts about distributions, frequencies, interactions, and relations have an irreducible explanatory contribution to make, but there is nothing comparable to the mind-brain relation. As a consequence, the metaphor of levels that underlies the layer-cake model does not really help to make sense of the issues that social scientists addressing social macro facts are facing. Giving it up will have a number of beneficial consequences.

First, there are some philosophical advantages. As I will argue in the next section, once we give up the image of levels, we get rid of the problem of causal exclusion that arises from the image of causally competing levels. There is no problem of downward causation as there are only causal influences from large-scale things to small-scale things and descriptions of large-scale things at various levels of abstraction. The problem is

replaced with the more down to earth problem of explanatory selection: Under which

description can we formulate the most robust claims about counterfactual dependence?

Secondly, we no longer have to face the problem of finding an acceptable definition of

the comprehensive individual level so that we can argue for or against methodological

individualism. We can start analyzing real social scientific explanations instead and focus

our attention on the possible contributions that large-scale things make to those on a

smaller scale and what kinds of causal mechanisms mediate these influences.

This change in framing also has some advantages when considering relations

between disciplines. The division of labor between psychology and the social sciences is

justified by differences in scale and the importance of large-scale relations and

interactions, not in terms of independent and autonomous levels of reality. This

guarantees that the social sciences will never be reduced to psychological sciences.

However, thinking in terms of scale also cuts down the false aspirations of disciplinary

autonomy. When the social scientists are denied their own autonomous level of reality,

the ideal of completely psychology-free social science becomes less appealing. It should

be an empirical matter whether the details of human cognition matter for social

explanation. It might be that is some cases it makes good mechanistic sense to

incorporate some processes on the sub-personal level in the explanatory theory. I will

return to this possibility in the final section.

2.4 Causation, constitution, and microfoundations

One prominent idea in the recent philosophy of biology debate about mechanisms has not

been employed in the philosophy of social sciences debate.[1] This is the distinction

between causation and constitution. Although the difference between constitutive and causal explanation has been noted earlier (Salmon 1984; see also Cummins 1983), it has only recently become a topic of systematic study (Craver 2007).

Both causation and constitution are relations of dependence (or determination), and they are easily confused. However, there are some crucial ontological differences. Causation is a relation between events; it is about changes in properties. Causation takes time, so we talk about causal processes. Finally, causation is characterized by the asymmetry of manipulation: The effect can be manipulated by manipulating the cause, but not the other way around (Woodward 2003).

In contrast, constitution relates properties. The properties (and relations) of parts constitute the properties of the system (sometimes also the relations to the environment are important). The whole is *made of* its parts and their relations. Unlike causation, constitution does not take time, and we do not talk about the process of constitution. Furthermore, the *relata* of constitution are not 'independent existences' (as Hume called them). For this reason we cannot characterize the relation of constitution with the help of the asymmetry of manipulation. For example, the molecular structure of glass constitutes its fragility: To be fragile is to have a particular molecular structure; the fragility is not a consequence of the molecular structure. However, there is another sort of asymmetry: the asymmetry of existence. The parts preexist the system in the sense that the parts can exist independently of the system, but the system cannot exist independently of its parts (although the system can exist independently of particular parts).

An interesting sort of regress characterizes both causation and constitution. In the case of causation, we talk about chains of causation. This is based on the idea that for

every event that is a cause, there is another event that is its cause. A similar idea applies to constitution; we assume that all parts can be further decomposed into their parts and their organization. We could call these chains of constitution. Now a tricky question is whether there exists a first cause that is not itself caused, and a similar problem can be stated concerning the ultimate building blocks of reality, but in this context we can leave them aside. There is no danger that such ultimate things will show up in the social sciences. However, these regress properties create chains of explanations, which are relevant from the point of view of the social sciences. The crucial thing in this context is to understand that although there is always an explanation for every social scientific explanatory factor, this does not imply that their explanatory status depends on us knowing the explanation for them. Both in the case of causation and constitution, an explanation presupposes that the *explanans* facts are the case, not that we have to have an explanation for those facts. I will return to this issue in the next section.

Explanation is about tracking relations of dependence. Although metaphysically the relations of constitution and causation are quite different, in terms of explanation the basic principles are quite similar. Both explanations attempt to track networks of counterfactual dependence. A causal explanation tells us how *the antecedent events* and *their organization* (timing and location) bring about the event to be explained. In contrast, a constitutive explanation describes how *the properties of the components* and *their organization* give rise to the system's properties.

