Mechanisms and Explanation

On the Origins of the New Mechanistic Philosophy

Bendik Hellem Aaby

Thesis presented for the degree of
MASTER OF PHILOSOPHY

Supervised by
Anders Strand

Department of Philosophy, Classics, History of Art and Ideas
Faculty of Humanities

UNIVERSITY OF OSLO
Spring 2016

Bendik Hellem Aaby

Mechanisms and Explanation

On the Origins of the New Mechanistic Philosophy

Abstract

Mechanisms are how things work. The new mechanistic approach to scientific explanation takes a description of how things work as *the* central explanatory aim. In this thesis I look at the debates around scientific explanation in the latter half of the 20th century. The deductive-nomological model sets the stage. By looking at the problems with this model I locate two normative criteria for an adequate explanation, explanatory relevance and asymmetry. This leads us to the causal-mechanical model, which marks the first step towards new mechanistic explanation. I then look at the logical empiricists' model for inter-level explanation, theory reduction. The reactions to this model from the systems tradition mark the second step towards new mechanistic explanation. I then look at the new mechanistic philosophy, going through the common views on explanation and the internal differences on the nature of mechanisms and mechanistic explanation. I argue that the general trend is move away from the *global* approach the philosophy of science, characterized by logical empiricism, towards a *local* approach that defends the autonomy of the special sciences. This is the historical part of my thesis. Parallel to the historical treatment, I argue for a meta-philosophical claim. A consequence of the local approach to the philosophy of science is that the concepts invoked are often restricted in their (intended) scope or generality. The meta-philosophical claim is that we need to pay explicit attention to such restriction so that we don't argue across purposes. I finish the thesis by suggesting that much of the current disagreement on new mechanistic philosophy, especially explanation, can be traced to the lack of explicit discussion on the restriction of the new mechanistic framework.

Mechanisms and Explanation: On the Origins of Mechanistic Philosophy

https://www.duo.uio.no

Acknowledgements

When I started my thesis in the autumn of 2014 I was a historian turned philosopher. I was uncertain of my interests, but luckily I approached Anders Strand at an early stage. No graduate student could ask for a more helpful supervisor. I was quickly introduced to interesting literature on topics I would have struggled to discover myself. I am endlessly grateful for his guidance, support, patience, and the amount time and attention I have received.

Further I would like to thank Gry Oftedal for providing me with feedback on early drafts on my thesis. I would also like to extend my gratitude to Dr. Veli-Pekka Parkinen for providing me with an extensive amount of relevant literature and valuable discussion in a study group on causation. Hans-Robin Solberg and Torbjørn Amundsen Lien also deserve special mentioned, they have provided valuable insight, each from their own perspective. I would also like to thank Mori Diakite for discussion and collaboration.

There are few whose work go as unacknowledged as parents, I extend my gratitude to my mother and father, as they show enormous support in all my endeavors.

I also want to register my appreciation for receiving a scholarship from the *Science Studies Colloquium Series*.

Lastly, a special thanks is in order to Conrad Bakka for his insightful comments, and recommended revisions. Few are as lucky to have such a good friend in the time-consuming process of finishing a master's thesis.

B.H.A

Oslo,
June 2016

Table of Contents

§1. Introduction

> A typical thesis of positivistic philosophy of science is that all true theories in the special sciences should reduce to physical theories in the long run.
>
> Fodor (1974: 97)

> We need a philosophy of science that can be pursued by *real* people in *real* situations in *real* time with the kinds of tools that we *actually* have—now or in a realistically possible future.
>
> Wimsatt (2007: 5)

§1.1 The two principal claims

This thesis has two principal claims, one historical and one meta-philosophical.

<u>*Historical claim*: </u>The new mechanistic philosophy can be situated as an extension of a pluralistic methodology in the philosophy of science that appeared due to a shift from a global approach, characterized by logical empiricism, to a local approach that treat the special sciences as autonomous.

<u>*Meta-philosophical claim*: </u>Paying attention to the restrictions that usually follow with a pluralistic methodology can help us evaluate and locate the arguments provided by the mechanists, the arguments against the new mechanists, and the disagreements within the mechanist literature.

I argue for the historical claim by tracing developments in the concepts of scientific explanation, reduction, and hierarchical organization. After accounting for the global nature of the logical empiricist approach to the explication these concepts, I argue that alternative local accounts were provided primarily by philosophers of the special sciences, which have a local approach to the philosophy of science. I argue for the meta-philosophical claim by showing that classical counter-examples to global approaches do not work as effectively to undermine local accounts. Further, I argue that differences between the mechanists themselves, especially in terminology, are largely a consequence of the local approach to philosophy of science. Lastly, I show that the shift from a global to a local approach makes it difficult to determine the intended scope of local frameworks, by using mechanistic explanation as an example, in the philosophy of science if one doesn't keep the meta-philosophical claim in mind.

I have chosen to focus on new mechanistic explanation because it follows in the tradition of local approaches to the philosophy of science, but it is unclear whether or not it is moving in the direction of a global approach. I think much of the disagreement in the new mechanistic literature rests on a failure to explicitly state if the generality of a global account is desired or not. I argue that can lead to unwillingness from non-mechanistic philosophers to approach the literature.

A note on the terminology invoked in the claims. By *global approach*, I mean an approach to philosophy of science that views science as, in principle, unified. Unification is achieved when a fundamental science (or theory) can adequately account for the diverse phenomena studied by all the different sciences. This is commonly referred to as theoretical monism. For the logical empiricist, the global methodology and theoretical monism have influenced the analysis of the central concepts in the philosophy of science. I discuss three examples of this: I look in detail at Hempel's deductive-nomological model of explanation (Hempel & Oppenheim 1948, Hempel 1965) in §2, theory reduction (Nagel 1961: ch. 11) in §3.3, and the global hierarchical organization of the constituents of the world which corresponds to the hierarchy of scientific disciplines (Oppenheim & Putnam 1948) in §3.5.

By *local approach*, I mean the approach to the philosophy of science that take the disunity of science as best characterizing the current, and presumably the foreseeable future, state of science. On such an approach, different scientific disciplines are autonomous. The local approach adopts a fully-fledged pluralism, where there is a plurality of theories, methodologies and motivations. It is generally the sciences that recommend what one should be pluralist on. The analysis of the central concepts in the thesis (i.e. explanation, reduction and hierarchical organization), on the local approach, should be understood in the context of the relevant scientific discipline were they are invoked and scientific practice can be a reliable justificatory authority. The local approach to analyzing the concepts mentioned above is generally, as we shall see, promoted by philosophers of the special sciences.

Salmon's (1984) causal-mechanical model of explanation is not easily situated either on the global or local approach. However, his emphasis on the explanatory relevance of causation has influenced the mechanists to a great extent, so even though he does not clearly emphasize a local or global approach, he needs to be included in a historical account of the origins of the mechanistic philosophy. As Bechtel (2006) notes:

> Salmon was one of the first twentieth century philosophers of science to revive the interest in mechanistic explanation. … Salmon's account offers a significant advance over seventeenth- and eighteenth-century mechanical philosophy. (Bechtel 2006: 25)

The advance on mechanistic philosophy was by broadening the category of what counts as mechanism, such that the seventeenth- and eighteenth-century requirement that only the strictly atomistic physical properties, such as shape and motion of particles was not the only properties that could figure in accounts of mechanisms.

§1.2 A brief account of the new mechanistic philosophy

In my thesis, I will primarily be concerned with the new mechanistic[1] approach to explanation and the ideas that lead up to it. It will therefore be useful to give a brief account of the mechanistic approach to explanation so that the historical discussion will be easier follow. The new mechanistic approach to explanation has become widely popular in the philosophy of the special sciences.[2] Roughly put, the mechanists argue that the behavior of complex systems can be explained by a description or account of the mechanisms that produce or underlie the behavior exhibited by the system. The behavior that is being explained is commonly referred to as the phenomenon, and the mechanism that produces the phenomenon are the explanatory relevant material *entities* and their *activities* that constitute the mechanism.

Take for example a simple mechanism, *gear-change on a bicycle*. In this case the relevant entities are: cassette, chain, derailleur, shifter, and they are engaged in certain activities: rotation (cassette and chain) and shifting (derailleur and shifter). Together, the entities and their activities constitute the mechanism, which in turn produces the

---

[1] '*New* mechanism' or '*new* mechanists' are meant to distinguish the approach from earlier talk about mechanisms and the mechanistic worldview inherited by Descartes, and classical mechanics. From here on, if it is not explicitly stated otherwise, 'mechanists', 'mechanisms' and related terms refer to the *new mechanistic* framework.

[2] The "special sciences" or "higher-level" sciences are commonly understood to be sciences that study phenomena at a higher level of organization than physics. The special sciences include everything from the life

[2] The "special sciences" or "higher-level" sciences are commonly understood to be sciences that study phenomena at a higher level of organization than physics. The special sciences include everything from the life sciences to the social sciences. Mechanistic explanation has chiefly been applied to the proximal disciplines in life sciences, e.g. Bechtel (2006) and Craver (2007), but there has been an increase in mechanistic research in the social sciences as well (Hedström & Ylikoski 2010)

phenomenon; the changing of gears. A further important point concerning mechanisms is the following:

> [The] entities and activities [that constitute a mechanism are] *organized* such that they are productive of *regular* changes from start or *set-up* to finish or *termination conditions*. (Machamer, Darden & Craver 2000, italics added)

The organization of the entities and activities is what makes the mechanism work. In the case of the gear change, the relevant parts must exhibit the right kind of organization in order to be able to change gear. If the derailleur is disconnected, the gears won't change. The regularity condition and the start to finish condition are conditions for the organization of the entities and activities.

