

## Chapter 2

# Causal Mechanisms: Lessons from the Life Sciences

Philip S. Gorski

The appeal to “causal mechanisms” has been a rallying cry for Critical Realism (CR) ever since its (re)emergence in the UK during the 1970s (Bhaskar 1975, 1979; Harré 1970; Harré and Madden 1975).<sup>1</sup> Causal mechanisms have likewise become a conceptual mainstay of the Morphogenetic/Morphostatic (MM) approach since its inception in the late 1970s (Archer 1979, 1985, 1995).

Over the last decade and a half, other voices have joined in as well, and the rallying cry has rapidly built into something of a chorus (Gerring 2008; Hedström and Ylikoski 2010; Mahoney 2001). In American social science, at least, mechanisms have now gone mainstream. Mechanisms talk pervades not only sociology but political science and economics as well.

The attraction of the mechanistic approach is clear enough. There is a widespread recognition that the search for social laws, even probabilistic ones, has proven futile. There is also a general if not universal sentiment that cultural interpretation does not exhaust social science; some form of causal explanation must also be a goal. In the present constellation, then, many social scientists are attracted to mechanistic explanation as a possible *via media* between nomothetic hubris and idiographic humility.

How should Critical Realists and Morphogenetic theorists respond to the sudden popularity of mechanisms talk? With some ambivalence, I will argue. On the one hand, they should welcome it, insofar as the turn towards mechanisms does involve a turn away from logical positivism. On the other hand, they should remain wary, because the mechanistic turn has not been as sharp or as final as it seems; many neo-mechanists are still half-positivist.

The remainder of the essay is in four parts. In Part I, I review the four most influential approaches to causal mechanisms within contemporary American

---

<sup>1</sup>I say “re-emergence”, because the term first appears in English in: Sellars, Roy Wood (1916).

P.S. Gorski (✉)

Department of Sociology, Yale University, New Haven, CT 06520-8265, USA

e-mail: [philip.gorski@yale.edu](mailto:philip.gorski@yale.edu)

sociology: mainstream, analytical, counterfactual and neo-pragmatist. I argue that they are not particularly critical nor even fully realist. In Part II, I subject the mechanisms concept to a critical-historical analysis. I argue that the mechanisms metaphor still carries a good deal of ontological baggage from the mechanistic worldview of the seventeenth century. It must therefore be used with a great deal of caution. In Part III, I survey the recent discussion of causal mechanisms within the philosophy of biology. I highlight a number of commonalities between biological and social mechanisms, but caution against easy analogies. In the conclusion, I argue in favor of a thicker and more pluralistic understanding of causation, not only for the social sciences, but also for the “special sciences” more generally.

## 2.1 Causal Mechanisms in American Sociology: Four Approaches

There is a folk version of mechanisms talk that one encounters quite often in American sociology and political science these days. I will call it the generic approach (GA). In the GA, causal mechanisms are a supplement to the positivist approach. “Cause” and “effect” are still defined in Humean or positivist terms, i.e., as “events” or “variables.” “Causal mechanisms” are then construed as the “causal chain” that connects them. The causal links are likewise conceptualized as “variables” or “events” and frequently characterized as “mediating” or “intervening.” Even a casual inspection of recent articles in the leading journals in the United States will quickly turn up many examples of this “association plus theory” version of causal mechanisms.

The historical origins of the GA are somewhat unclear. Knight, Morgan and Winship trace them back to Paul Lazarsfeld’s notion of “M-accounts” (Knight and Winship 2013; Morgan and Winship 2007). Such accounts introduce a third variable that might cast further light on the causal connections that underlie a statistical association and help disambiguate the direction of causal influence (Kendall and Lazarsfeld 1950; Lazarsfeld 1955). For example, if one discovered a correlation between being male and having car accidents, one might introduce an additional variable such as “miles driven” and see what effect this had on the association (Hagenaar 2004).

A decade later, Hubert Blalock (1967), Otis Dudley Duncan (1966) and others drew on the pioneering work of biologist Sewall Wright (1921) in order to incorporate intervening variables into multiple regression analysis via path analytic techniques (Sewell et al. 1969, 1970). In this way, it was argued, one could indeed get from correlation to causation (Blalock 1968; Land 1969).

While another famous member of the Columbia Sociology Department did explicitly invoke the mechanisms idea (Merton 1949), Lazarsfeld himself did not, nor did later advocates of path analysis, with one important exception: Raymond Boudon. Harkening back to Merton, Boudon would argue that the only way to get

from statistical association to causal inference was via “generative mechanisms” derived from social theory (Boudon 1974, 1976, 1991). By “theory”, of course, Boudon meant rational choice theory.

But Boudon’s proto-GA approach was vehemently rejected by leading members of the Wisconsin School such as Robert Hauser, who insisted that a causal explanation just was a statistical model, nothing more, nothing less (Hauser 1976). Hauser’s view prevailed, at least within mainstream sociology, and in the decades that followed, most sociologists stopped talking about causal mechanisms. Why? One likely reason is that rapid advances in computing power and statistical software made it so easy to do “kitchen sink” regression analyses (i.e., to throw in every variable “including the kitchen sink”). Be that as it may, this much is certain: the GA approach to causal mechanisms long antedates the current-day revival of mechanistic thinking – indeed, it entirely antedates widespread use of the mechanisms concept.

Apart from a few neo-Marxists familiar with CR (Brooks 1989; Isaac 1987a, b; Wright 1987, 1997) there was relatively little explicit talk about causal mechanisms within American sociology between the mid-1970s and the mid-1990s.<sup>2</sup> This changed abruptly around the turn of the millennium, due mainly to the efforts of an interdisciplinary group of predominantly Scandinavian social scientists centered around Peter Hedström. Placing themselves in the lineage of Lazarsfeld and Merton, they sought to revive the agenda first set forth by Boudon two decades before, namely, a loose-jointed version of rational choice theory in which individual actors were the basic building blocks of all causal mechanisms, the “cogs and wheels” inside of the “black box” connecting causal variables (Elster 1983, 1989, 1999; Hedström and Swedberg 1996, 1998; Sørensen 1998). Perhaps because of the deep resistance to rational choice within American sociology, the Hedström group later restyled their approach as “Analytical Sociology” (Demeulenaere 2011; Hedström 2005, 2008; Hedström and Bearman 2009a; Hedström and Ylikoski 2010).

Unlike Boudon, Hedström and many of his followers do not envision causal mechanisms as a mere add-on for statistical analysis (i.e., as a way of giving greater “depth” to regression modeling). On the contrary, they have become increasingly opposed to variables-oriented sociology as such. Instead, following James Coleman (1994), they see it as a way of putting sociological analysis on firmer ontological foundations – namely, methodologically individualistic ones. All social phenomena, they insist, can ultimately be reduced to “micro-level” interactions between individuals, with their “desires, beliefs, and opportunities.” They do include “macro-level phenomena” in their basic model, but only as perceived “constraints” on individual action. They explicitly reject strong, ontological versions of social-structural emergence in favor of weak, epistemic understandings of property emergence. In other words, they regard social structures as real only if and insofar as social actors behave as if they are real. Accordingly, they distinguish between three basic categories of social mechanisms: (1) “macro-micro” or “situational”; (2) “micro-micro” or “individual action” and (3) “micro-macro” or “transformational” (Hedström and Swedberg 1996: 297). While the AS approach appears similar to the

<sup>2</sup>For one important exception, see: Stinchcombe, Arthur L. (1991).

MM approach at a schematic level, the resemblance is only superficial, not only because social structures are treated as weakly emergent but also because human persons are treated as “rational actors.”

The third approach to causal mechanisms currently on offer within American sociology is “counterfactual dependency” (CFDp). Its chief advocates have been Christopher Winship and his students (Elwert and Winship 2010; Elwert 2013; Morgan 2001, 2013; Morgan and Winship 2007). The principal architects of CFDp have been James Woodward (2002, 2003, 2004, 2011), on the philosophical side, and Judea Pearl, on the statistical side (Pearl 2009, 2010), though other philosophers and statisticians have lately begun contributing to this literature as well (Bollen and Pearl 2013; Hitchcock 2001; Hoerl et al. 2011; Spirtes et al. 1993).