In both cases we are looking for the difference-makers: The criterion of explanatory selection is counterfactual. As the precise *explanandum* is best characterized in contrastive terms (why x is the case rather than x*), we are interested in the differences

that would have made the difference we are interested in ([Woodward 2003](); [Ylikoski 2007](); Northcott this volume). In the case of causation these differences are in antecedent events; in the case of constitution these differences are in the properties of parts (or in their organization). Also in both cases it makes sense to ask a further question: Why does the counterfactual dependence hold? The answers to these questions will in both cases draw from the same body of mechanical knowledge, so it is understandable that in the philosophy of biology debates both explanations are called mechanical explanations. So, despite the important metaphysical differences, the same basic ideas about explanation can be applied to both cases.

Not only are the principles of explanatory relevance similar, so are the explanatory questions. This leads easily to confusion. Consider the question: "Why is this glass fragile?" The question is ambivalent: It could either mean "How did the glass become fragile?" or it could mean "What makes the glass fragile?" The first question is causal; the latter question constitutive. The answer to the causal question will tell us about the causal history of the glass—it will specify the crucial features of the process that led to the object being fragile rather than robust. The answer to the constitutive question will not focus on earlier events. It will detail the relevant aspects of the object's molecular structure that makes it fragile. So while the explanation-seeking questions may look the same, the request for explanatory information is quite different. Without a clear understanding of the differences between causation and constitution, some confusion is bound to occur. This is also the case in philosophy of social sciences. For example, it is quite a different thing to explain how a regime became stable than to explain what makes it stable. While some of the facts cited by both explanations might be the same, they are

addressing different *explananda*: One is focused on how the causal capacity was acquired and the other on the basis of that causal capacity. A social scientist is usually interested in both questions, but she should not confuse them with each other.

For all social macro properties, one can ask both constitutive and causal why- and how-questions. (Although for some statistical properties the constitutive questions are relatively trivial.) The first sort of questions asks how the macro properties are constituted by the micro-level entities, activities, and relations. The aim is to track how the details of macro-level facts depend on the micro details. The question is often how the macro facts would have been different if some of the micro facts had been different in some specific way. These questions can also be characterized in terms of interventions: How would the macro facts change if some of the micro facts were changed? Notice that here intervention is a causal notion (all change happens in time), but the dependence of interest is constitutive.

A clear example of constitutive explanation is an explanation for the difference in the problem-solving capacities of two groups. The crucial difference might be in the properties of the members, such as their intelligence or social skills. Alternatively, the pivotal factors might be the informal social norms that characterize the interactions within the group or its formal organization. Of course, the explanation may also be found in some combination of these factors. Just like in this example, the usual *explananda* of constitutive explanations are causal capacities and dispositions of the whole. The constitutive explanation tells us what gives the whole (population, group, organization, or society) those properties, and the answer is found in the causal capacities of the parts and their organization.

The *explanantia* in constitutive explanations are always at the micro level. As the explanation attempts to capture what the whole is made of, an appeal to the properties of the whole does not really make sense. In this sense, the methodological individualists, and other reductionists, have been on the right track. On the other hand, the explanation of macro properties does not in any way diminish their reality: The wholes are as real as their parts. This implies that those methodological individualists who have suggested that a micro explanation somehow eliminates the macro properties are either metaphysically confused or just choosing their words badly. The talk about macro reducing to micro makes as little sense as the talk about reducing effects to their causes.

The causal questions about the macro social properties are concerned with their origin, persistence, and change. These explanations are tracking counterfactual dependencies between events. How would have the outcome been different if some of the causes had been different in some specified manner? What kind of difference would an intervention on some antecedent facts make? The *explanantia* in these causal explanations are always antecedent events.