An account of the mechanism that produces or underlies a phenomenon is said to *explain* the phenomenon. Craver and Darden (2013), among others, believe that most phenomena of interest in the life sciences (at least the proximal disciplines) can be adequately explained by an appeal to the underlying mechanism.

The mechanistic approach to scientific explanation is supposed to capture how scientists, especially in sciences like biology, neuroscience, histology, etc., *actually* explain phenomena (Craver and Darden 2013, Craver 2007, Bechtel 2006). Such a desideratum for an adequate account of scientific explanation can be contrasted with earlier models of explanation, such as the deductive-nomological model[3] (Hempel and Oppenheim 1948) and unificationist accounts (e.g. Kitcher 1989), where scientific practice and *actual* scientific explanation is of a secondary concern.[4]

As we shall see in §2, the modern debates around scientific explanation and related topics in the philosophy of science started with problems and discussions around the *DN*-model (Salmon 1989). Because of the influence of the *DN*-model upon the philosophy of science in general and especially on the philosophical debates around scientific explanation, I discuss the problems that led philosophers of science to dismiss "the received view" (i.e.

---

[3] '*DN*-model' will serve as an abbreviation for the deductive-nomological model. I follow Salmon (1989) in referring to Hempel's deductive-nomological model (and its additions) as 'Hempel's covering-law model', and not as '*the* covering-law model'. Not much will hinge upon this distinction, but it emphasizes that Hempel's model is *one kind* of covering law model.

[4] Some might here argue that the *DN*-model figures prominently in explanations in physics. The reason why this is of little concern to us here is that the mechanistic approach to explanation is first and foremost thought to capture actual scientific practice in the special sciences, especially in the proximal disciplines in the life sciences.

Hempel's DN-model) of scientific explanation and the ideas that sprung out of the discussion of these problems that led to the mechanistic approach to scientific explanation.

There are especially two motivating reasons behind the mechanistic approach to explanation: First, the appeal to scientific practice as a guide to adequate explanations. For example, biologists or psychologists seldom appeal to laws when they are explaining, and if they are appealing to generalizations (which can substitute laws in a covering-law model), they seldom present explanations as arguments. As I will argue below, there are several problems with the "received view" that has in part contributed to the development of the mechanistic account of scientific explanation. Salmon's (1984) causal-mechanical model, which I discuss in §3.2, can be seen as an attempt to meet these difficulties.

Second, the mechanistic framework is believed to be able to account for the problematic nature of multi- or inter-level explanation. The logical empiricists invoked, theory reduction (Nagel 1961: ch. 11) and global hierarchical organization (Oppenheim & Putnam 1948) as concepts to deal with the problematic nature of inter-level theories. In §3.3, I discuss Nagel's theory reduction, including briefly the most famous counter-example (the argument from multiple realizability). In §3.4, I look at the system approach to reduction and inter-level explanation as an alternative. In §3.5, I discuss the notion of *level*. In §3.6, I sum up the historical account, and show that the mechanists have adopted many of the key ideas from the reactions to the logical empiricists.

The mechanistic framework is, however, not as unified as the previous paragraphs may suggest. After I have given a detailed account of the general mechanistic framework in §4.2, which (presumably) all the mechanists would accept, I discuss several different explications of the mechanistic framework found in the literature. In §4.3, i show that some of these differences are due to the local nature of their approach. That is, because the mechanistic framework is imposed on local scientific disciplines, some of the mechanists use different terms for the same purpose, in order to yield the right kind of connotation. This is an example of how my meta-philosophical claim can be helpful.

Some of the differences, however, are more severe. For instance, Craver (2007) and Bechtel (2006) differ on the ontological status of mechanistic explanations. Adopting the distinction from Salmon (1984), Craver has an ontic view of explanation, while Bechtel has an epistemic view. The ontic conception holds the explanations are "out there", they are actual existing entities that we discover when we account for the components of a mechanism. The epistemic conception holds that explanations are *representations* of (models, arguments, diagrams, etc.) mechanism.

The nature of causation also differs between the mechanists; there are three different major accounts on the nature of causation; the mechanism view (Glennan 1996), the counterfactual view (Craver 2007: ch. 3), and the kinds of causing view (Machamer 2004, Bogen 2005).

The last difference I discuss the different understandings of the *regularity*-condition. There are several views on how we should understand regularity with respect to mechanisms. In §4.3.2, I discuss three views found in the mechanistic literature: i) that the regularity condition picks out *organizational* regularity (Machamer, Darden & Craver 2000), where this is characterized by the fact that the organization of the components of a mechanism yields predictable behavior of the mechanism from set-up to termination condition. ii) *Repeated* regularity (Leuridan 2010), the view that mechanism should exhibit recurrence, and iii) *epistemic* regularity (Machamer 2004, Bogen 2005), the view that a mechanism does not have to exhibit regularity (it can be stochastic), but that higher-level regularity is epistemically valuable for generalizations.[5] After I have discussed these differences, and taking a cue from Levy (2013), I suggest that if we look at the motivation and underlying methodology of the different mechanists, it is not surprising that we find these differences. Thus, I employ the meta-philosophical claim in order to locate the nature of their dispute. However, I will not suggest any form of unification of the views, just noting that there are perfectly good reasons why they disagree.

I end my thesis, in §4.4, by contrasting the global with the local approach to philosophy of science with a view to the meta-philosophical claim. Further, I argue that the restrictions of the local approach, i.e. that it is relative to a context or scientific discipline, have certain methodological and meta-philosophical consequences. First, I argue that classical counter-examples as those invoked against the logical empiricist may not have the desired effect on local approaches. Second, I argue that it is of significant importance to be explcit about the restriction of the domain of discourse that usually follow form a local approach. Last, I end my thesis by suggesting that much of the debate both within and outside

---

[5] Andersen (2012) offers a taxonomy of regularity in mechanisms. The idea is that the regularity condition is relative to the context of the mechanism. This allows for a mechanism to exhibit repeated, organizational and/ or epistemic (stochastic) regularity. It should be noted that she uses different terminology: 'reliable, but not exceptionless', 'sporadic', and 'infrequent'. Where the first and the last correspond roughly to my 'repeated' and 'stochastic'. Organizational regularity however, says nothing about the relative recurrence of the mechanism. So organizational regularity could exhibit all the kinds of regularity in Andersen's taxonomy. I will not go further into this discussion, but we could note that this is an example of the meta-philosophical claim. The difference on the regularity can be seen as amendable by viewing the condition of regularity as *relative to the context* the mechanism is situated in, e.g., the notion of regularity will be different in analyzing speciation as mechanism in evolutionary biology contra protein synthesis in molecular biology.

of the mechanistic literature can be seen as the lack of discussion of the intended generality of the mechanistic framework.

§1.3 What the thesis is not

The historical part of my thesis will concern the key ideas that contributed to the rise of mechanistic philosophy of science. I will *not* be concerned with how the concept of mechanism has gone through a conceptual change from Descartes until now, nor will I review in detail how mechanisms have been employed in the sciences (apart from where it is relevant).[6] It is also not meant as a survey of all the relevant literature that has led up to the mechanistic philosophy, such a task would be daunting, to say the least. I have picked out the discussion and ideas that I have found important and interesting for understanding the development of the mechanistic philosophy and a general trend towards a local approach to the philosophy of science as it is today.

---

[6] For an excellent review of the conceptual change that "mechanism" has undergone, see Bechtel & Richardson (2010 [1993]). For a review of how "mechanism" has been used in biology, see Nicholson (2011).

2 Our Logical Empiricist Heritage

> Although philosophy of science is plainly recognizable in the ancient world, in the Scientific Revolution, and in the nineteenth century, the central concerns of today's philosophy of science are primarily a legacy of logical empiricism.
>
> Lange (2007: 3)

§2.1 Philosophical reflection on scientific explanation

To understand what is special about mechanistic explanation we need to look back to the middle of the 20th century. Most of the current debates in philosophy of science can be traced to reactions to the widely popular *DN*-model of scientific explanation due to Hempel & Oppenheim (1948). This not to say that philosophy of science and interest in (scientific) explanation is nowhere to be found in earlier times. We find a distinction between explanatory knowledge as opposed to *merely* descriptive knowledge in Aristotle, what he calls knowledge-*why* and knowledge-*that.* We even find the idea that explanations or answers to *why*-questions should take the form of deductive arguments (Salmon 1989). This as we shall see, is in accordance with the *DN*-model due to Hempel.

In Mill (1843) we find another idea that is prominent in Hempel's model of scientific explanation, the idea that to explain a phenomenon is to be able to discover a law under which the phenomenon can be subsumed (Wilson 2007). Not only are the ideas about explanatory knowledge of the world in Aristotle and Mill prominent in Hempel's account, Hempel also shares the empiricist epistemological attitude of Mill, and the appeal to a secure logical foundation (similarly to Aristotle)[7] upon which such an epistemology can be placed.

While the logical positivist movement can be seen, at least historically, as a linguistic refinement of British empiricism, the logical empiricism of Hempel was a more nuanced version of the movement, where laws and generalizations, and appeal to other unobservable entities where viewed as an integral part of science (Fetzer 2014).[8] One of the many purposes of the *DN*-model of scientific explanation was its supposedly *general* or global applicability. The *DN*-model, with its logical structure, was supposed to capture all the explanations in the

---

[7] However, the logic Hempel utilizes is not Aristotelian syllogistic logic, but modern formal logic due to Frege, Russell and Whitehead.