David Hume is sometimes presented as the founder of this approach (Menzies 2009). In a key passage of the *Essay Concerning Human Understanding*, Hume offered the following definition of causation: “We may define a cause to be an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second. Or, in other words, *where, if the first object had not been, the second never had existed.*” Hume’s second locution may be read as implying a relation of counterfactual dependency. Other scholars (Hausman 1981; Sekhon 2004), including Pearl (2009) himself, trace the origins of CFDp to John Stuart Mill’s “method of difference” (Mill 1986). That method, to recall, involves comparing two similar cases that yield different outcomes to find the one major difference that distinguishes them; this will be the key cause.

But the revival of counterfactual reasoning within contemporary philosophy is mainly due to the influence of David Lewis. In his early writings on counterfactuals, Lewis defined causation as follows: “Where *c* and *e* are two distinct possible events, *e* causally depends on *c* if and only if, if *c* were to occur *e* would occur; and if *c* were not to occur *e* would not occur” (Lewis 1973a, b). Following Hume, then, Lewis understands cause and effect as “events.” He further stipulates that these events must be “independent” of one another. Critics immediately discovered a number of problems with Lewis’ approach. What if *c* and *e* are both caused by *b* (spuriousness)? What if *c* is caused by *b* (transitivity)? What if *e* can be caused by *b* and/or *c* (overdetermination)? What if *c* is prevented by *b* (preemption)? How can one be certain that *c* caused *e*, rather than the other way around (temporal asymmetry)? How can one be certain that *c* could not have occurred (possible worlds)?

In his later writings, Lewis attempted to deal with these problems by redefining counterfactual dependency in terms of continuous variation rather than discrete events, such that if, how, when *c* occurs will affect if, how and when *e* occurs (Lewis 2000). Conceived in this way, counterfactual semantics were easily combined with statistical analysis of “potential outcomes” (the so-called Neyman-Holland-Rubin model). CFDp was born.

The potential outcomes approach was originally developed by horticulturalists and epidemiologists, who were interested in the average effects of specific interventions on a particular population (e.g., the use of a new fertilizer or drug). In other words, the goals were practical rather than scientific. But what if the analyst was interested in typical causes rather than average effects? Judea Pearl has argued that one can get from observed effects to underlying causes by combining counterfactual

reasoning with three further elements: causal models, causal graphs and structural equations. By “causal models”, he understands pictorial representations of a causal system. His stock examples are human-made physical set-ups, such as electrical circuits. By causal graphs, he understands “directed acyclic graphs” (DAGs) which represent causal processes via arrow diagrams, with each arrow standing in for a causal variable. Structural equation models are then used to test for the presence, strength and direction of the causal effects. It is very important to note that Pearl’s approach requires a number of highly restrictive assumptions that are not often realized in the social world. Otherwise, the statistical tests will not be sufficient to establish the direction and magnitude of the effects, nor can the DAG be assumed to be an accurate model of the actual causal process.

The fourth and final approach to causal mechanisms that I wish to touch on is the neo-pragmatist one (NP) recently sketched out in a well-received article by Neil Gross (2009). There, Gross proposes that we define “social mechanisms” as “composed of chains or aggregations of actors confronting problem situations and mobilizing more or less habitual responses” (2009: 368). This leads to a research agenda which “entails breaking complex social phenomena into their component parts to see how aggregations or chains of actors employing habits to resolve problem situations to bring about systematic effects” (2009: 375). As Gross himself notes, the NP approach is fairly similar to AS, insofar as it tries to explain higher-order social properties in terms of lower-order individual-level processes. It also resembles AS in another respect, which Gross does not highlight: namely, in its rejection of strong social emergence. There are at least two important differences between NP and AS. One is that the NP rejects the utilitarian rational-actor model in favor of a practical, habitual-actor model. Another is that it invokes micro-macro explanations to explain stability (morphostasis), rather than transformation (morphogenesis), as in AS. The implication, never made explicit by Gross, is that system-level change – morphogenesis – is the result of “creative action” (Joas 1996) that responds to “problem situations” that challenge habitual routines. While the NP approach is much less developed at present than either AS or CFDp, both philosophically and methodologically, neo-pragmatism qua social theory is certainly very much *en vogue* amongst younger, theoretically minded American sociologists today.

For the Critical Realist, however, none of these approaches can be considered fully realist, if by “fully realist” we mean epistemologically, ontologically and ethically realist. For example, the AS approach is epistemologically realist but not ontologically realist. It squarely rejects empiricist and positivist understandings of causation as a constant conjunction or probabilistic association between events or variables. And it firmly embraces a realist view of causation in terms of mechanisms. However, it is ontologically ir-realist to the extent that it allows only an epistemic form of social emergence understood as higher-order properties that are perceived by social actors. As I have shown elsewhere (Gorski 2009), this renders AS – and all such efforts to combine social realism with methodological individualism – ontologically incoherent. How so? On the one hand, AS admits the existence of non-observable sub-individual level entities and process (e.g., conflicting desires and rational choices) while denying the existence of supra-individual entities and processes on the grounds that they are not observable. In this regard, AS is still empiricist and not fully realist.

With CFDp, we encounter the reverse situation: it is ontologically realist but epistemologically antirealist. It is ontologically realist to the degree that it at least tacitly allows for “downward” or “macro-micro” forms of causation. For example, CFDp analyses of social mobility often look at the impact that a “macro” variable such as education or neighborhood has on a “micro” variable such as life chances or average income (Morgan 2001; Sharkey and Elwert 2011). Still CFDp remains epistemologically irrealist insofar as it conceptualizes causation as a probabilistic relationship between variables, rather than as a processual relationship between active entities. And this generates epistemological confusion within CFDp. For instance, empirical analyses within the CFDp framework frequently confuse mechanisms with models. Specifically, they present DAGs as causal mechanisms rather than as statistical models of those mechanisms.

Gross’ NP approach might be described as “Hume lite.” It is doubly irrealist but not strongly so. The epistemological irrealism manifests itself in the unspoken equation of mechanisms with regularities. Consequently, it cannot allow for suppressed, inactive or intermittent mechanisms. In other words, neo-pragmatism is another form of actualism. The ontological irrealism reveals itself in an easygoing form of methodological individualism, which is skeptical about the existence of supra-individual social structures. It leads to two programmatic difficulties. One is that social transformation cannot be explained in terms of causal mechanisms; it can only be accounted for in terms of “creative action” (Joas 1996). Of course, creative action can be a mechanism of social transformation (Sewell 1996). But it is hardly the only one. The other shortcoming is that it must explain morphostasis purely in terms of habitual action, because it lacks, or rather eschews, any notion of social structures that might generate or reproduce habits.

What is more, none of the four approaches is morally realist. They see causal mechanisms as an integral feature of a good explanation; but they do not attribute any critical function to them. Both CR and MM see the proper identification of causal mechanisms as the *sine qua non* of a good explanation. But they also believe that mechanistic analysis can function as a form of social critique, and in at least two different ways. The first is what Bhaskar has called “explanatory critique” (Bhaskar 1986, 2002). This involves the identification of social mechanisms whose operation is systematically misrecognized by, and therefore concealed from, the social actors themselves, where such misrecognition is crucial to the continued operation of the mechanism. This form of critique is hardly specific to CR of course. After all, the paradigmatic example of an explanatory critique is Marx’s analysis of the extraction of surplus value (Marx et al. 1976). Of course, social actors do not always have a “false consciousness” about the social structures they are enmeshed in; sometimes, they understand them quite well and enter into them more or less voluntarily, *faute de mieux*. I therefore propose that we distinguish a second form of mechanistic analysis. Let us call it “eudemonistic critique.” It involves showing that a particular form of life-conduct or social organization limits or prevents the realization of certain human capacities or relational goods – and that it does so unnecessarily.