This is the context in which confusion between constitution and causation can create trouble. If we are considering simple causal statements about causal dependence, individualists tend to make the claim that the causes have to be at the micro level. However, nothing in the notion of causation implies that the real causal work is always to be found at the micro level. Of course, the notion of constitution implies that every time we have a cause at a macro level, we also have micro level facts that constitute it. If we stick to the counterfactual criterion of explanatory selection, as I think we should, there is no a priori reason to privilege micro-level causes (Woodward 2003, 2008). It is sufficient

that there is an appropriate counterfactual dependence between the macro variable and the *explanandum*. Of course, in many cases the justification of a claim about this causal dependence might require some knowledge of the underlying mechanisms. However, this observation about the justification of a causal claim should not be confused with the claim itself. Similarly, although adding mechanistic details to the explanation will involve references to micro-level processes, this does not imply that the macro facts will lose their explanatory relevance. They will still be possible difference-makers and legitimate explanatory factors. In other words, although the information about the relevant mechanistic details improves the explanation significantly, it does not remove the causal relevance of the initial invariance involving macro-level facts.

In the counterfactual account of causal relevance, the location of explanatory relevance at the micro or macro level is a contingent matter that depends on the *explananda* that one is addressing. There is no reason to assume that the most invariant counterfactual dependence (with respect to the contrastively specified *explanandum*) will always be found at the micro level. Similarly, one has to give up the often presented suggestion that levels of explanation should match so that macro would always explain macro and micro would always explain micro. The issues of explanatory relevance (how the explanatory factors are selected, at which level of abstraction they are described, etc.) are always determined by the facts of the case and the details of the intended *explanandum*, not by generic philosophical arguments.

2.4.1 The proper role of microfoundations

Is the above argument about the legitimacy of macro-level causal facts compatible with the mechanistic call for microfoundations? I want to argue that it is fully compatible with the core ideas of mechanism-based thinking. Contrary to the common assumption, the point of mechanistic microfoundations is not that we have more real causes at the micro level, but to have a better grasp of the explanatory dependence underlying the causal relation involving macro variables. Consequently, the advocates of mechanism-based explanations should not call into question the reality of macro-level causal relations. Instead, they should emphasize the importance of microfoundations for understanding these dependencies. There are a number of reasons why microfoundations are important.

First, all causal relations involving macro properties are mechanism-mediated causal relations. Understanding how the dependence involving macro variables is constituted helps to understand why that particular dependence holds ([Ylikoski 2011](#)). It also integrates the piece of causal information contained in the macro-level generalization to other pieces of explanatory knowledge (Ylikoski and Kuorikoski 2010). This is certainly a form of explanatory understanding that we should be interested in if we take the notion of explanatory social science seriously.

However, the utility of this information is not limited to the expanded theoretical understanding. It also often tells about the conditions under which the causal dependence in question will hold. There are three dimensions to this knowledge. First, there is knowledge about the range of values of the *explanandum* variable that are possible without the dependence breaking apart. Second, there is knowledge about the sensitivity of the dependence to changes in background conditions. Finally, there is possible knowledge about alternative interventions that could bring about similar effects. Without

knowledge of these issues, the explanatory use of the macro-level explanatory generalization can be very risky business. It is very difficult to extrapolate to other cases without understanding the background mechanisms (Ylikoski 2011; see also Cartwright, this volume, Kincaid, this volume).

Apart from an expanded understanding and the security of an explanatory claim, the insight into the underlying mechanisms might also help to improve the explanatory generalization. With the help of a mechanistic understanding, one might be able to make the *explanandum* more precise or to reformulate the explanatory generalization in such a manner that it allows a broader range of values of the *explanandum* variables or background conditions (Ylikoski 2011).

These considerations justify the presumption that microfoundations are important for proper explanatory understanding. However, they do not demolish the explanatory relevance of macro facts. On the contrary, they put them in the right context as the mechanisms bridge the large-scale micro facts to causal interactions between persons and to their decision-making processes. I think this is the point James Coleman (1990) attempted to make with his often misunderstood graph.

[Insert Figure 2.1 here: Macro-Micro Linkages]

Following Hedström and Swedberg (1998, 23), I refer to the arrows in figure 2.1 as situational mechanisms (arrow 1), action-formation mechanisms (arrow 2), and transformational mechanisms (arrow 3). The situational mechanisms describe how social structures constrain individuals' action sand cultural environments shape their desires and beliefs, the action-formation mechanisms describe how individuals choose their preferred

courses of action among the feasible alternatives, and the transformational mechanisms describe how individual actions produce various intended and unintended social outcomes.