[8] Hempel was, despite being part of it, one of the most austere critics of the logical positivist movement. Importantly, he did not accept *the verifiability criterion for meaningfulness*, because this criterion renders statements involving laws (and other concepts referring to unobservable entities that cannot be derived from concepts referring to observable entities) meaningless.

sciences. It was just a matter of deriving the deductive argument from its other form. Functional explanation in biology is a good example. Hempel (1965: 297–331) and Nagel (1961: ch. 12) both thought that functional explanation could be formulated as deductive arguments.[9] Let's turn our attention to the *DN*-model of scientific explanation, locating some of the problems with a purely logical model of explanation by looking at some general worries and counter-examples. The reactions to these problems led to the development of alternative approaches to explanation and theorizing about the nature of science, which as we shall see contain some of the key ideas of the mechanistic approach to explanation.

§2.2 Hempel's covering law model of scientific explanation

There are two key elements in the *DN*-model of scientific explanation, the *explanandum* and the *explanans*. The *explanandum* is the phenomenon to be explained, while the *explanans* are the class of sentences that together constitute the explanation of the *explanandum.* In other words, the *explanandum* is *that which is being* explained, while the *explanans* is *that which does the explaining*. Further, there are two sets of conditions that have to be met in order for the *explanans* to be an adequate explanation of the *explanandum.* (Hempel & Oppenheim 1948: 136–137)

First, the deductive conditions: The class of sentences that constitute the *explanans* must figure as premises in a deductive argument, where the *explanandum* figures as the conclusion. The *explanans*-sentences must be true and the argument must be sound.

Second, the nomological conditions: The class of sentences in the *explanans* must contain at least one law (of nature) or a universal generalization. The law or generalization has to be an essential premise in the argument. That is, the "law-premise" needs to be essential in the sense that the derivation of the *explanandum* from the *explanans* would not be valid without it (1948: 137).

Usually, when we want an explanation of something it is of a singular event, the *why*-question we want answered is concerned with something *particular*. This means that something more than sentences that are universal generalization (such as laws of nature or regularities) is needed for a derivation of the *explanandum*-phenomenon from the class of sentences included in the *explanans*. Such sentences, which provide information of the

---

[9] There were, as Hempel and Nagel noticed, difficulties in extracting the gist of the functional argument without loosing the validity of the deductive argument or making it vacuous. Hempel (1965: 297–331) and Nagel (1961: ch. 12) each offer slightly different approaches. I will not go into these difficulties in my thesis, just noting that Hempel and Nagel were aware of them and that framing functional arguments as deductive arguments has largely been abandoned in the philosophy of biology and in naturalistic functional theories of mind.

particular event, are called *antecedent condition*. These are sentences are descriptive. Hempel and Oppenheim (1948: 138) give us the following schemata for the formal structure of a deductive-nomological explanation:

(P1) $C_1, C_2\ldots C_n$	(P1) Statements of antecedent conditions
(P2) $L_1, L_2\ldots L_n$	(P2) General laws
(C) E	(C) Description of the empirical phenomena to be explained

(P1) and (P2) are the premises of the argument, the *explanans*, and (C) is the conclusion, the *explanandum*. (C) is derived, by logical deduction, from (P1) and (P2) (1948: 139). The formal structure of a deductive-nomological explanation may be successful in some cases, but in does not conform well to explanations in the special sciences.

During the latter part of the 20th century the *DN*-model was heavily discussed in the philosophy of science. Some of these discussions were on problematic concepts that already were topics of their (e.g. laws of nature), but significantly more attention was given to them in light of the discussion of the problematic cases deductive-nomological explanation. The problematic cases and counter-examples to the *DN*-model were not only from philosophers with rival accounts of explanation. Most of the critique was from people accepting the *DN*-model as *the* account of scientific explanation. The first couple of decades after the *DN*-model was proposed consisted of creating counter-examples and problems, and then finding solutions to these problems (sometimes adding to the framework, e.g. *Inductive-Statistical* model). This was when the *DN*-model (or including its additions) was considered the "received view" of scientific explanation (Salmon 1989).[10] I will consider the following problems with the *DN*-model as they point out important shortcomings of the application of *DN*-model to explanation in the special sciences:

(§2.2.1) Explanations without universal laws or regularities
(§2.2.2) The asymmetry of explanation and appeal to causation
(§2.2.3) Explanatory irrelevancies

---

[10] In some areas of philosophy of science, especially in the philosophy of physics, it is still seen as the "received view". I thank Anders Strand for pointing this out to me.

The problem cases point to the limits of a strictly logical framework of explanation. Extending the framework to include more than just deductive arguments that appeal to laws can solve some of these problems, while some problematic cases can only be adequately explained with alternative models of scientific explanation. I start with cases of explanations that do not involve universal statements such as laws or regularities and touch upon, briefly, problems with laws in general and in the special sciences.

§2.2.1 Explanations without laws or regularities

Before heading on to examples of *why*-questions that cannot be explained by an appeal to laws (and through deduction) let us briefly touch upon the notion of law as it is understood by Hempel and many of the advocates of the *DN*-model. To borrow from Woodward (2014), the basic intuitions about laws and generalizations that are in play in the *DN*-model are, roughly, that laws are true generalizations, and that true generalizations come in two classes, those that are true by "accident" and those that are laws.[11] We can distinguish these with two examples. An accidentally true generalization may be, "all the articles in my office are printed on white paper", and a law or universal (to some extent at least) generalization may be, "all gasses expand when heated under constant pressure" (to use Hempel's (1965: 339) own example). According to the advocates of the *DN*-model, what separates these two types of true generalizations in terms of explanatory power is that with a law, one can, along with some information about some particular sample of gas that has been heated under constant pressure, give an explanation as to *why* the sample of gas expanded. While with the accidentally true generalization along with information about one article in my office, one *cannot* give an explanation as to *why* the article in question is printed on white paper.

The striking feature about what Hempel calls laws is that they are thought to be *exceptionless*, while one can easily see how "fragile" the accidentally true generalization is to exceptions. If I brought an article printed on yellow paper to my office tomorrow, the generalization would be rendered false. But how can one tell them apart? There are other cases where the accidentally true generalization is harder to distinguish from universally true generalizations (i.e. laws), take for example the two statements. (G) "All gold spheres are less than a mile in diameter" and (U) "All uranium spheres are less than a mile in diameter". What distinguishes these two claims? Well, both are (presumably) true, it is highly unlikely

---

[11] Laws are sometimes formulated as those generalizations that are *universally* true. See Salmon (1989: 13–24) or Hempel (1965: 354–359) for a more comprehensive discussion around the different classes of generalization statements.

that there actually exists a gold sphere with a larger than a mile in diameter, however, it is physically possible for it to exist. The critical mass of uranium, however, makes it physically impossible for it to be one mile in diameter (van Frassen 1989). So there is a difference between these claims, and it can be spelled out using modal terms. One could appeal to necessity, counterfactual dependencies, etc. However, because of his empiricist sentiment, Hempel found modal characterization of lawhood wanting and he remarks that any attempt to fully characterize the criteria for what it takes to be a law seems somewhat unmanageable (Hempel 1965: 354–359). In the subsequent decades[12] several different accounts of lawhood have been proposed, Lewis (1973, 1986) proposed a "systems" account of laws, where one laws figures as axioms (and derived laws are theorems) in a deductive system or systems. Armstrong (1983) and Tooley (1977) appeal to laws as a relation of contingent necessitation between universals. Where, as I understand it, the universals are *contingently* universals by the fact that the universe (or a relevant domain that contain the accepted universals) *could* be different, and if it where different, the relation between the universals could also be different (and, presumably, the universals themselves). That the relation is one of *necessitation* can be understood as the fact that when the universe turned out the way it did, the relation between the universals was "set" and the relation holds by necessity after this.[13] These are examples of some of the proposed accounts of laws in the recent literature. However, many philosophers of science, especially philosophers of the special sciences, regard an antireductionist account of lawhood, where appeal to nomic, often modal, concepts (such as counterfactual conditional, disposition and/ or causation) is sufficient to describe lawhood (Carroll 2010). Many philosophers of the special sciences aren't particularly distressed by the lack of a commonly accepted definition of lawhood, as the sciences they are concerned with seldom refer to laws in their explanation, or at the very least these "laws"[14] can be cashed out in terms of what Hempel took as the central underlying assumption of lawhood, namely, that laws are exceptionless generalized statements, which describe regularities (Hempel 1965: 335–347).

Another problem with the nomological conditions in the *DN*-model is explanations that appeal to statistics, or statistical laws. The statistical laws are to be contrasted with deterministic laws. Are statistical laws, such as "smoking increases the risk of lung cancer"

---

[12] That is, in the decades following Hempel & Oppenheim (1948, and Hempel 1942, 1965).

[13] This could be called *physical necessitation* as opposed to *logical necessitation*. It is, for example, not logically necessary that $E = mc^2$, but it is physically necessary or the relation between 'E' and 'mc^2' is identical by physical necessitation.

[14] These can have the status of *ceteris paribus* laws or generalized descriptions of regularities.

explanatory? Hempel accepted statistical laws as explanatory (Hempel 1965: 376). But a problem for the deductive part of the model arises when an explanation of singular or particular cases is wanted. One cannot *deduce* that "Mr. Smith got lung cancer from smoking" from the antecedent conditions: "Mr. Smith is a human male with a respiratory system", "Mr. Smith smokes", etc. and the statistical law, "smoking increases the risk of lung cancer". As a deduction with these conditions and laws as premises do not yield the conclusion: "Mr. Smith got lung cancer from smoking". It leaves us with, if all premises are true, an invalid argument. Hempel therefore divided them into two categories: *Deductive-Statistical (DS)* and *Inductive-Statistical (IS)* (Hempel 1965: 376–381). The *DS*-model can give an explanation of statistical regularities, while the *IS*-model can give an explanation of particular facts, based on statistics and probability. The idea in *IS*-explanations is, roughly, that if the *explanans* confers high probability (more than 0.5) on the *explanandum*-outcome, the *explanans* is explanatory. There many famous counter-examples to the *IS*-model, and I will bring two of them into attention here: the "male pregnancy" counter-example due to Salmon (1971) and the "flagpole" counter-example due to Bromberger (1966).[15]

§2.2.2 Explanatory irrelevancies

Since the *DN*-model takes the form of an argument, a successful explanation could potentially cite *irrelevant* information. Salmon (1971: 34) gives us an example:

(P1) All males who take birth control pills regularly fail to get pregnant (statistical law).