## 2.2 Causal Mechanisms and the Physicalist Imaginary: A Critical-Historical Analysis

For the Critical Realist, the recent move towards mechanistic analysis in American sociology marks a welcome departure from the sort of positivist empiricism that so dominated the discipline for over half a century. However, as we have just seen, the realist train has not quite made it out of the Humean station; it remains half-stuck in various forms of skeptical irrealism – epistemological, ontological or moral. Why? There are many reasons, of course, including commitments to: certain positivist-inspired methodological techniques; a deeply individualistic ethico-political framework; and a sharp distinction between “facts” and “values.”<sup>3</sup> In this section, I would like to argue that there is also another deeper and less obvious reason: contemporary approaches to social mechanisms are tacitly structured by a *physicalist imaginary* whose roots lie in the “mechanical philosophy” of the seventeenth century. In concluding, I will contend that CR itself has not entirely disentangled itself from this imaginary.

I borrow the notion of an “imaginary” from Charles Taylor. In *A Secular Age*, for example, Taylor defines a “social imaginary” as “the way that we collectively imagine, even pre-theoretically, our social life” (Taylor 2007: 146). Elsewhere, he defines it more colloquially as: “the ways people imagine their social existence, how they fit together with others, how things go on between them and their fellows, the expectations that are normally met, and the deeper normative notions and images that underlie these expectations” (Taylor 2002: 23).

Following Taylor we might also speak of a “natural imaginary.” By “natural imaginary”, I mean something like “the way that we collectively imagine, even pre-theoretically, the natural world.” Or, more expansively, “how we collectively and often pretheoretically envision the ontological furniture of the natural world, how it is ordered, and where human beings fit into these arrangements.”

Taylor argues that social imaginaries are historically and culturally variable. I would argue that this applies to natural imaginaries as well. The physicalist imaginary, for example, is an early modern revival of the ancient atomism of Democritus and Lucretius, supplemented by the geometrical formalism of Pythagoras and Plato (Funkenstein 1986; Gaukroger 2006; Shapin et al. 1985; Wilson 2008). Put simply, it presumes that the world is “really” composed of elementary particles that interact in a deterministic manner that can be captured mathematically, and that everything else in the world is ultimately epiphenomenal. I refer to this as the “physicalist imaginary” because the resurgence of atomistic metaphysics in the seventeenth Century was largely a philosophical response to the triumph of celestial mechanics, which combined atomism and Pythagoreanism. I say “philosophical”, because its chief proponents were men like Descartes and Hobbes, who may have fancied themselves physicists, but who are now known to us mainly as philosophers, not least because their physical theories were catastrophic failures, whereas those whom

---

<sup>3</sup>For a critique of the latter, see: Gorski, Philip S. (2013). and the literature cited therein.

we nowadays remember as physicists did not firmly embrace this atomistic metaphysics, either because they thought scientific knowledge was founded on experimentation (e.g., Galileo and Boyle) and/or because they were ultimately committed to a non-mechanistic metaphysics of some sort (e.g., Newton, who believed that divine intervention was necessary to maintain cosmic order).

Though they are mostly implicit in modern-day social theory, the basic metaphysical assumptions of the physicalist imaginary are made very explicit in the atomistic physics of early modern mechanism. Four of the key assumptions are as follows: (1) all things really consist of atoms; (2) all change is just the motion and collision of atoms; (3) all such motions and collisions obey the laws of geometry; (4) therefore, all events are fully determined in advance.

While many early modern scientists came to believe that this was an accurate description of the physical world, Hobbes and other neo-Epicurean philosophers argued that the social world could also be understood in exactly the same way (Martinich 2005). It, too, was comprised of “atoms” (i.e., individuals). Individual behavior was driven by internal “motions” (i.e., desires for objects). Human interaction was subject to “natural law” (in the Grotian sense of “self-preservation”). Therefore, social life was also fully deterministic.

The central ambition of the neo-Epicureans was to do to the Aristotelian view of human society what mechanistic physics had done to the Aristotelian cosmos, namely to supplant it. In this way, they hoped to unify the sciences by placing them on the same metaphysical foundation: atomism. In so doing, they transformed a natural imaginary into a social one, giving birth to the physicalist worldview that still underpins much work in sociology and in the social sciences more broadly.

The physicalist imaginary is deeply embedded in the modern social imaginary, so deeply in fact that it is worth recalling the preceding natural imaginary it replaced, namely, the Aristotelian world picture that underwrote medieval natural philosophy (Feser 2004; Sachs 2004). Let me quickly draw out four important points of contrast. (1) Hylemorphic ontology: in Aristotelian metaphysics, the world was comprised, not of atoms, but of “substances”, complex combinations of matter and form, which were hierarchically ordered. The physicalist imaginary was derived from a flat and monistic ontology in which there was only one substance. (2) Causal Pluralism. In Aristotelian natural philosophy, an adequate explanation invoked four different types of causation: material, efficient, formal and final. Early modern mechanism reduced the four types of causation to one: efficient. (3) Powerful Particulars. Different substances behave in characteristic ways. There are no “laws of motion” that apply equally to all realms of being (physical, biological, social and so on). In the physicalist imaginary by contrast particular powers are lumped together into the unifying category of “cause.” (4) Human Freedom. One of the characteristics powers of human persons is to act according to reason; another is to live in society. In the physicalist imaginary, by contrast, human beings are just so many billiard balls, jostling into one another.

Whereas the early modern mechanists projected a physicalist ontology onto the human world, Aristotelianism did more or less the reverse: it understood the cosmos as a living entity, suffused with agency and purpose (Lear 1988). In sum, the shift



from the Aristotelian cosmos to the mechanical world-picture involved stripping away: (1) the ontological category of form and therefore also of substance; (2) material, formal and final aspects of causation; all causation was reconceived as efficient causation; (3) the shift from a purposive to a deterministic view of the natural order. (4) the shift from a biocentric to a physiocentric cosmos.

The continuing influence of the physicalist imaginary on the social sciences can be seen in various ways. One is the enduring power of certain metaphysical prejudices. Two are particularly consequential: ontological smallism and causal deflation. By “ontological smallism” (Wilson 2004: 22–24), I mean the pervasive tendency within scientific discourse to privilege the small over the large in all realms of study. The unspoken assumption is that things at larger scales can only be explained in terms of things at a smaller scale and never the reverse. The classic expressions of smallism in the social sciences are some form of “methodological individualism” and its Siamese twin, methodological reductionism.<sup>4</sup> By “causal deflation”, I mean the tendency to squeeze all forms of causation into the model of efficient causation and, even more, to (re)conceive of efficient causation in a purely mechanistic fashion (i.e., as a direct transfer of energy from one entity to another via physical contact). One common manifestation of deflation is the widespread practice of representing all causation in terms of nodes and arrows. Social science smallism creates an epistemological privilege for reductionism in all disciplines that in turn justifies a disciplinary hierarchy in which intellectual status is directly correlated with smallist commitments. The smaller the primary objects of study in your discipline or sub-field, the more scientific your research is, and the higher your status. Causal deflation, meanwhile, compels social scientists to translate all manner of causal relations into an efficientist language, renders any form of causation not involving physical contact (e.g., collective memory) inherently “spooky” and “unscientific”, and blinds researchers to the diversity and specificity of causal relations in the social world.

CR cannot be charged with smallism. It has been committed to strong emergence and ontological stratification since its inception. However, it has not entirely freed itself of the smallest prejudice. For example, the recurring trope of “underlying mechanisms” carries the unfortunate connotation that mechanisms operate at the micro-scale. There is, as well, a small remnant of causal deflation. The MM approach does draw a clear distinction between “macro” and “micro” causation, to be sure. But macro-to-micro causation is often represented in terms of efficient causation: structure at  $T_1$  impacts agency at  $T_2$ . No doubt! But not only. Structure also influences agency *synchronically* by constraining and enabling certain agentic powers. What is needed, then, is: (1) an understanding of social mechanisms that is fully shorn of the mechanistic metaphysics of the physicalist imaginary; and (2) an understanding of social causation that is more attentive to: (a) different *forms* of social causation; and (b) specific types of causal *powers* in the social world. Recent work in the philosophy of biology can help move CR and MM towards all of these goals.