Coleman was critical of nonmechanistic explanations that remain at the level of macro regularities. However, there is no reason to assume that he was denying the causal relevance of macro social facts. Rather, his point was to make it clear that proper sociological understanding requires that we understand both the mechanisms by which large-scale social facts influence the local decision-making processes of individual agents (the situational mechanisms) and the mechanisms by which individual actions create and influence macro social facts (the transformational mechanisms). He was calling for mechanisms that bridge the levels, not just descriptions that somehow reduce the macro facts to individual level facts. Only when we understand the relevant mechanisms, do we have a satisfactory theoretical grasp of the social phenomena in question.

Coleman's criticism of Weber's (partial) explanation of the emergence of modern capitalism in Western Europe illuminates these points. Weber started with an idea that was commonplace in late nineteenth-century Germany: There is a close connection between Protestantism, entrepreneurism, and the rise of capitalism. To substantiate this vague explanatory suggestion, Weber asked what changes the emergence of Protestantism brought about in the beliefs, desires, and communal practices of individual agents. This question has both causal and constitutive dimensions that are not clear in Coleman's analysis. However, Coleman's focus is on Weber's second causal question: How did these changed life practices of individuals influence economic activities and institutions and how did these changes in turn facilitate the formation of modern

capitalism? Coleman's central point was that Weber was not clear enough about this last passage of the causal chain. He was not able to give a sufficiently clear account of the transformative mechanisms that connected the Protestant ethic to the rise of modern capitalism. In other words, Weber was not able to show how the changes at the micro level (the life practices of Protestants) bought about a major macro-level outcome (the early forms of modern capitalism). As the crucial mechanism is lacking, so is the legitimacy of Weber's causal claim about history.

Here it is important to see the difference between the justificatory and explanatory roles of mechanisms. Coleman's analysis shows why it is legitimate to challenge Weber's causal claim. Knowledge of the causal mechanisms have an important role in the justification of historical causal claims, so pointing to the missing details of the causal chain constitutes a challenge to the legitimacy of the causal claim. However, this criticism of a singular causal claim does not imply that Coleman generally considers macro-level facts to be nonexplanatory or causally impotent. He is simply challenging the justification of this particular historical hypothesis.

2.5 Intentional fundamentalism

Arguments for methodological individualism often appeal to the special explanatory status of intentional explanations. I call this position intentional fundamentalism. According to intentional fundamentalism, the proper level of explanation in the social sciences is the level of the intentional action of individual agents. The intentional fundamentalist assumes that explanations given at the level of individual action are especially satisfactory, fundamental, or even ultimate. In contrast to explanations that

refer to supra-individual social structures, properties, or mechanisms, there is no need to provide microfoundations for intentional explanations. They provide rock-bottom explanations. In other words, according to intentional fundamentalism, the intentional explanations of individual actions are *privileged* explanations.

Although intentional fundamentalism can take various forms, it is often related to rational choice theory. French social theorist Raymond Boudon (1998, 177) expresses the idea clearly: "When a sociological phenomenon is made the outcome of individual reasons, one does not need to ask further questions." The idea is that in the case of supra-individual explanations there is always a black box that has to be opened before the explanation is acceptable, but in the case of intentional explanation there is no such a problem: "The explanation is final" (Boudon 1998, 172). Diego Gambetta appeals to same sort of finality (1998, 104): "Not only will a rational choice explanation be parsimonious and generalizable; it will also be the end of the story."[2]

My claim in this section is that intentional fundamentalism is not compatible with the causal mechanistic account of explanation. As intentional fundamentalism is often advocated by rational choice theorists and as many believe that rational choice explanations are the best examples of mechanical explanations in the social sciences, this incompatibility claim is of some interest. If my argument is valid, it suggests that the relation of rational choice theory and a mechanism-based philosophy of science requires some rethinking. It also implies that one common argument for methodological individualism is much less credible than is commonly assumed.

2.5.1 The regress argument

To make sense of intentional fundamentalism, we should start with *the explanatory regress argument for methodological individualism*. Methodological individualists often make the case that nonindividualist explanations are either explanatorily deficient or not explanatory at all. At most, they allow that explanations referring to macro social facts are placeholders for proper (individualistic) explanatory factors. In this view, the explanatory contribution of supra-individual explanations is at best derived: They are explanatory because they are (in principle) backed up by a truly explanatory story. This is the regress of explanations argument: Unless grounded at the lower level, explanations at the macro level are not acceptable. The underlying general principle is the following:

> [P] A genuine explanation requires that the *explanans* is itself explained or is self-explanatory.