(P2) John Jones is a male who has been taking birth control pills regularly (antecedent conditions).

(C) John Jones will fail to get pregnant.

The argument above satisfies the *IS*-model of explanation, but it doesn't seem particularly explanatory. Of course, though John Jones will fail to get pregnant, he doesn't have to take birth control pills in order to fail to get pregnant, it is simply physiologically impossible for John Jones to get pregnant because he is a man. There are at least two lessons to be learned here. First, information that is irrelevant should not be included in the class of *explanans*-

---

[15] See Salmon (1989: 46–50) for a list of the seven most famous counter-examples.

sentences. However, the *IS*-model has no way of distinguishing irrelevant information from relevant explanatory information.[16] Second, and this is similar to the flagpole example in the next section, the irrelevance of the intake of birth control pills could be highlighted by citing the causal process that ensue when a man (or John Jones if we opt for a singular causal analysis) takes birth control pills. We could also refer to the constitutive factors involved. A man (or John Jones) is constituted in a way such that it is physically impossible for him to bear a child. Similar considerations (i.e. causal and constitutive) take part in the flagpole example. But before turning to that example, let us briefly touch upon a suggested account that could amend for the explanatory irrelevancies in statistical explanations.

Salmon (1971) proposed that a *Statistical-Relevance* model for scientific explanations could rule out statistically irrelevant factors. In short, the statistical-relevance model states that in a class or population, *C,* an attribute a_1 will be *statistically relevant* to a different attribute a_2 if and only if the probability of a_2 conditional on *C* and a_1 is different from the probability of a_2 conditional on *C* alone.[17]

The key idea is that while statistically irrelevant properties are not explanatory, statistically relevant properties are explanatory. The fact that a property is making a difference to the *explanandum*-phenomenon is "packed out" by appealing to the statistical relevance of the property on the *explanandum*-phenomenon (Woodward 2014: §3.1). However, as we shall see in the next example, statistical relevance might sometimes do the trick, but an appeal to causes (and constitution) seems to be needed to account for the statistical relevance.

§2.2.3 The asymmetry of explanation and the explanatory relevance of causation

The flagpole example of Bromberger (1966) points out two problems for Hempel's covering law model. First, it shows us that many explanations are asymmetric, that is, cases where a derivation of the *explanandum*-phenomena from the class of *explanans*-sentences is explanatory, but a "backward" derivation is not explanatory even though it satisfies the adequacy criteria of the *DN*-model for explanation. Second, it shows us that an appeal to *causation* is in many cases vital for explanatory success. These two points are closely related, as causal relations are *by definition* asymmetric.

---

[16] At least not without further qualifications and/ or criteria.

[17] In formula: $P(a_2 \mid C \,.\, a_1) \neq P(a_2 \mid C)$

The flagpole counter-example goes as follows. Let's assume that you ask the question: "Why is the shadow cast by that flagpole *x* meters long?" A successful *DN*-explanation could use the antecedent conditions, "the flagpole is placed at point *p*", "the height of the flagpole is *h*", "the angel of the sun above the horizon is *a*", in conjunction with the laws about the rectilinear propagation of light, to derive the length *l* of the shadow cast by the flagpole. However, a "backward" derivation of the height *h* of the flagpole from *p*, *a*, *l,*[18] and the laws of the propagation of light, would, according to the *DN*-model, yield a successful explanation of the height *h* of the flagpole. This does not seem explanatory, however, since the length of the shadow of the flagpole (and the other antecedent conditions) does not causally explain why the flagpole has the height it has. To phrase it in the form of a *why*-question: "Why is the flagpole *x* meters high?" In order to answer this question, appealing to the length of its shadow tells us nothing, yet the "backward" derivation counts as a successful explanation on the *DN*- account.

What is missing here? The most common answer is that the height of the flagpole is best explained by appealing to what *caused* it to have that length. Explanations like "be*cause* that was the specification made by the individual who ordered it", "be*cause* it was made out of an oak tree that was *x* meters long" or "be*cause* that is the highest a flagpole made out of tree can be while still being robust enough to withstand the projected environmental effects", seem to be able to give us an answer to the *why*-question. The height of the flagpole could be given by an explanation that refers to the causal history of flagpole coming into being; this is called an *etiological* explanation. Alternatively, it could be explained by the constitution of the flagpole; such an explanation is called *constitutive*. For Salmon (1984) both etiological and constitutive explanations are causal.

The problems of explanatory asymmetry and irrelevance were some of the most important cases that showed the explanatory inadequacy of Hempel's covering-law model. During the latter half of the 20th century several new accounts that were suggested. The most famous of which are the pragmatic theory (van Fraasen 1980, Achinstein 1983), the unificationist model (Kitcher 1989, Friedman 1974) and the causal-mechanical model (Salmon 1984, 1994, 1997). However, since I am discussing the modern history of explanation in order to locate the key ideas that led to the mechanistic framework, I will focus on the causal-mechanical model out of these three in the following section. This model has influenced Glennan (2002), Craver (2007: ch. 5) and Bechtel (2006: 33). Bechtel is

---

[18] The point (*p*) at which the flagpole is placed, the angle (*a*) of the sun above the horizon, and the length (*l*) of the shadow cast by the flagpole.

perhaps the philosopher who is least influenced by Salmon, but as we saw in §1.2, he points out that Salmon (1984) revived and expanded the philosophical interest in mechanisms. In the next chapter, we focus on the discussions that laid the *positive* foundations for the mechanist approach, as opposed to this section which has focused on the views that where abandoned.

§3 Causality, Reduction, Emergence and Complex Organization

> We explain the successful activity of one homunculus not by idly positing a second homunculus within it that successfully performs the activity, but by positing *a team* consisting of several smaller, individually less talented and more specialized homunculi—and detailing the ways in which the team members cooperate in order to produce their joint or corporate output.
>
> Lycan (1987: 40)

§3.1 Introduction

In the second half of the 20th century, several philosophers reacted to the logical empiricist approach to philosophy of science. Salmon (1984) reacted primarily to the epistemic nature of logical empiricism, viewing scientific explanation as ontic and causal. Salmon (1984) is somewhat the "odd one out" of the positive contributors to the mechanistic philosophy as he is primarily concerned with physics, while the rest of the philosophers that I include as positive contributors are philosophers of the special sciences who advocating a *local* approach to the philosophy of science.

We shall begin, in §3.2, with an account of Salmon's (1984) causal-mechanical model of explanation. However, we are not done with logical empiricism yet. In accounting for the local alternative to reduction provided by philosophers in the systems tradition, we need first to discuss what this approach to reduction was a reaction to. They reacted to the epistemic and theoretically monistic account of theory reduction due to Nagel (1961: ch. 11). In §3.3, I give an account for the Nagelian approach to reduction so that we can appreciate and understand the motivation and methodological shift away from the global approach of logical empiricism, by the system tradition. I also include a short account of the most famous argument against theory reduction, the argument from multiple realization, even though it is not necessary for the system approach. It is, however, an argument for the autonomy of the special sciences, as such it is a defense of a local methodology, though from a different perspective.

The next step, in §3.4, is to discuss the novelties of the systems approach to reduction an inter-level explanation. The mechanistic framework shares many of the ideas developed in this tradition. For example, Bechtel (2006: ch. 1–2), Craver (2007: ch. 5), Glennan (2002), and Machamer, Darden & Craver (2000) hold that mechanisms are complex systems, and that the central task inter-level explanation is *decomposition* and *localization*.[19] Similarly they hold that complex organization of parts in a system can yield novel properties at a higher level. It should be noted, however, that decomposition for the mechanists is thought to be non-reductive (e.g. Craver 2007: 15–16 & ch. 4), while the system tradition generally speak of it as scientific reduction, where the mechanists usually call it integration (e.g. Craver 2007: ch. 7). This difference is merely terminological and can be amended by paying attention to the methodology.[20] What Craver (2007: 15–16) calls a reductive approach is a fundamentalist approach, either as theoretical monist or as denial of any form of ontological emergence (see §3.4). The system approach and the mechanists agree that novel properties of a system can appear because of *organization* of the parts of a system, but denies that such properties are irreducible. We shall call this kind of emergence *organizational emergence*. We shall return to the similarities (and differences) between the systems tradition and the mechanists, and Salmon (1984) and the mechanists in §3.6.

Before that, in §3.5, I turn to the notion of *level*, and different approaches to hierarchical organization, not in the systems tradition. Yet again we will turn to the logical empiricists, namely Oppenheim & Putnam (1948), and see that their version of hierarchical organization is global as well. The discussion levels and hierarchies is, as with the argument from multiple realization, not directly related to the mechanists. However, it provides an example that the shift in methodology from a global approach to a local is trend in the philosophy of the special sciences *in general*, and not just in the system tradition and mechanistic philosophy.