---

<sup>4</sup>Though in recent years, the relationship has become uncomfortable, as individualism has been challenged by even more radical forms of reductionism (e.g., neuronal or genetic).

### 2.3 Causal Mechanisms in the Life Sciences: The Chicago School Approach

While most Critical Realists will happily agree that the physicalist imaginary is ontologically inadequate, few analytic metaphysicians would join them. Amongst philosophers, particularly philosophers of mind, ontological smallism and causal deflation are still very much the order of the day. In the philosophy of science, however, and especially in the philosophy of biology, other views have been gaining ground, heterodox views that are more consonant with CR and MM. Since much of this work has been done by William Wimsatt and his students at the University of Chicago, I will refer to this approach as “the Chicago School” – not to be confused with the Chicago School in Economics, of course, which is mechanism par excellence! The Chicago School provides very powerful arguments against the physicalist imaginary and a useful starting point for reconstructing our social imaginary along Critical Realist lines.

The area of philosophy where the validity of smallism and deflation have been most heavily debated is probably the philosophy of mind. The central question in this literature concerns the relationship between mind and brain. And the most popular answer is probably Jaegwon Kim’s notion of “supervenience” (Kim 1979, 1987, 1993, 2002). While some of the early twentieth century emergentists used “supervenience” as a synonym for emergence (Broad 1929), present-day philosophers of mind typically present it as an alternative to emergence. Let us assume, as CR does, a “stratified” or “layer-cake” ontology. For simplicity’s sake, let us further assume two strata or layers, “A” and “B”, where “A” is higher and “B” is lower. “The core idea of supervenience is captured by the slogan that there cannot be an *A*-difference without a *B*-difference” (McLaughlin 1995, 2005, 2006). For example, let us imagine that mental states (level A) “supervene” on brain states (level B). This means that any change in mental state (feelings of pain or hunger, thoughts of exercise or dinner, and so on) will correspond to a change in brain state. The attraction of this approach for philosophers of mind is that it saves the qualia – the “secondary qualities” of subjective experience (e.g., sweetness, redness, perhaps even beauty) (Searle 1998, 2004) – but without abandoning physicalism. This is because supervenience allows for a “weak” or “epistemic” form of emergence. It allows for emergent properties (e.g., qualia) that can be exhaustively explained in terms of lower-order physical entities and processes (e.g., the firing of neurons). Some sociologists have also been attracted to supervenience for similar reasons: it allows one to defend “methodological individualism” without denying the existence of macro-social properties (Healy 1998; Hedström and Bearman 2009b; Sawyer 2002, 2005). On this account, there may be “social facts” (e.g., birth rates, crime rates and so on), but they will “supervene” on individual activity. In other words, higher order processes and properties are nothing but aggregations of lower order ones.

The problem with supervenience, as Kim himself has recently concluded, is that it does not in fact provide the sort of stable middle ground between Cartesian dualism and reductive physicalism that it promises (Kim 1999, 2005, 2011). To see why,

consider the mind/brain relationship again. If we accept that all mental states supervene on brain states in Kim's sense, then why bother studying mental states at all? Why not just focus on the brain? After all, the supervenience account strongly implies that the "real" causal action will be at the level of the brain, anyway; mental states are ultimately just epiphenomenal. To suggest otherwise, Kim argues, entails the possibility that mental states might have causal powers independent of brain states, opening the door to "downward causation" (Andersen 2000; Campbell 1974; Murphy et al. 2009). By "downward causation", Kim understands a form of efficient causation in which A properties cause B properties. For example, one's mental state at  $T_1$  would exert "downward causation" on one's brain state at  $T_2$  (Kim 2000). And this, says Kim with rather considerable alarm, would seem to threaten the "physical closure" of the world, because it implies that mental processes might sometimes overrule or even violate physical laws. With the dissolution of supervenience Kim concludes, there are really only two stable positions left in the philosophy of mind: reductive physicalism and metaphysical dualism. And only one of these is scientifically legitimate, namely, physicalism. Is he right?

The recent work of the Chicago School suggests not. In a series of articles, William Wimsatt paves the way by turning the tables on reductive physicalists. He asks: What would it really mean for a higher order system property to really just be nothing but an aggregation of lower order processes? (Wimsatt 1985, 1997). Wimsatt enumerates four conditions which must all be fulfilled: (1) Inter-Substitution: internal rearrangements or external substitutions of system parts will not affect system properties; (2) Qualitative Similarity: Increases in the size or scale of the system have no influence on its system properties; (3) De/Recomposition. The system can be disassembled and reassembled without any loss of system properties. (4) Linearity: "There are no cooperative or inhibitory interactions amongst the parts of the system for this property" (1997: 386). As Wimsatt rightly notes, there are precious few systems that actually fulfill all four of these criteria. The proverbial heap of sand might come close. But even a pile of stones might not, since the exact shape and arrangement of the stones and the size of the pile might affect their stability (Mumford 2012). In Wimsatt's words: "Very few system properties are aggregative, suggesting that emergence, defined as failure of aggregativity, is extremely common – the rule, rather than the exception" (1997: 382).

While Wimsatt has argued strongly against reductive physicalism, other members of his Chicago School have strongly criticized nomothetic understandings of scientific knowledge. At least in the biological sciences, they contend, explanations usually appeal to mechanisms rather than laws. But just what is a biological mechanism? In a much cited paper, Peter Machamer, Lindley Darden and Carl Craver (hereafter: MDC) offer the following definition: "Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions" (Machamer et al. 2000: 3). Let us examine their definition a little more closely. The first thing to note is that it includes both "entities" and "activities." By means of this "dual ontology", MDC seek to incorporate the insights of both "substantialist" definitions of causal mechanisms that focus on the dispositional properties of natural kinds (Cartwright 1989; Ellis and Lierse 1994; Ellis 2001;

Mumford 1998) and those of “process ontologies” that give relations pride of place (Latour 2013; Rescher 1996, 2000; Stengers 2011; Whitehead 1978). It is also worth noting that MDC themselves give priority to activities, and for much the same reasons as MM gives priority to practice in its conception of persons, namely: First, they argue that we learn about the causal nexus of the world through our own activity in the world, regardless of whether “we” means scientific researchers or young children. Second, they argue that entities exert their powers only via activities (Machamer 2004). This leads to a third point: causation is first and foremost about “production”, not conjunction, correlation, or relevance, as the Humeans and neo-Humeans have variously asserted (Glennan 1994). Now, as some critics pointed out (Bogen 2005), the reference to “regularity” might seem to put MDC back in the Humean camp, with its actualist prejudices. However, members of the Chicago School quickly clarified that mechanisms are always regular in their *activities* but not necessarily in their *occurrences* (Craver 2006; Darden 2006; Glennan 2010). Consider the “fight or flight response.” Its operation may be regular, but its initiation is irregular. Another attractive feature of MDC’s definition that bears emphasis is their notion of start-up and finishing conditions. The advantage of this formulation is that it captures the temporal dimension of causal mechanisms without recourse to an events-ontology.

If one commonality between the Chicago School and CR/MM approaches is a commitment to mechanistic explanation and a rejection of nomothetic ones, another is a strong embrace of a layered ontology and a concomitant suspicion of ontological smallism. Biological mechanisms can rarely be fully described within a flat ontology. This is because the entities that comprise them often vary significantly in size or scale. What is more, the activities of some of these entities may depend upon those of various sub-mechanisms as well. Thus, descriptions of biological mechanisms routinely distinguish between various “levels” and frequently specify “inter-level” processes. While decomposition is often a helpful strategy for investigating mechanisms (Bechtel and Richardson 1993), so is re-composition: discovering what role a particular entity plays in a larger system can illuminate why it has the particular powers or structure that it does. Consequently, the direction of investigation in the life sciences is sometimes from large to small, rather than the other ways around. Within most areas of the life sciences, however, scientific investigation operates within a certain scalar range. MDC refer to this operative range as “topping off” and “bottoming out.” The top and bottom levels in a given field are defined through the interplay of disciplinary convention and explanatory relevance. That is to say, that researchers typically have a tacit feel for the scalar range within which they typically search for causal mechanisms, a “personal knowledge” based on their scientific training and theoretical tools. However, they will sometimes breach or move these ontological boundaries in search of a fuller description of the mechanisms they are investigating.