In short, the explanatory buck has to stop somewhere.

The principle [P] is general, and it raises the possibility of an explanatory regress that is only halted at a fundamental (physical) level. This would be highly unintuitive, so for the intentional fundamentalist the buck stops at the level of (self-interested) rational intentional action. This level is treated as inherently understandable, as shown in the above quotations from Boudon. The inherent intelligibility of intentional action explains why the search for microfoundations should stop at the level of the individual. The special status of intentional explanation also makes the explanatory regress argument safe for the methodological individualist: He can use the argument's full force against anti-individualists who cannot make a similar claim about a privileged status, and it does not challenge the legitimacy of his favored explanatory factors.

The fundamentalist argument for individualism fails for a number of reasons. The first reason is that the principle [P] is not valid. The explanatory relation between the *explanans* and the *explanandum* is independent from the question of whether the *explanans* is itself explained. An explanation of X in terms of Y presupposes that Y is the case, but it does not presuppose that Y is itself explained. Of course, it would be great also to have an explanation for Y, but this is a separate issue from the legitimacy of the explanatory relationship between Y and X. The distinctness of these issues implies that the regress does not begin.

Why whould anyone believe in [P]? One plausible suggestion is the following: The belief in [P] arises from a straightforward confusion between justification-seeking and explanation-seeking why-questions. It makes sense to ask how well justified are those things that one appeals to in justification of one's beliefs. It also makes sense to ask whether one is justified in believing the things that one appeals to in one's explanation. However, justifying one's belief in Y, is not the same as explaining why Y is the case.

2.5.2 Intentional explanations without a special status

Another reason for the failure of the regress argument is that intentional explanations lack the special properties assumed by the argument. If one accepts the mechanistic account of explanation, as many advocates of rational choice sociology do, such a special status does not make any sense. The assumption that human deliberation is a black box that should not be opened is more in line with nineteenth-century hermeneutic romanticism than with causally oriented social science. Of course, the chain of mechanistic explanations will

end somewhere (if there is such a thing as a fundamental level), but that stopping point is not the level of individual rational action.

A mechanistic explanation appeals to micro-level processes, but nothing in the notion of mechanistic explanation implies that these micro things would always be facts about the intentional actions of individuals. Mechanisms that cite supra-individual entities or properties are certainly possible (Mayntz 2004). For example, various filtering mechanisms that are analogical to natural selection are difficult to understand other than as population-wide processes, and when the units that are selected are organizations (for example, firms), it is natural to conceive the mechanism as supra-individual. Similarly, the crucial parts of the explanatory mechanism could well be located below the level of intentional psychology. For example, various facts about human information processing—for example, implicit biases (see Kelly and Mallon, this volume)—could well be relevant for explanatory understanding of intentional action. There is no valid reason to give up mechanistic thinking in the case of intentional action.

Another reason to challenge intentional fundamentalism is the implicit realism of mechanistic thinking. For mechanists, explanation is factive. It is not enough that the explanation saves the phenomenon: It should also represent the essential features of the actual causal structure that produces the observed phenomena. So, if the explanation refers to the goals, preferences, or beliefs of agents, the agents should indeed have those mental states. Mere as-if storytelling does not suffice for a mechanistic explanation as it does not capture the relevant parts of the causal process. This realist attitude goes against the instrumentalist attitude common among many rational choice theorists. The fact that one can rationalize any behavior does not imply that those rationalizations are also the

correct causal explanations for those behaviors. Similarly, the human fluency in coming up with intentional accounts for our behavior is not a reason for regarding them as superior explanations.

It is important to understand the limited nature of my argument. I am not denying that intentional explanations are, and will be, an indispensable part of the social scientific explanatory repertoire. For me, intentional explanations are legitimate causal explanations. Furthermore, the intentional attitudes of individuals play an important role in most mechanism-based explanations of social phenomena. The only thing I am challenging is the supposed special explanatory status of intentional or rational accounts of human action. In the mechanistic account of explanation, the importance of certain sorts of explanatory factors is not a basis for their privileged status.