§3.2 The causal-mechanical model of explanation

As we saw in §2.2.2, Salmon (1973) proposed a statistical relevance model to account for the problem of explanatory irrelevance that faced Hempel's covering law model. However, Salmon (1984) later gave up his statistical relevance model for a causal-mechanical approach to explanation. It should be mentioned however, that Salmon did not view his statistical-relevance model as problematic or incoherent, he just noted that it was insufficient as an adequate account for scientific explanation. He writes:

> Subsequent reflection has convinced me that subsumption of the forgoing sort [subsumption under statistical relevance relations] is only part—not all—of what is involved in explanation of particular

---

[19] See also Bechtel & Richardson (2010 [1993]) and Craver (2001).

[20] That is, we invoke the meta-philosophical claim.

> facts. It now seems to me that explanation is a two-tiered affair. At the most basic level, it is necessary, for purposes of explanation, to subsume the event-to-be-explained under an appropriate set of statistical relevance relations, much as was required under the S-R model. At the second level, it seems to me, the statistical relevance relations that are invoked at the first level must be explained in terms of *causal* relations. The explanation, on this view, is incomplete until the causal components of the second level have been provided. This constitutes a sharp divergence from the approach of Hempel, who explicitly rejects the demand for causal laws (1965, pp. 352–354).[21] (Salmon 1984: 22)

Salmon immediately remarks that the statistical relevance model could be seen as "statistical analysis" and have important uses, but falls short of capturing scientific understanding.

Salmon separates himself from Hempel's covering law model by arguing that he offers an *ontic* model of explanation, while Hempel's model is *epistemic*.[22] This can be exemplified by Salmon's suggestion to replace Hempel's wish to locate empirical phenomena in the *nomic* nexus (nomic being a class of epistemological concepts) with locating empirical phenomena in the *causal* nexus (the causal nexus refers to *ontological* patterns/processes in the world). Now that we have contrasted Salmon (1984) to Hempel, let us look at the causal-mechanical model of explanation in detail.

Salmon distinguishes between two kinds of explanation: (1) *etiological* explanations and (2) *constitutive* explanations.[23] Both kinds of explanation are causal in nature. It should be noted that Salmon has a particular view of ontology and causation. It is a causal process theory, commonly called the "mark transmission account of causation". A process, for Salmon, is an entity maintaining a persistent structure through space-time; a *causal* process is a process capable of *transmitting* a mark or change in such structures. For example, inflicting

---

[21] Interestingly, Wimsatt (1976b) accepts Salmon's (1971) SR-model for explanation with some provisions. One of which is that in finding the statistically relevant factors/ partitions, "we are doing so with the aim of partitioning the reference class into *kinds of mechanisms*, or kinds of cases involving a given mechanism. (I am giving a realist interpretation to his model). … [A] search for factors … ties in naturally with a view of [functionally reductive] explanation as a search for the mechanisms which produce a given phenomenon, and as an account of how they do it." (Wimsatt 1976b: 686). Wimsatt, as we shall see, precedes Salmon (1984) own view! This seems to have gone unnoticed in Salmon (1984). Wimsatt's provision also indicates a strikingly similar view of explanation as the mechanists. Though not spelled out in as much detail (see for example Craver 2007: 121–122, Craver use 'exhibits' where Wimsatt use 'produce', or Bechtel & Richardson (2010 [1993]: 17) use "behavior of the system" where Wimsatt use "phenomenon").

[22] Salmon (1984: ch. 4) separates three different basic conceptions of scientific explanation, *epistemic*, *modal*, and *ontic*. In short (in deterministic cases), on the epistemic conception there is a relation of *logical* necessity between the class of *explanans*-sentences and the *explanandum*-description (deductive logic). On the modal conception, a relation of *physical* necessity holds between particular antecedent conditions and the explanandum-phenomena (general laws). On the ontic conception "we might say that *to explain an event is to exhibit it as occupying its* (nomologically necessary) *place in the discernable parts of the world*." (1984: 18). I shall have more to say about the distinction between *epistemic* and *ontic* conceptions later on.

[23] One could also phrase this as two *aspects* of an explanation if the *explanandum*-phenomenon is the same for both the *etiological* and the *constitutive* explanation. They will just answer different *why*-questions regarding the same phenomena.

force on a car can "mark" the car with a dent or a scratch, but inflicting force on the car's shadow will not "mark" the shadow but rather the structure on which the shadow appears. In this way Salmon tired to differentiate between processes that are *causal* with those that are merely e*piphenomenal*. Lastly, a causal *interactions* are the intersections between causal processes where the "marking" or change in structure occurs (Salmon 1984: ch. 5 & 6). The causal interactions are thought to account for Hume's *secret connexion*, i.e. what *makes* the structure change (Salmon 1984: 169). In Salmon (1984), the causal interactions were defined in terms of a counterfactual criterion of mark transmission. However, in (Salmon 1994) he develops an account of causal interactions as the exchange of conserved quantities. I will not go further into Salmon's account of the nature of causation here, as (most of) the mechanist we will discuss later on have rejected this type of causal theory, and rely on non-reductive accounts of causation (Craver 2015).[24]

Even though it is the constitutive aspect of Salmon's causal model of explanation that has been most influential on the mechanist, it will be helpful for later discussions to explicate the etiological aspect as well.

Etiological explanations are the descriptions of the causal relations that led up to the *explanandum*-event. In other words, it is a causal history of the antecedent (causal) conditions that led up to the event that is being explained. It answers *why*-question of the sort "why did the *explanandum*-event occur?" or "why is this the case rather than that?" by an account of the *explanandum*-event's causal history. Salmon's own example is the (etiological) explanation of a carbon dated 30,000-year-old bone with human carvings found in Alaska. The problem is that archeological and paleontological evidence does not record any human activity in Alaska until 12,000 BC. In order to solve this problem, Salmon suggests that a look at the relevant causal history of the bone could yield an explanation. Because of the difficulty in verifying or falsifying the causal history of a particular bone, this explanation is conjectural, but many archeological and paleontological explanations share this feature. An explanation of the human carvings on a 30,000-year-old bone based on the causal history of the bone and its interaction with its environment (including the existing archeological and paleontological evidence) could be as follows: "A mammoth was trapped in a crevasse during

---

[24] See Salmon (1984: ch. 5, 1994) or Dowe (1992) for further discussion of causal process theory. Glennan (1996) is a notable exception to the non-reductive tendency towards causation in the mechanistic literature. For examples of non-reductive causal theory see Bogen (2005), Machamer (2004) and Craver (2007: ch. 4). Craver bases his analysis of causality on Woodward's (2003) interventionist theory of causation.

an ice-age and was frozen. 25,000 years later the ice receded and a prehistoric human found the bound and fashioned it into a tool."

Such explanations are perhaps more often present in sciences that explain historically. Evolutionary explanations (or explanations by natural selection) often take such a form as well, and they are not only present in evolutionary biology, but figures in psychology, anthropology, sociology, etc. as well.

Constitutive explanations, on the other hand, do not refer to the causal history or antecedent conditions that led up to the *explanandum*-phenomenon. These explanations describe or account for the underlying causal mechanisms that *constitute* the explanandum-phenomenon. Let's illustrate with an example: "Why does salt dissolve in water?" In this case an appropriate explanation refers to the chemical properties of the molecules that *constitute* the water and the salt. Water and salt are constituted of polar molecules. Polar molecules are molecules in which the electrons are unequally distributed between the atoms. This causes a partial charge on the atoms of the molecules. In water the electrons oscillate to a greater extent around the oxygen atom, making the oxygen atom partially negatively charge and the hydrogen atom partially positively charged. Likewise, in salt the electrons oscillate to a greater extent around the chlorine atom, making the chlorine atom partially negatively charged and the sodium atom partially positively charged. When the salt is put into the water, the negatively charged side of the water molecules surrounds the positively charged sodium, and the positively charged side of the water molecule surrounds the negatively charged chlorine. This breaks the bond between the sodium and chlorine atom and the salt dissolves. We should note that, for Salmon (1984: 270), molecules are causal processes. They participate in causal interactions, (e.g. surrounds, oscillates, bonds). So in this explanation the molecules that constitute the water and the salt are the *underlying causal mechanisms*.

Constitutive explanations explain by showing *how* the *explanandum*-phenomenon occurs. It answers *why*-question by appealing to the (sometimes sufficient and/or necessary conditions) underlying causal relations that *produce* the *explanandum*-phenomenon. An easy and intuitive, though not entirely accurate way to distinguish these two kinds of explanations is be saying that constitutive explanations tell you "how it works", while etiological explanations tell you "why it's there". It is the constitutive explanations that are of special importance to the mechanists, as this kind of explanation is necessary in accounting for the components of a mechanism.