While the Chicago School approach provides a powerful critique of reductive physicalism and ontological smallism, premised on a mechanistic epistemology and a layered ontology, it has thus far been less successful in effecting a reflation of causality. To be sure, MDC’s distinction between entities and activities does point towards the Aristotelian distinction between material and efficient causation. What

is more, MDC's frequent references to the "organization" and "function" of mechanisms and systems gestures towards the categories of "formal" and "final" causation. But in the end, MDC understand causation exclusively in terms of activity, which is to say, in terms of effective causation. Consider the following passage:

In our view, the phrase 'top down causation' is often used to describe a perfectly coherent and familiar relationship between the activities of wholes and the behaviors of their components, but the relationship is not a causal relationship. Likewise, the phrase 'bottom-up causation' does not, properly speaking, pick out a causal relationship. Rather, in unobjectionable cases both phrases describe mechanistically mediated effects. Mechanistically mediated effects are hybrids of constitutive and causal relations in a mechanism... (Craver and Bechtel 2007: 547)

Elsewhere, however, they note that (1) the operation of a mechanism typically depends upon the causal powers of its constituent parts; (2) the organizational form of a causal mechanism both constrains and enables the causal powers of its constituent parts; and (3) the analysis of a mechanism generally requires knowledge of its end state or function (Craver 2001). Why we should not see these relations as causal ones – specifically, material, formal and final – is not at all clear, at least not to me.

## 2.4 Ontological Dis/Analogies Between Biological and Social Mechanisms

The Chicago School approach provides some useful arguments against reductive physicalism. Specifically, it delivers an open challenge to the natural imaginary bequeathed to us by the mechanistic thinkers of the seventeenth century. For them, the work of science was like watching a game of billiards. All the action takes place in a two-dimensional closed system and consists of energy transfers between point particles resulting in motion that obeys the basic rules of Euclidean geometry. Or so the observer may infer after watching repeated rounds of the game. That such interactions presume not only a closed system but human intervention – that the interactions themselves are, in this sense, humanly created – is quickly forgotten. Let us call this the billiard-ball ontology.

Of course, it is no longer clear how far the billiard-ball ontology actually obtains for the atomic world, not to speak of the quantum world. Be that as it may, it is quite clear that the billiard-ball ontology generates a highly inadequate understanding of the biological realm. Let us simply note some gross contrasts to establish this point. To begin, no biological system is perfectly closed. Indeed, one definition of a living organism is that it absorbs external energy in order to sustain internal order. Further, interactions between biological systems typically involve much more than energy transfers. A cell can become infected by a malicious virus, for instance, and an ecosystem can be invaded by a new species.

A second major point of difference is that biological processes occur in four dimensions, rather than two. The third dimension is the spatial dimension of physical scale. Biological entities vary enormously in size from small proteins through

mid-sized organisms to vast ecosystems, and many important causal mechanisms are cross-scalar. The fourth dimension is the temporal one of historical time. Of course, time also matters in the mechanistic world of the billiard-ball ontology but in a purely physical rather than genuinely historical sense. Collisions between billiard balls occur in time. And they lead to new configurations. But the basic parameters of the system and the laws of interaction governing its components do not change. Not so in the biological realm. There, new entities may emerge over time (e.g., molecules, mutations, species, behaviors, niches and so on), creating the possibilities of fundamentally new types of powers and interactions: organisms that can walk or fly, populations that can split or migrate and so on. Meanwhile, the sorts of change that are likely to occur and endure are constrained by changes that have taken place in the past. For example, evolutionary adaptations are constrained by the body plans of organisms (Stadler et al. 2001). Finally, at least some biological entities engage in purposive activity, oriented, at minimum, to physical survival and biological reproduction. In short, material, formal and final causality play a much larger role in the biological world than in the physical world at least as that world is conceived in the mechanistic imaginary.<sup>5</sup>

Let us now turn from the dis-analogies between the physical and biological realms to the analogies between the biological and social realms. There are many. They, too, can be conceived in terms of Aristotle's four types of causation. Let us begin with the material. Human inventiveness continues to bring new entities into the world thereby creating the possibility of new structures and mechanisms. The transportation and communications revolutions of the modern era provide many illustrations (steamships, automobiles, telephones, the internet and so on). Of course, one can also think of artifacts in instrumental terms, as technical means to human ends and, in this way, fold them back into the category of efficient causation. And indeed, that is how most social scientists do tend to think about artifacts, perhaps for that very reason. However, one can – and should – also think of them as material causes of new forms of social organization. Didn't the invention of the railroad contribute to the development of national consciousness? Wasn't the mass produced automobile a material cause of American suburbanization? Wasn't the creation of the internet a material cause of new forms of social networks?

The second analogy, already touched on, concerns formal causation. One of the most common and consequential types of formal causation in the social world is "path dependency" in which established forms of social organization place powerful downstream constraints on subsequent developmental trajectories (Mahoney 2000; Pierson 2004). The paradigmatic example is the QWERTY keyboard. But there are other types as well. Sociologists of organization have long noted the strong

---

<sup>5</sup>The foregoing examples also help us to see where the implications of the Chicago School go somewhat beyond the conclusions drawn by MDC. As I noted at the conclusion of the previous section, MDC acknowledge the significance of what were traditionally referred to as material and formal causes, but decline to refer to them as causes, presumably because they operate synchronically, whereas causation (since Hume) is presumed to be diachronic. Note, however, that the examples just given suggest that material and formal causation may also operate diachronically.

tendencies towards structural “isomorphism” within any given social “field” (higher education, automobile companies, etc.) (Powell and DiMaggio 1991). One reason may be that there are certain well-established ways of doing things in certain domains of social life and new entrants into the domain tend to imitate them to one degree or another. But again, social scientists often tend to conceive of social forms in instrumental or strategic terms so as to subsume them into models of efficient causation. But isn’t this too simple? Don’t social forms also constrain actors’ strategies? Indeed, don’t the dominant forms even “choose” or at least advantage some actors over others? If so, then perhaps it is best to speak of formal causation.

That we should find the Aristotelian schema helpful for thinking about causation in the biological world is hardly surprising. After all, it developed out of Aristotle’s zoological researches. More interesting, perhaps is that we should find it generative in the social domain as well. Of course, biological analogies have a long history in social science. They were frequently deployed by earlier generations, from Spencer and Durkheim through Malinowski and Parsons. But the well-known shortcomings of evolutionary and functionalist approaches should also give us pause and prompt us to reflect on the dis-analogies as well. They, too, are not hard to find.

Let us begin with the material causes of biological and social structures. The building blocks of biological structures are primarily, naturally occurring, material substances, including animal species and populations. By contrast, even moderately complex social structures are minimally composed of: (1) human persons (qua “actors” or “agents”); (2) physical artifacts (e.g., machines, buildings); (3) symbolic systems (e.g., rituals, rules). The contrast should not be overdrawn, of course. Some animals do build things, typically shelters. And some of the higher mammals also appear to be capable of a fairly high degree of intra-species communication. But these latter capacities are far more developed in human animals, opening qualitatively different possibilities. The crucial point is that social structures contain a much higher proportion of artifactual and symbolic elements than one finds even in the most highly developed communities of non-human social animals (e.g., social insects and primates).

Now, let us turn from matter to form. In the biological domain, the form of a structure is often the result of the spatio-temporal organization of naturally occurring matter, such that the microphysical arrangement of the component parts constrains the causal powers of those parts while creating new, higher-order causal powers. In the social domain, by contrast, the form of a structure (also) involves symbolically mediated relations between human persons and artifacts, which coordinate and magnify the causal powers of individual actors. As a result, social structures cannot be properly understood in a purely spatio-temporal manner. We could not understand or even categorize a human institution (e.g., a “bank” or a “college”) simply by observing the placement and movement of persons in a building. To this degree, the old interpretivist critique of “behavioral” social science was spot on. However, interpretivists sometimes imply that social structures are reducible to human interactions, and this is not quite right either. Why? Because of the artifactual element. Buildings, for example, are important to the operations of banks

and schools, because they constrain and enable certain patterns of interaction and cooperation.