Neither should my rejection of intentional fundamentalism be regarded as a wholesale attack on the use of rational choice theory. For many social scientific purposes, a rather simple version of intentional psychology is both preferable and sufficient. For example, when one is attempting to make sense of social complexity, it is understandable that social scientists attempt to keep the psychological assumptions of their models very simple. Such idealizations are fully legitimate if they do not lead to a gross misrepresentation of the causal mechanism under consideration. However, the practical necessity of these idealizations does not constitute a justification for accepting intentional fundamentalism.

Furthermore, my argument should not be regarded as an argument against the claim that there should exist a division of labor between the social sciences and the sciences of cognition. However, it follows from the flexibility of mechanistic levels that

the boundaries of this division of labor are adjustable and not fixed. It is inherent in the idea of mechanistic explanation that all the gaps between levels of analysis[3] are ultimately to be bridged by mechanistic interfield theories. So the challenge for the social sciences is not to define their objects of study in such a way that they are in no way touched by psychological sciences, but to look at ways in which social and cognitive mechanisms can be meaningfully combined. This is not as easy as it sounds, as recent attempts to combine neuroscience and economics show (Kuorikoski and Ylikoski 2010).

2.6 Conclusions

In this chapter, I have attempted to show what consequences the mechanism-based account of explanation would have on issues traditionally discussed under the title of methodological individualism. Borrowing some ideas developed by philosophers who have studied the mechanistic explanation in the biological sciences, I have argued that we should give up the notion of a unique, privileged, and comprehensive individual level that has been a presupposition of the individualism debates. In addition, I have argued that rather than employing metaphors borrowed from the philosophy of mind for micro-macro relations, we should pay closer attention to how real macro social facts figure in social scientific theories and explanations. There the micro-macro issue is more an issue of bridging large-scale social facts to small-scale social interactions rather than that of finding a way to see relations between autonomous levels of reality.

Notes

1. There are some exceptions. For example, Wendt (1998) distinguishes between causation and constitution. However, his discussion of constitution is very confused. His

notion of constitution covers not only the constitution of causal capacities, but also causal preconditions, definitions, and other conceptual relations. The standard philosophy of science notion that I am using is limited only to the constitution of causal capacities.

2. The key issue here is not whether these authors would ultimately subscribe to intentional fundamentalism. I am only claiming that in these passages they argue as if intentional fundamentalism is correct.

References

Bechtel, William. 2006. *Discovering Cell Mechanisms. The Creation of Modern Cell Biology*. New York: Cambridge University Press.

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

Boudon, Raymond. 1998. "Social Mechanisms Without Black Boxes." In *Social Mechanisms: An Analytical Approach to Social Theory*, eds. Peter Hedström and Richard Swedberg, 172–203. Cambridge: Cambridge University Press.

Burt, Ronald. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.

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

Craver, Carl. 2007. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Clarendon Press.

Cummins, Robert. 1983. *The Nature of Psychological Explanation*. Cambridge, MA: Bradford/The MIT Press.

Darden, Lindley. 2006. *Reasoning in Biological Discoveries. Essays on Mechanisms, Interfifield Relations, and Anomaly Resolution*. Cambridge: Cambridge University Press.

Elster, Jon. 1989. *Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.

Elster, Jon. 2007. *Explaining Social Behavior. More Nuts and Bolts for the Social Science*. Cambridge: Cambridge University Press.

Gambetta, Diego. 1998. "Concatenations of Mechanisms." In *Social Mechanisms: An Analytical Approach to Social Theory*, eds. Peter Hedström and Richard Swedberg, 102–24. Cambridge: Cambridge University Press.

Harré, Rom. 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 Peter and Peter Bearman, eds. 2009. *The Oxford Handbook of Analytical Sociology*. Oxford: Oxford University Press.

Hedström Peter and Richard Swedberg. 1998. "Social Mechanisms: An Introductory Essay." In *Social Mechanisms: An Analytical Approach to Social Theory*, eds. Peter Hedström and Richard Swedberg, 1–31. Cambridge: Cambridge University Press.

Hedström, Peter and Petri Ylikoski. 2010. "Causal Mechanisms in the Social Sciences." *Annual Review of Sociology* 36: 49–67.

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

Horgan, Terence. 1993. "From Supervenience to Superdupervenience: Meeting the Demands of a Material World." *Mind* 102: 555–86.

Kim, Jaegwon. 1993. *Supervenience and Mind*. Cambridge: Cambridge University Press.