§3.3 Theory reduction

In order to satisfactorily understand why the system approach is an interesting development in the philosophy of science we need an account of what it was a reaction to. The logical empiricist, as we saw §2, viewed explanation as arguments, and the outstanding examples where explanation as *deductive* arguments. So when Nagel (1961: ch. 11) developed a theory of reduction, it was made in the logical empiricist image. The inter-level relations between theories should be reduced by logical derivation, and the ultimate derivation is from the fundamental science, which for the logical empiricist was fundamental physics. The logical empiricists are *in-principle* theoretical monists. That is, if it were possible the unity of science would be achieved if all theories could be formulated in the language of a fundamental science. The most famous argument against to theory reduction, the argument from multiple realization (Fodor 1974), is an argument against the logical empiricist unity of science. Even though multiple realization has had a limited influence on the system approach (and consequently only indirectly on the mechanists), it will be valuable to give a brief account of the concept and argument so we can talk about it freely at later occasions. As it turns out, multiple realizability is possible in the system approach to reduction. However, let us return to Nagel's account of reduction. This is his preliminary remark on the concept of theory reduction:

> Reduction, in the sense in which the word is here employed, is the explanation of a theory or a set of experimental laws established in one area of inquiry, by a theory usually though not invariably formulated for some other domain. … [W]e shall call the set of theories or experimental laws that is reduced to another theory the "secondary science," and the theory to which the reduction is effected or proposed the "primary science. (Nagel 1961: 338)

Such reductions come in two forms, *homogeneous* and *heterogeneous*. Homogeneous reduction involves the reduction of theories in which the secondary science employs descriptive terms that are the same (or have approximately the same meaning) as in the primary science. The reduction may for example be to a theory that has a larger domain or range of application. Nagel's example illustrates such a reduction. The theories Galileo formulated in order to explain the physics of free-falling terrestrial bodies where integrated in Newtonian mechanics and gravitational theory. With Newtonian mechanics and gravitational theory it became possible not only to explain the physics of motion of terrestrial bodies, but also of celestial bodies.[25] Thus, Galileo's theories could be reduced to Newtonian mechanics and gravitational theory (an area of inquiry with a larger domain or larger range of application). Homogeneous reductions are reductions that establish deductive relations between to sets of statements that use descriptive terms that have the same (or an approximately similar) vocabulary, i.e. a homogeneous vocabulary. Nagel sees such reductions as indications of normal scientific development, and regard them as unproblematic (1961: 339).

Heterogeneous reductions, on the other hand, are more controversial. These are reductions where the secondary science employs different descriptive terms than the primary science, i.e. they employ heterogeneous vocabularies. Theories in the secondary sciences are designed to explain *qualitatively different phenomena* than the secondary sciences. On such a reduction, the theories from the primary science will explain the macroscopic processes, usually explained by the secondary science, by reference to the microscopic constituents that make up those processes.[26] In other words, in a heterogeneous reduction, not only will the descriptive terms, i.e. theories, be different, but the explanation of the primary science might even refer to *phenomena* that are not referred to in the secondary science as well. In order to successfully establish deductive relations between sets of statements with *heterogeneous* vocabularies, criteria are needed for the deductions to be valid and sound.[27]

There are two formal conditions that together form the necessary and sufficient conditions for heterogeneous reduction. These are the *condition of connectability* and *condition of derivability* (1961: 354). The condition of derivability states that the descriptive statements (e.g. laws, theories) in the theory of the secondary science are deducible, by first order logic, from the descriptive statements (e.g. *fundamental* laws, theories) in the primary science. The condition of connectability is supposed to make derivability possible when the secondary science includes descriptive terms that do not figure in the primary science (or

---

[25] In Galileo's time, the study of celestial bodies was a different discipline than the study of terrestrial bodies (Nagel 1961: 339).

[26] The "constitution for those macroscopic processes" (1961: 340) is not further defined, but as I understand Nagel, these will be variables that are given a value relative to entities in the currently accepted theory of the primary science (in the relevant context). Because of the logical nature of theory reduction, it is only in *instances* of reduction that we can talk about the properties of the constituents or what it takes for something to constitute a higher-level phenomenon. However, if the goal is to reduce everything to the most fundamental science, then the constituents will *ipso facto* be the entities or processes of the fundamental science (presumably some kind of theoretical physics). This will be important when we turn to the rejection of theory reduction by an argument from organizational emergence (see §3.4.1).

[27] Nagel uses the reduction of thermodynamics to statistical mechanics (in conjunction with the kinetic theory of matter) as an example of a successful heterogeneous reduction. It is showed that if you assume the mean kinetic theory of gases, in conjunction probabilistic approach for analyzing the location and motion of the individual molecules (in a "phase-cell") of an ideal gas in a container, you will be able to derive by Newtonian mechanics (with statistical mathematical procedures) the Boyle-Charles' law (1961: 342–345).

when their meaning is different). The problem is; what is the nature of the connection between two heterogeneous vocabularies? In other words, what is supposed to do the job of connecting the theory of secondary science to the descriptive terms in the theory in the primary science so derivability is possible? According to Nagel, in heterogeneous reduction, it is assumed that there are statements that include descriptive terms from both the secondary and the primary science that makes a derivation of the theory in the secondary science possible from only terms in the primary science and the relevant statements that include descriptive terms from both the secondary and the primary science, these are commonly called "bridge laws", "reduction functions" or "connectability assumptions" (Sarkar 1998: 25).[28] These bridge laws are universally quantified biconditional *or* conditional[29] statements that can be interpreted as conventions or as facts. However, reference to empirical evidence is needed for bridge laws to count as facts (Nagel 1961: 354–355, Sarkar 1998: 25).[30] summing up, inter-theoretic reduction is the reduction of a theory, T_S, by logical deduction from a more fundamental theory, T_P. If T_S is formulated with descriptive terms that do not figure in T_P, bridge laws that enable a logical deduction are necessary.

Multiple realizability arguments attack the belief that natural kinds in higher-level theory can be adequately described in terms of the vocabulary of the lower-level theory and bridge laws. It is denied by an appeal to cases where the natural kind in the higher-level theory can be realized by several different kinds of lower-level entities (or microscopic constituents), making derivation invalid. These kind of arguments originated in the philosophy of mind with Putnam's (1967) argument against type-identity theory of mental states with physical states. Putnam argues against the reduction of *types* of mental states to *types* of physical states, that is, against the view that some type of mental state, such as pain, is identical to some type of physical state, such as a particular pattern of neural activation. This, however, still allows for particular instances of mental states, such as an instance of pain, to be identical to a particular instance of a physical state, such as a particular instance of a pattern of neural activation.

---

[28] These are *not* Nagel's terms; nowhere in his book (1961) does Nagel refer to bridge laws. The notion of "connectability assumptions" is, as I understand Nagel, perhaps what is closest to capture what Nagel has in mind. However, since "bridge laws" have become the standard term, I will employ that.

[29] The connectability condition is entailed by the derivability condition, at least according to Nagel. Thus the bridge laws do not have to take the form of biconditionals as many have believed (e.g. Kim 1992), but may be conditionals (l961: 355, n. 5 – thanks to Anders Strand for pointing this out).

[30] Since Nagel has an epistemological view of reduction, i.e. that it describes the logical relations about our knowledge, whether or not the bridge laws are facts or conventions will depend upon the context, and they are not mutually exclusive either. It could be a convention that, after empirical evidence was provided, turned out to be a fact.

Fodor (1974) generalized the same kind of argument to be an argument against reduction of natural kinds in the special sciences to natural kinds in physics. One of his examples is Gresham's law in economics. It states that currency of lower intrinsic value tends to circulate more freely than currency of higher intrinsic value and equal nominal value. In a formulation with ha little more punch; "Bad money drives out good". According to Fodor (1974), whether the bad money whether the bad money is paper-bills, coins, or food stamps and the good money is gold, silver or baseball cards have no influence on Gresham's law, the law holds with a multitude of different physical constitution. So a logical derivation of Gresham's law by appeal to it microscopic or lower-level physical constituents would be invalid (unless the bridge laws consisted of a infinite disjunction of actual and possible physical realizations of the bad and the good money). Multiple realizability arguments have had an enormous influence in philosophy, especially as arguments for non-reductive physicalism in the philosophy of mind and as arguments for the explanatory autonomy of the special sciences in the philosophy of science. The systems approach to reduction, as we will now turn our attention to, is not directly an argument against theory reduction, but is presented as an alternative take on reduction, which retains the explanatory autonomy of the higher-level sciences (some authors explicitly endorse theoretical pluralism, e.g. Wimsatt (1972), while it is for the most part implicit). The guiding idea is that scientists, especially biologists, frequently appeal to reduction but in a very different way than how the logical empiricist takes reduction to occur in the sciences (Wimsatt 2007 [2000]).

§3.4 Systems approach as an alternative take on reduction

The systems approach to reduction and inter-level explanation, as I mentioned above, can be seen as a reaction to the (global) logical empiricists *DN*-model of explanation and the theory reduction that followed. Several philosophers of the special sciences saw an appeal to reduction and explanation *as it figures* in the special sciences as a central guide to the question of reduction and inter-level explanation.[31] The change of perspective is characterized by the view that inter-level explanation can be thought of as analysis of the functions or behavior of a complex system at a high level of organization in virtue of articulating or describing how the lower-level parts (in conjunction with their interactions and organization) that constitute the system are capable of realizing or producing the relevant

---

[31] See fn. 32 for references.

function or system-behavior.[32] These approaches where especially influential in the philosophy of biology and psychology, and has been labeled the *systems approach* to or *systems tradition* of explanation (Craver 2007: 109). It is a reductive approach, because it explains higher-level phenomena by appealing to lower-level constituents. However, the systems approach emphasizes that the higher-level phenomena cannot be *fully* explained by the lower-level constituents alone, there are several restraining conditions or boundary conditions placed on the system at a higher level that must be appealed to or included in order for a complete explanation to be achieved (Campbell 1974, Wimsatt 1972). One does not achieve an adequate explanation just by describing the lowest-level, one need to appeal to the hierarchical organization of the system, where each level may be restrained by lower- *and* higher-level boundaries. In science (at least the special sciences) the usual way of describing reduction is more akin to the reductive approach mentioned here and *not* by theory reduction (Wimsatt 1976b). Biology, for example, usually studies wholes that are *more* than the sum of its parts. This is a key idea for the system approach, namely that if the organization of constituents at the receding levels in a hierarchical system is complex, the whole can be more than the sum of its parts. Such hierarchically organized systems are called *complex* systems, and these are what the system approach are interested in.[33] Hierarchically organized systems that are organized in such a way that the whole is equal to the sum of its parts are called simple systems (Wimsatt 1972, 2007 [2000]). Wimsatt (1972) distinguishes between two types of complexity and simplicity in a system, *descriptive* and *interactional*. Essentially, interactional complexity is a myriad of different causal interactions between the parts of a system, and interactional simplicity describes a system with few interactions between its parts. Descriptive complexity is complex compositional organization of a myriad of different parts, which make up stable sub-assemblies in a hierarchical system. Living systems are the principal examples. Descriptive simplicity is a whole which is equal to the sum of its parts, the prime example being a polymer, which has a simple hierarchical organization. (Simon

---

[32] There are different notion of how to *analyze* the complexity or lower-level parts of a system. In the philosophy of biology Wimsatt (1972, 1976), following Simon (1962, 1996 [1969]: ch. 8), argues for an analysis by (nearly) *decomposing* the complex systems into parts and/or process. Kauffmann (1971) argues for pluralistic analysis by articulation of parts explanations of a system. In the philosophy of psychology, Cummins (1975, 2000) argues for a functional analysis of relevant parts that realize or produce the phenomenon, while Dennett (1994) argues it should be analyzed by way of reverse engineering. These different approaches all share an overarching methodological approach – namely the appeal to the way parts and their interactions are *organized* to produce the phenomenon to be explained.