What about final causes? Since Darwin, final causation has been declared anathema within the biological sciences. Of course, as critical observers such as Etienne Gilson noted early on, Darwinian theory cannot really do without something very much like final causation (Gilson 1984). For instance, at the level of the organism, it must presume a will to survive and/or reproduce. Meanwhile, at the level of the species, it must presume something like a developmental tendency towards adaptation and/or fitness. Be that as it may, the reflective capacities generated by human language mean that human behavior is not fully intelligible without reference to some sort of finality (Sehon 2005). Why? Because human beings are forever making plans and telling stories (Ricoeur 1990). In making plans, they reflexively seek to pursue their concerns and attempt to relate present actions to future purposes. And in telling stories, they relate past actions to future purposes.

In this section, I have argued that structures and mechanisms in the biological and social realms are not easily handled within the framework of the billiard ball ontology with its deflated view of causation. More positively, I have argued that an Aristotelian approach to ontology and causation provides a far more fruitful starting point, because it restores the scalar and historical dimension to structure and the material, formal and final aspects of causation. Whether it provides an appropriate ending point is beyond the scope of this essay. This much seems certain however: a non-reductive ontology and a pluralist approach to causation would help to resolve some of the persistent aporia that the physicalist imaginary has bequeathed to the modern social sciences.

## 2.5 Conclusion: Mechanisms or Powers?

CR and MM first embraced the mechanisms concept as an alternative to the nomothetic model of scientific explanation championed by logical positivists. Should they continue to do so? I am of two minds about this. On the one hand, there are good intellectual reasons for abandoning the concept. At the same time, there are also good pragmatic and theoretical ones for retaining it.

The fundamental problem with the mechanisms concept is that it primes a whole series of fallacious assumptions about social ontology, specifically: smallism, physicalism, invisibilism and sequentialism. We have already encountered the first two. By smallism, I mean the tendency to privilege smaller units of analysis over larger ones. By physicalism, I mean the tendency to conceptualize interactions in terms of physical contact and energy transfers. However fruitful these heuristics may have proven for the development of mechanistic physics, they have now outlived their usefulness within physics and have proven less useful in the biological domain, where causal mechanisms may include units that differ significantly in scale and even less useful in the social domain where interactions between units are symbolically as well as physically mediated. The mechanisms concept also tacitly implies



that causal processes are invisible to social actors and can only be revealed by social analysts. This is not always the case. Social actors may be quite aware that they are enmeshed in an exploitative relationship. Indeed, they can and do create institutions for the express purpose of dominating others! The final fallacy I would like to touch on is sequentialism. The mechanisms concept implies that a past occurrence can only affect the present by means of a spatio-temporally connected series of physically specifiable interactions. This is manifestly untrue in social life. The most obvious and important counter-example concerns memory. Various forms of social memory – individual and collective, neuronal and historical, traumatic or foundational – may exert an ongoing effect across time (Assmann 2011; Halbwachs 1992; Olick and Robbins 1998). As a result, past events can become a part of a social mechanism, and dead people can be social actors.

The mechanistic approach was supposed to seal the break with logical positivism. In this, it has not succeeded. Why? Because logical positivism is also premised on the physicalist imaginary. This is why anti- and demi-realist versions of causal mechanisms have proliferated in recent years. Crypto-positivists, neo-Humeans, and semi-realists have all embraced the mechanisms concept as a halfway house between nomothetic and fully realist forms of social science.

And yet, it is perhaps for this very reason that proponents of CR and MM should think twice before letting go of the mechanisms concept. It has become an important focus of intellectual debate, with contending schools attempting to impose their preferred definitions on it. There are also important theoretical and political reasons not to let go just yet: the mechanisms concept reminds us that there are fairly regular but non-observable processes in the social world, even today. For example, however much the technological means of modern capitalism have been transformed, the inner logic of capital accumulation is not really as different as some observers suggest. Nor should the rapidity with which capital – and information – now circulate around the globe lead us to imagine that these mechanisms and structures have all dissolved into “contingencies” and “flows” whose only properties are “risk” and “acceleration.” Social life is not that fleeting, even in the morphogenetic society. For all these reasons, social scientists should not give up trying to identify the causal mechanisms that shape our world. To do so would be an abdication of their proper vocation.

## References

- Andersen, P. B. (2000). *Downward causation: Minds, bodies and matter*. Aarhus: Aarhus University Press.
- Archer, M. S. (1979 [2013]). *Social origins of educational systems*. London: Sage.
- Archer, M. S. (1985 [1982]). Structuration vs. morphogenesis. In S. N. Eisenstadt & H. J. Helle (Ed.), *Macro-sociological theory. Perspectives on sociological theory: Vol. 1. Sage studies in international sociology*. New York: Sage.
- Archer, M. S. (1995). *Realist social theory: The morphogenetic approach*. Cambridge/New York: Cambridge University Press.

- Assmann, J. (2011). *Cultural memory and early civilization: Writing, remembrance, and political imagination*. New York: Cambridge University Press.
- Bechtel, W., & Richardson, R. C. (1993). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton: Princeton University Press.
- Bhaskar, R. (1975). *A realist theory of science*. London: Verso.
- Bhaskar, R. (1979). *The possibility of naturalism: A philosophical critique of the contemporary human sciences*. Atlantic Highlands: Humanities Press.
- Bhaskar, R. (1986). *Scientific realism and human emancipation*. London: Verso.
- Bhaskar, R. (2002). *From science to emancipation: Alienation and the actuality of enlightenment*. New Delhi: Sage Publications.
- Blalock, H. M. (1967). Path coefficients versus regression coefficients. *American Journal of Sociology*, 72(6), 675–676.
- Blalock, H. M. (1968). Theory building and causal inferences. In H. M. Blalock & Ann Blalock (Eds.), *Methodology in social research* (pp. 155–198). New York: McGraw-Hill.
- Bogen, J. (2005). Regularities and causality; Generalizations and causal explanations. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 397–420.
- Bollen, K. A. & Pearl, J. (2013). Eight myths about causality and structural equation models. In *Handbook of causal analysis for social research* (pp. 301–328). Dordrecht: Springer.
- Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in Western society*. New York: Wiley.
- Boudon, R. (1976). Comment on Hauser's review of education, opportunity, and social inequality. *American Journal of Sociology*, 81(5), 1175–1187.
- Boudon, R. (1991) What middle-range theories are. *Contemporary Sociology: An International Journal of Reviews*, 20(4, July 1991), 519–522.
- Broad, C. D. (1929). *Mind and its place in nature*. New York: Harcourt, Brace and Company.
- Brooks, E. C. (1989). Realism in social science. *Berkeley Journal of Sociology*, 34(ArticleType: book-review/Issue Title: Symposium on the Foundations of Radical Social Science/Full publication date: 1989/Copyright © 1989 Regents of the University of California):271–280. doi:[10.2307/41035419](https://doi.org/10.2307/41035419).
- Campbell, D. T. (1974). Downward causation. In F. J. Ayala & T. Dobzhansky (Eds.), *Studies in the philosophy of biology* (pp. 179–176). Berkeley/Los Angeles: University of California Press.
- Cartwright, N. (1989). *Nature's capacities and their measurement*. Oxford: Clarendon Press/Oxford University Press.
- Coleman, J. S. (1994). *Foundations of social theory*. Cambridge, MA: Harvard University Press.
- Craver, C. F. (2001). Role functions, mechanisms, and hierarchy. *Philosophy of Science*, 20(4, July 1991), 53–74.
- Craver, C. F. (2006). When mechanistic models explain. *Synthese*, 153(3), 355–376.
- Craver, C. F., & Bechtel, W. (2007). Top-down causation without top-down causes. *Biology and Philosophy*, 22(4), 547–563.
- Darden, L. (2006). *Reasoning in biological discoveries*. New York: Cambridge University Press.
- Demeulenaere, P. (2011). *Analytical sociology and social mechanisms*. Cambridge: Cambridge University Press.
- Duncan, O. D. (1966). Path analysis: Sociological examples. *American Journal of Sociology*, 72(1), 1–16.
- Ellis, B. (2001). *Scientific essentialism*. Cambridge: Cambridge University Press.
- Ellis, B., & Lierse, C. (1994). Dispositional essentialism. *Australasian Journal of Philosophy*, 72(1), 27–45.
- Elster, J. (1983). *Explaining technical change: A case study in the philosophy of science*. Cambridge: Cambridge University Press.
- Elster, J. (1989). *Nuts and bolts for the social sciences*. Cambridge: Cambridge University Press.
- Elster, J. (1999). *Alchemies of the mind: Rationality and the emotions*. Cambridge: Cambridge University Press.