Kim, Jaegwon. 1998. *Mind in a Physical World*. Cambridge, MA: The MIT Press.

Kincaid, Harold. 1996. *Philosophical Foundations of the Social Sciences. Analyzing Controversies in Social Research*. Cambridge: Cambridge University Press.

Kuorikoski, Jaakko. 2009. "Two Concepts of Mechanism: Componential Causal System and Abstract Form of Interaction." *International Studies in the Philosophy of Science* 23 (2): 143–60.

Kuorikoski, Jaakko and Petri Ylikoski. 2010. "Explanatory Relevance Across Disciplinary Boundaries—The Case of Neuroeconomics." *Journal of Economic Methodology* 17 (2): 219–28.

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

Mayntz, Renate. 2004. "Mechanisms in the Analysis of Social Macro-Phenomena." *Philosophy of the Social Sciences* 34 (2): 237–59.

McCauley, Robert. 2007. "Reduction: Models of Cross-Scientific Relations and Their Implications for the Psychology-Neuroscience Interface." In *The Handbook of Philosophy of Science. Philosophy of Psychology and Cognitive Science*, ed. Paul Thagard, 105–58. Amsterdam: North Holland.

Oppenheim, Paul and Hilary Putnam. 1958. "Unity of Science as a Working Hypothesis." In *Concepts, Theories, and the Mind—Body Problem, Minnesota Studies in the Philosophy of Science II*, eds. Herbert Feigl, Michael Scriven, and Grover Maxwell, 3–36. Minneapolis: University of Minnesota Press.

Richardson, Robert. 2007. "Reduction without the Structures." In *The Matter of the Mind. Philosophical Essays on Psychology, Neuroscience, and Reduction*, eds. Maurice Schouten and Huib Looren de Jong, 123–45. Oxford: Blackwell.

Salmon, Wesley. 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.

Sawyer, R. Keith. 2005. *Social Emergence. Societies as Complex Systems*. Cambridge: Cambridge University Press.

Scott, John. 2000. *Social Network Analysis. A Handbook*. 2nd ed. London: Sage.

Sperber, Dan. 2006. *Explaining Culture: A Naturalistic Approach*. Oxford: Blackwell.

Thagard Paul. 1999. *How Scientists Explain Disease*. Princeton, NJ: Princeton University Press.

Turner, Stephen. 1994. *The Social Theory of Practices. Tradition, Tacit Knowledge and Presupposition*. Cambridge: Polity Press.

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

Wendt, Alexander. 1998. "On Constitution and Causation in International Relations." *Review of International Studies* 24 (5): 101–17.

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

Woodward, James. 2002. "What Is a Mechanism? A Counterfactual Account." *Philosophy of Science* 69 (3): S366–S377.

Woodward, James. 2003. *Making Things Happen. A Theory of Causal Explanation*. Oxford: Oxford University Press.

Woodward, James. 2008. "Mental Causation and Neural Mechanisms. In Being Reduced." *Essays on Reduction, Explanation, and Causation*, eds. J. Hohwy and J. Kallestrup, 218–62. Oxford: Oxford University Press.

Wright, Cory and William Bechtel. 2007. "Mechanisms and Psychological Explanation." In *The Handbook of Philosophy of Science. Philosophy of Psychology and Cognitive Science*, ed. Paul Thagard, 31–79. Amsterdam: North Holland.

Ylikoski, Petri. 2003. "Explaining Practices." *Protosociology* 18: 316–30.

Ylikoski, Petri. 2007. "The Idea of Contrastive Explanandum." In *Rethinking Explanation*, eds. Johannes Persson and Petri Ylikoski, 27–42. Dordrecht: Springer.

Ylikoski, Petri. 2011. "Social Mechanisms and Explanatory Relevance." In *From Social Mechanisms to Analytical Sociology*, ed. Pierre Demeulenaere, 154–72. Cambridge: Cambridge University Press.

Ylikoski, Petri and Jaakko Kuorikoski. 2010. "Dissecting Explanatory Power." *Philosophical Studies* 148 (2): 201–19.

Zahle, Julie. 2006. "Holism and Supervenience." In *The Handbook of Philosophy of Science. Philosophy of Anthropology and Sociology*, eds. Stephen P. Turner and Mark Risjord, 311–41. Amsterdam: North Holland.