[33] A clarification is in order here. While the logical empiricist theory reduction relates laws, theories and properties at different levels, the systems approach concerned with the individual parts and their relations at the different level of a system. So the *relata* in the hierarchy (what are placed at different level in a hierarchy), are ontological entities for the system-approach and (primarily) epistemic entities for the logical empiricists.

1962, Wimsatt 1972). Simon (1962), who is probably the most influential writer in complex systems theory, characterizes complexity roughly as follows:

> [B]y complex systems I mean one made up of a large number of parts that interact in a non-simple way. In such system the whole is more than the sum of the parts, not in an ultimate metaphysical sense [e.g. vitalism], but in the more important pragmatic sense that, given the properties and parts and the laws of their interaction, it is not a trivial matter to infer the properties of the whole. In the face of complexity, an in-principle reductionist may be at the same time a pragmatic holist. (Simon 1962: 468)

Such an account of complexity led to the denial that (system) properties can only be epistemologically or nominally emergent, allowing for system properties or behavior of complex system to be organizationally emergent (see §3.3.1). Simple-system emergence (i.e. cases where one can infer deductively the emergent property by its constituents) is still emergence it is just thought to be an uninteresting form of emergence (Wimsatt 2007 [2000]).

We have thus distinguished three reactions to the logical empiricist framework for reduction. First, *the ontological reaction*, i.e., the acceptance of organizational emergence, second *the epistemological reaction*, i.e., the view that *complete* (reductive) explanation is achieved not only by lower-level entities and their interactions, but also by the *organization* of the lower-level parts and interactions and the boundary conditions of the system. Lastly, we have *the methodological reaction*, i.e., the view that philosophy of science is, and should be, *local*. That is, one should allow for theoretical pluralism, and the autonomy of the special sciences. Another feature that regularly appear with the local approach to philosophy of science is the view that scientific practice is a valuable guide, especially to reduction and explanation. In the following sections I go through these reactions. I will try to keep the discussion at such a general level that most or all of the philosophers of the systems approach would accept the doctrines and views I assert them to have.[34]

§3.4.1 The ontological reaction

Emergence for the logical empiricist is formulated as the thesis that a hierarchical organized system exhibit properties at a higher level of organization, by virtue of complex inter- and

---

[34] See fn. 32. I will, in other words, try to show what I take to be the common features between the views of these authors. I will not, however, refer to all the authors each time a reference is needed, but I will refer to the papers were the discussion is particularly similar to mine. Suffice to say, the authors mentioned in fn. 32 are all, to some degree or another, representative for the views I discuss here. I will, however, refer to particular parts of my discussion to the authors that I lean most on.

intra-level relations, which are *non-predictable* from properties found only at lower levels (Nagel 1961: 367, Hempel 1965: 258–264), call this *epistemological* emergence.

Another way of characterizing emergence is the thesis that the mere putting together of the lower-level parts that constitute a hierarchically organized system *will not* produce the properties exhibited at the higher-level of organization (Simon: 1996 [1969], ch. 7), call this *ontological* emergence. As an approach to explanation ontological emergence is a strong, global and stratified *holism*.[35] The epistemological conception of emergence is not compatible with any form of reductionist approach, and neither of the philosophers mentioned in this section have any wish to hold on to it. Epistemological emergence is largely a product of theoretic reduction due to Nagel (1961: 366–380)[36] and is "nothing more than a more than temporary confessions of ignorance" (Wimsatt 2000 [2007]: 274]).[37] If a standard reduction is successful, then the postulated emergence will be explained away. The notion of ontological emergence as I have characterized it here is also too strong, the reason it is too strong is because it excludes the possibility of explaining the higher-level behavior of a systems by appeal to its parts and processes. Simon notes:

> Applied to living systems, the strong claim [ontological emergence] … implies a vitalism that is wholly antithetical to modern molecular biology. Applied to minds in particular, it is used to support both the claim that machines cannot think and the claim that thinking involves more than the arrangement and behavior of neurons. Applied to complex systems in general, it postulates new system properties and relations among subsystems that had no place in the components; hence it calls for emergence, a "creative" principle. Mechanistic explanations of emergence are rejected. (Simon 1996 [1969]: 170)

Such implications for emergent properties (or system behavior) are not compatible with the system approach either. The core idea behind the system approach is, as noted, that the interactions among system components can account for system behavior. It is therefore necessary, if the system approach is to distinguish itself from the *reductive* approach and the

---

[35] By 'stratified' in relation to holism I mean the dividing of objects into discrete wholes (strata), where each strata or whole exhibit novel properties not reducible to the parts of which it is composed (i.e. *ontological* emergence). By global I mean that the holism extends to all the objects of the universe. I will refer to this view as *holism*, exempting myself from other uses of the term. What I call the *holistic approach* below, is an approach to explanation that considers the behavior or properties of a system as a whole as thoroughly *non-reductive*, meaning both that it is more than the sum of its parts and organization *and* that it can only be explained by an appeal to a theoretical language at the system-level.

[36] It should be noted that Nagel distinguishes between two types or formulations of emergence. (1) emergence as the *non-predictability* of higher-level properties, and (2) emergence as a diachronic cosmogonic process. We shall be concerned with the first, as this is the type of emergence that initiated the systems approach. Most notable in criticism of the epistemological view of emergence is Wimsatt (1972, 1976, 1980, 2007).

[37] This is Wimsatt paraphrasing what a classical reductionist would take (epistemological) emergence to really be.

holistic approach (see fn. 35), to formulate a notion of emergence that is balanced between these two extremes. On the one hand, it needs to keep a (though restricted) reductionist sentiment, while on the other it must allow for novel properties (realized by component parts and their relations). A third formulation of emergence is needed, one that provides some degree of autonomy of the different levels of organization or, in other words retaining the boundary conditions for objects or components within a system (Campbell 1974, Wimsatt 1972), while retaining a reductionist sentiment. So, *organizational emergence*, the systems approach's formulation of emergence (Simon 1999 [1969] ch. 7), can be formulated as follows:[38] Emergence is the existence of higher-level properties or behaviors of a system, which is realized by (or explained by) *mutual relations* between lower-level parts or components that do not exist in isolation (Simon 1996 [1969]: ch. 7). Wimsatt (1976, 1979, 2000) follows Simon (1962, 1996 [1969]) and views emergence as the dependence of a system property on the *organization* or *arrangements* of its parts. The whole is greater than the sum of its parts, but *not* due to some mystical vital essence or something external to the system, but rather because all the relevant parts and interactions of the system are arranged in a certain manner such that the system property is realized or produced.

Let's look at a simple example; two celestial bodies are arranged such that they exhibit gravitational attraction. Each body, by itself (i.e. in isolation), does not exhibit gravitational attraction, so the gravitational attraction between them depends on the arrangement plus the constitution of the bodies. In more complex systems such 'novel' system properties are realized at several levels of organization, so the reduction is only possible at the immediately lower-level, and that level again is only reducible to the next lower-level and so on.[39]

Organizational emergence is first and foremost an ontological doctrine;[40] it says something about how the higher-level properties of a system are *realized* or *produced*. While this is in opposition to the standard reductionist epistemological conception of emergence, Wimsatt (1974) maintains that it is still a reductionist approach. A lower level of organization, and the intra-level organization of the components located at that level, can be appealed to in order to *explain* a higher-level phenomenon. However, it will not be "fully fledged" reduction (in the epistemological sense), because these explanations appeal to the

---

[38] Simon (1996 [1969]: ch. 7) calls this emergence in a weak sense, I have chosen to characterize it as organizational emergence, as 'weak' emergence is used differently is modern philosophy.

[39] This is characteristic of mechanistic hierarchies as well. See §4.2.1

[40] It is ontological by the fact that it says something about the nature of complex system. It is not to be confused with what I have called *ontological emergence*, which is a strong, non-reductive formulation of emergence.

realization of the higher-level properties by the lower-level parts and their interactions *and* their organization. The explanations will therefore need terms appropriate for the relevant levels, and these terms must therefore be relative to the system-property being explained, and each given level corresponds (presumably) to one or more scientific disciplines.[41] This leads us over to the *epistemological* reaction to the logical empiricists account of (reductive) explanation.

§3.4.2 The epistemological reaction

A consequence of organizational emergence is that there are certain inter-level relations (i.e. mutual relations that obtains in virtue of the complex organization) that along with the lower-level parts and interactions realize or produce the higher-level system property or behavior. This means that in order to give a complete explanation of a the higher-level phenomenon, one must enlist the both the lower-level parts and interactions (the reductive aspect), the mutual relations that obtain between them when they are organized in a particular way and thereby realize or produce the higher-level phenomenon (the organizational aspect) and lastly the function[42] or behavior of the higher-level phenomenon (the functional aspect).