- Elwert, F. (2013). Graphical causal models. In *Handbook of causal analysis for social research* (pp. 245–273). Dordrecht: Springer.
- Elwert, F. & Winship, C. (2010) Effect heterogeneity and bias in main-effects-only regression models. In *Heuristics, probability and causality: A tribute to Judea Pearl* (pp. 327–336). College Publications.
- Feser, E. (2004). *Aquinas*. Oxford: Oneworld.
- Funkenstein, A. (1986). *Theology and the scientific imagination from the middle ages to the seventeenth century*. Princeton: Princeton University Press.
- Gaukroger, S. (2006). *The emergence of a scientific culture science and the shaping of modernity 1210–1685*. Oxford: Oxford University Press.
- Gerring, J. (2008). The mechanistic worldview: Thinking inside the box. *British Journal of Political Science*, 38(01), 161–179.
- Gilson, E. (1984). *From Aristotle to Darwin and back again: A journey in final causality, species and evolution* (J. Lyon, Trans.). Notre Dame: University of Notre Dame Press.
- Glennan, S. (1994). Why there can't be a logic of induction. In *PSA: Proceedings of the Biennial meeting of the philosophy of science association*. Volume 1: Contributed papers (pp. 78–86). University of Chicago Press.
- Glennan, S. (2010). Mechanisms. In H. Beebe, C. Hitchcock, & P. Menzies (Eds.), *The Oxford handbook of causation*. Oxford: Oxford University Press.
- Gorski, P. S. (2009). The ecipes model: A critical realist approach to causal mechanisms. In P. Hedström & B. R. Wittrock (Eds.), *Frontiers of sociology* (pp. 147–194). Leiden: Brill.
- Gorski, P. S. (2013). Beyond the fact/value distinction: Ethical naturalism and the social sciences. *Society*, 50(6), 543–553.
- Gross, N. (2009). A pragmatist theory of social mechanisms. *American Sociological Review*, 74(3), 358–379.
- Hagenaar, J. A. (2004). *Elaboration (Lazarsfeld's method of)*. *The Sage encyclopedia of social science research methods*. Thousand Oaks: Sage Publications.
- Halbwachs, M. (1992). *On collective memory*. Chicago: University of Chicago Press.
- Harré, R. (1970). *The principles of scientific thinking*. Chicago: University of Chicago Press.
- Harré, R., & Madden, E. H. (1975). *Causal powers: A theory of natural necessity*. Ottawa: Rowman and Littlefield.
- Hauser, R. M. (1976). On Boudon's model of social mobility. *American Journal of Sociology*, 81(4), 911–928.
- Hausman, D. M. (1981). John Stuart Mill's philosophy of economics. *Philosophy of Science*, 48(3), 363–85.
- Healy, K. (1998). Conceptualising constraint: Mouzelis, Archer and the concept of social structure. *Sociology*, 32(03), 509–522.
- Hedström, P. (2005). *Dissecting the social: On the principles of analytical sociology*. Cambridge: Cambridge University Press.
- Hedström, P. (2008). Studying mechanisms to strengthen causal inferences in quantitative research. In Janet M. Box-Steffensmeier, Henry E. Brady & David Collier (Eds.), *The Oxford handbook of political methodology* (pp. 319–335). Oxford: Oxford University Press.
- Hedström, P. & Bearman, P. (2009a) *The Oxford handbook of analytical sociology*. Oxford/New York: Oxford University Press
- Hedström, P. & Bearman, P. (2009b). What is analytical sociology all about? An introductory essay. In *The Oxford handbook of analytical sociology* (pp. 3–24).
- Hedström, P., & Swedberg, R. (1996). Social mechanisms. *Acta Sociologica*, 39(3), 281–308.
- Hedström, P., & Swedberg, R. (1998). *Social mechanisms: An analytical approach to social theory*. Cambridge: Cambridge University Press.
- Hedström, P., & Ylikoski, P. (2010). Causal mechanisms in the social sciences. *Annual Review of Sociology*, 36, 49–67.
- Hitchcock, C. (2001). The intransitivity of causation revealed in equations and graphs. *The Journal of Philosophy*, 98(6), 273–299.

- Hoerl, C., McCormack, T., & Beck, S. R. (2011). *Understanding counterfactuals, understanding causation: Issues in philosophy and psychology*. Oxford: Oxford University Press.
- Isaac, J. C. (1987a). Beyond the three faces of power: A realist critique. *Polity*, 20, 4–31.
- Isaac, J. C. (1987b). *Power and Marxist theory: A realist view*. Ithaca: Cornell University Press.
- Joas, H. (1996). *The creativity of action*. Cambridge: Polity Press.
- Kendall, P. L., & Lazarsfeld, P. F. (Eds.). (1950). *Continuities in social research: Studies in the scope and method of "The American Soldier"*. New York: Free Press.
- Kim, J. (1979). Causality, identity, and supervenience in the mind-body problem. *Midwest Studies in Philosophy*, 4(1), 31–49.
- Kim, J. (1987). "Strong" and "Global" supervenience revisited. *Philosophy and Phenomenological Research*, 48(2), 315–326.
- Kim, J. (1993). Can supervenience and 'Non-Strict Laws' save anomalous monism? In his *Mental causation* (pp. 19–26). Cambridge and New York.
- Kim, J. (1999). Making sense of emergence. *Philosophical Studies*, 95(1), 3–36.
- Kim, J. (2000). Making sense of downward causation. In *Downward causation* (pp. 305–321). Denmark: Aarhus University Press.
- Kim, J. (ed) (2002). *Supervenience*. Brookfield, Vermont: Ashgate Publishing.
- Kim, J. (2005). *Physicalism, or something near enough*. Princeton: Princeton University Press.
- Kim, J. (2011). *Philosophy of mind*. Boulder: Westview Press.
- Knight, C. R., & Winship, C. (2013). The causal implications of mechanistic thinking: Identification using directed acyclic graphs. In S. L. Morgan (Ed.), *Handbook of causal analysis for social research* (pp. 275–300). New York: Springer.
- Land, K. C. (1969). Principles of path analysis. *Sociological Methodology*, 1, 3–37.
- Latour, B. (2013). *An inquiry into modes of existence: An anthropology of the moderns* (C. Porter, Trans.). Cambridge: Harvard University Press.
- Lazarsfeld, P. F. (1955). Interpretation of statistical relations as a research operation. In his *The language of social research* (115–125). Glencoe, IL: The Free Press.
- Leahy, J. (1988). *Aristotle: The desire to understand*. Cambridge: Cambridge University Press.
- Lewis, D. (1973a). Causation. *The Journal of Philosophy*, 70(17), 556–567.
- Lewis, D. (1973b). *Counterfactuals*. Cambridge, MA: Harvard University Press.
- Lewis, D. (2000). Causation as influence. *The Journal of Philosophy*, 97(4), 182–197.
- Machamer, P. (2004). Activities and causation: The metaphysics and epistemology of mechanisms. *International Studies in the Philosophy of Science*, 18(1), 27–39.
- Machamer, P., Darden, L., & Craver, F. C. (2000). Thinking about mechanisms. *Philosophy of Science*, 67, 1–25.
- Mahoney, J. (2000). Path dependence in historical sociology. *Theory and Society*, 29(4), 507.
- Mahoney, J. (2001). Beyond correlational analysis: Recent innovations in theory and method. *Sociological Forum*, 16(3), 575.
- Martinich, A. P. (2005). *Hobbes*. New York: Routledge.
- Marx, K., Fernbach, D., Fowkes, D., & Mandel, E. (1976). *Capital: A critique of political economy*. Harmondsworth: Penguin, in association with New Left Review.
- McLaughlin, B. (1995). Varieties of supervenience. In Elias E. Savellos & Ümit D. Yalçın (Eds.), *Supervenience: New essays* (pp. 16–59). Cambridge: Cambridge University Press.
- McLaughlin, B. (2005). Supervenience. In E. L. Zalta (Ed.), *Stanford encyclopedia of philosophy*. <http://plato.stanford.edu/entries/supervenience/>
- McLaughlin, B. (2006). Supervenience. In Lynne Nadel (Ed.), *Encyclopedia of cognitive science*. Hoboken, NJ: Wiley.
- Menzies, P. (2009). Counterfactual theories of causation. In E. L. Zalta (Ed.), *Stanford encyclopedia of philosophy*. <http://plato.stanford.edu/archives/spr2001/entries/causation-counterfactual>. Palo Alto: Stanford University.
- Merton, R. K. (1949). On sociological theories of the middle range. In R. K. Merton (Ed.), *Social theory and social structure* (pp. 39–53). New York: The Free Press.
- Mill, J. S. (1986). *A system of logic, ratiocinative and inductive: Being a connected view of the principles of evidence and the methods of scientific investigation*. Charlottesville: Lincoln-Rembrandt Pub.

- Morgan, S. L. (2001). Counterfactuals, causal effect heterogeneity, and the catholic school effect on learning. *Sociology of Education*, 74, 341–374.
- Morgan, S. L. (Ed.). (2013). *Handbook of causal analysis for social research*. Dordrecht: Springer.
- Morgan, S. L., & Winship, C. (2007). *Counterfactuals and causal inference: Methods and principles for social research*. New York: Cambridge University Press.
- Mumford, S. (1998). *Dispositions*. Oxford: Oxford University Press.
- Mumford, S. (2012). *Metaphysics: A very short introduction*. Oxford: Oxford University Press.
- Murphy, N., Ellis, G. F. R., & O'Connor, T. (Eds.). (2009). *Downward causation and the neurobiology of free will*. Berlin: Springer.
- Olick, J. K., & Robbins, J. (1998). Social memory studies: From “Collective Memory” to the historical sociology of mnemonic practices. *Annual Review of Sociology*, 24(1), 105–140.
- Pearl, J. (2009). *Causality: Models, reasoning, and inference*. Cambridge: Cambridge University Press.
- Pearl, J. (2010). An introduction to causal inference. *The International Journal of Biostatistics*, 6(2), 353–379.
- Pierson, P. (2004). *Politics in time: History, institutions, and social analysis*. Princeton: Princeton University Press.
- Powell, W. W., & DiMaggio, P. (1991). *The new institutionalism in organizational analysis*. Chicago: University of Chicago Press.
- Rescher, N. (1996). *Process metaphysics: An introduction to process philosophy*. Albany: State University of New York Press.
- Rescher, N. (2000). *Process philosophy: A survey of basic issues*. Pittsburgh: University of Pittsburgh Press.
- Ricœur, P. (1990). *Time and narrative*. Chicago: University of Chicago Press.
- Sachs, J. (2004). *Aristotle's physics*. New Brunswick: Rutgers University Press.
- Sawyer, R. K. (2002). Nonreductive individualism part I—Supervenience and wild disjunction. *Philosophy of the Social Sciences*, 32(4), 537–559.
- Sawyer, R. K. (2005). *Social emergence: Societies as complex systems*. New York: Cambridge University Press.
- Searle, J. R. (1998). *Mind, language and society: Philosophy in the real world*. New York: Basic Books.
- Searle, J. R. (2004). *Mind: A brief introduction*. Oxford: Oxford University Press.
- Sehon, S. (2005). *Teleological realism: Mind, agency and explanation*. Cambridge: MIT Press.
- Sekhon, J. S. (2004). Quality meets quantity: Case studies, conditional probability, and counterfactuals. *Perspectives on Politics*, 2, 281–294.
- Sellars, R. W. (1916). *Critical realism: A study of the nature and conditions of knowledge*. New York: Rand McNally & Company.
- Sewell, W. H., Jr. (1996). Historical events as transformations of structures: Inventing revolution at the bastille. *Theory and Society*, 25(6), 841–881.
- Sewell, W. H., Haller, A. O., & Portes, A. (1969). The educational and early occupational attainment process. *American Sociological Review*, 34(1), 82–92.
- Sewell, W. H., Haller, A. O., & Ohlendor, G. W. (1970). The educational and early occupational status attainment process: Replication and revision. *American Sociological Review*, 35(6), 1014–1027.
- Shapin, S., Schaffer, S., & Hobbes, T. (1985). *Leviathan and the air-pump*. Princeton: Princeton University Press.
- Sharkey, P., & Elwert, F. (2011). The legacy of disadvantage: Multigenerational neighborhood effects on cognitive ability. *American Journal of Sociology*, 116(6), 1934.
- Sørensen, A. B. (1998). Theoretical mechanisms and the empirical study of social processes. In P. Hedström, & R. Swedberg (Eds.), *Social mechanisms: An analytical approach to social theory* (pp. 238–266). Cambridge: Cambridge University Press.
- Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causality, prediction, and search*. Cambridge, MA: MIT Press.

- Stadler, B. M., Stadler, P. F., Wagner, G. P., & Fontana, W. (2001). The topology of the possible: Formal spaces underlying patterns of evolutionary change. *Journal of Theoretical Biology*, 213(2), 241–274.
- Stengers, I. (2011). *Thinking with Whitehead: A free and wild creation of concepts* (M. Chase, Trans.). Cambridge: Harvard University Press.
- Stinchcombe, A. L. (1991). The conditions of fruitfulness of theorizing about mechanisms in social science. *Philosophy of the Social Sciences*, 21(3), 367–388.
- Taylor, C. (2002). Modern social imaginaries. *Public Culture*, 14(1), 91–124.
- Taylor, C. (2007). *A secular age*. Cambridge, MA: Belknap Press of Harvard University Press.
- Whitehead, A. N. (1978). *Process and reality: An essay in cosmology*. New York: Free Press.
- Wilson, R. A. (2004). *Boundaries of the mind: The individual in the fragile sciences-cognition*. Cambridge: Cambridge University Press.
- Wilson, C. (2008). *Epicureanism at the origins of modernity*. Oxford: Oxford University Press.
- Wimsatt, W. C. (1985). Forms of aggregativity. In Alan Donagan, Anthony N. Perovich Jr. & Michael V. Wedin (Eds.), *Human nature and natural knowledge* (pp. 259–291). Dordrecht: Springer.
- Wimsatt, W. C. (1997). Aggregativity: Reductive heuristics for finding emergence. *Philosophy of Science*, 64(Supplement), S372–S384.
- Woodward, J. (2002). What is a mechanism? A counterfactual account. *Philosophy of Science*, 69(S3), S366–S377.
- Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.
- Woodward, J. (2004). Counterfactuals and causal explanation. *International Studies in the Philosophy of Science*, 18(1), 41–72.
- Woodward, J. (2011). Mechanisms revisited. *Synthese*, 183(3), 409–427.
- Wright, S. (1921). Systems of mating. I. The biometric relations between parent and offspring. *Genetics*, 6(2), 111–123.
- Wright, E. O. (1987). Reflections on classes. *Berkeley Journal of Sociology*, 32, 19–49.
- Wright, E. O. (1997). Rethinking once again the concept of class structure. In his *Reworking class* (pp. 41–72). Cambridge: Cornell University Press.