(A) The *reductive aspect* is the process of *localizing* the lower-level parts and interactions that are the *relevant* constitutive[43] and causal conditions for the *inter-level* relation to obtain (Wimsatt 1972, Cummins 1975). In a very simple example, the production of cars in a factory), the lower-level parts and interactions that produce the phenomena (i.e. production of, say, cars) are the workers, the parts they work with (components) and the assembly line. The causal relationship is the movement of the assembly line and the workers doing their job. The constitutive relationship is the parts' *disposition* for assembly (and the fact that the worker, assembly line and parts are at size which makes the relevant causal relationship possible) (Cummins 1975).[44]

---

[41] An example of this can be seen with the attempt at theory reduction of Mendelian genetics to molecular genetics. See Hull (1972) a brief account of the problems relating to terms used at different levels of organization in this debate. Darden (2006: ch. 4) discusses these problems in relation to mechanistic philosophy, arguing that while reduction (that is, theory-reduction) is not possible, mechanistic *integration* is.

[42] The system properties or behavior in biological systems are often *functions*. Several of the philosophers in the system tradition explicitly deal with functional analysis *as* a form of system-approach to explanation (e.g. Cummins 1975, 2000).

[43] I here use the word constitutive, and *not* compositional, in order to emphasize that it is the facts about the part's *structur*e and not their *organization* in the system that is under discussion.

[44] See Cummins (1975: 757–758) for a discussion on dispositions of parts relative to functions. The assembly line example is adapted from Cummins, (1975: section 3.2)

(B) The *organizational aspect* is the higher-level feature that obtains when the causal and constitutive conditions are met and are *organized* in a particular way such that the system property is realized (Wimsatt 1972). In the factory example, it is the finished product, i.e., a car. Note here that if we were to switch out the workers (or give them different orders) and supply them with different parts we could get different products.[45] The products could be slightly different cars or even airplanes or dolls. It is the *organization* of the parts that determine the finished products, thus multiple realizability is not a problem.

(C) The *functional aspect*: The higher-level phenomenon is realized by (A) and (B), and because of (B) we cannot reduce it to merely a description of the parts and their interaction. Thus, the functional aspect explains what the higher-level phenomenon does in the system, e.g. what function it has in the system. In the factory example, it explains why the factory makes car. This answer to this could be because that's what the CEO decided to make, and that's why he bought the assembly line and the relevant the parts, and hired the workers with the right competence. Such answers are often called *etiological* approaches to functional explanation. Another answer to why the factory makes cars is to cite (1) and (2), i.e., to explain it by an appeal to the reductive and organizational aspect. In this case factory makes car because of how the assembly line, workers, parts and how factory is organized. This is commonly called the *dispositional* approach to functional explanation (McLauglin 2001: ch. 5 & 6).

In the system tradition, there is no agreement on whether one is better, more right, etc. than the other. Some claim that the etiological approach is insufficient and even parasitic upon the dispositional (Cummins 1975),[46] while others seem to prefer the etiological approach (Simon 1963)[47]. I will not take sides in this debate. Godfrey-Smith (1993) suggests that they're both useful, each for their own use, and I'm inclined to think the same.

§3.4.3 The methodological reaction

The reaction to what I have characterized as the global methodology of logical empiricism can be seen as an encouragement to view the practice of the special sciences, in particular biology and psychology, as helpful in understanding the different aspects of the philosophy

---

[45] We could even just ask them assemble the parts differently and we would get a car that looks and behaves differently.

[46] See especially Cummins (1975: 749, fn. 5), but also Cummins (2000).

[47] See especially his discussion on the evolution of complex systems and the parable of *Tempus* and *Hora*. However, Simon (1963) does nowhere exclude the possibility that some phenomena are best explained by a *dispositional* approach. Indeed, his view seems to encourage such approaches when the explanation does not include an evolutionary aspect.

of science. As we have seen, the system approach modeled their approach to reduction, not as a primarily logical and epistemic activity, but as the decomposition of actual living systems. They are giving an abstract account of how scientists approach reduction (Wimsatt 1976ab). There are several important consequences of we have characterized as organizational emergence.

First, as we seen, organizational emergence makes the higher-level system property not *completely* explicable in lower-level terms. This warrants the explanatory autonomy of the special sciences concerned with these higher-level system properties. We can show, by systems reduction, what the sufficient conditions for the system property is, but an explanation of the function or behavior of the system property relative to external factors or as a part of a larger system, must be at the higher-level.

Second, different systems or different parts of systems are decomposed differently. This yields a pluralistic methodology of decomposition and localization, not privileging either a bottom-up or a top-down approach to the discovery of the parts of a system.

For example, if one is interested in finding out how to build a bicycle one may start studying the bicycle's behavior as a whole. When one has found a peculiar or interesting function, say the fact that the pedals don't move forward when the bike is moving forward, but they move backwards when the bike is moving backwards, we locate the parts, their interaction and organization. In this example, it would be reasonable to start with the hub as a system, and "freewheeling"[48] as the particular system property or function. Then, after understanding how the hub enables the freewheeling of the cassette, one could look at the consequences of this at level of the drivetrain,[49] and after successfully understanding it the component parts, interaction and organization of the system of the drivetrain we could go to a higher level still, the organization of the parts and their interactions of the complete bike. Thus, we have located the sufficient conditions for one interesting function of the bike. This doesn't mean that we can build a complete bike; we have to locate the sufficient conditions for all the relevant functions. This example is meant to illustrate the collaboration between top-down (locating a function of the bike assembled completely – the highest level of organization) *and* bottom-up (locating the parts, interactions and their organization – starting at the lowest relevant level of organization). Dennett (1994) calls this reverse engineering,

---

[48] This is the function of the hub that enables the peculiar function, i.e., the fact that when the wheel is moving forward, the cassette is not moving forward with it, which (at a higher level of organization) would make the pedals move forward relative to the wheel as well.

[49] The drivetrain is the system that enables the transmission of power from the pedals through the axel which enables forward momentum.

and argues for the applicability of it in studying human behavior. Wimsatt (1972) argues that decomposition and localization (and consequently bottom-up and top-down methodology) is how biologists analyze living systems. Mendelian genetics and molecular genetics is a relevant example, where Mendelian genetics is top-down, while molecular genetics is (primarily) bottom-up. This example also illustrates the explanatory autonomy of the different scientific disciplines (see fn. 41).[50]

To sum up, the methodological reaction to the global logical empiricist approach to philosophy of science can be illustrated by three important and interrelated methodological "turns". First *pluralism*, that is, especially explanatory and methodological pluralism (though presumably other kinds of pluralisms encouraged by scientific practice would be welcome as well). Second, an *appeal to scientific practice* as a guide in solving problems in the philosophy of science. Third, the autonomy of the different approaches especially explanatory autonomy of the special sciences and different disciplines. These are the essential features of what I have called the *local* approach to the philosophy of science. As the next section will participate in illustrating, I believe this is common trend in the philosophy of science, not only in the systems tradition. However, I think it is natural to place the mechanists as a continuation of this tradition.

§3.5 The concept of level and alternative approaches to hierarchical organization

As we saw in §3.4, the notion of hierarchical organization essential in describing complex system. However, there are many other approaches to the notion of level in the philosophy of science that do not figure in the systems approach. We look at one of these proposals, levels of processes (Churchland and Sejnowski 1994). However, I star with going through some the intuitions usually involved when thinking about hierarchies and then describe the structure of compositional hierarchies.[51] After this I turn my attention to the global approaches to hierarchical organization, the two main contenders for the global approach are the British emergentists (Alexander 1920, Broad 1925 & Mill 1843) and the logical empiricists

---

[50] In his (Wimsatt 2007) book, Wimsatt has collected most of his work since the 1970s until today. A substantial part of his book is to argue that philosophers of science should pay attention to how scientist deal with complexity, and use the similar heuristic techniques in dealing with the problems that confront us as *philosophers* of science. He argues that we should not try to construct a global framework for all of the problems to be subsumed under, but rather we should embrace pluralism, as the scientists, especially biologists, have done before us.

[51] Salthe (2012) distinguished between two different logical structures of hierarchies, the compositional (which we will discuss here) and the subsumptive. The prime example of a subsumption hierarchy is the Linnaean taxonomic hierarchy invoked to describe the relation between organisms in biology. I will, however, not discuss the subsumptive hierarchies. A pictorial example, however, can be seen in *figure 2* below illustrating Craver (2007) taxonomy of levels.

(Oppenheim & Putnam 1948). There are, as we shall see several problems with a global approach to hierarchical organization. I end the discussion with levels of processes (Chruchland and Sejnowski 1994).

§ 3.5.1 Intuitions on and the abstract structure of hierarchies

Differentiating phenomena or reality into *levels* entail a hierarchy, a top-level and a bottom-level[52] and entities at the different places (i.e. the levels themselves) in the hierarchy. There are many other features that are often presumed when one talks about levels in science, e.g., that lower-level entities are smaller in size than higher-level entities, that different levels are investigated by different techniques, that different scientific fields and disciplines are concerned with different levels, and that the arrangement of levels are stratified and monolithic. It is also often natural organize reality into global hierarchy, i.e. a hierarchy for the whole of reality.[53] Global conception of levels has been suggested by the British emergentists (e.g. Alexander 1920, Broad 1925 and Mill 1843), and (some) of the logical empiricists (Oppenheim & Putnam 1958) in their discussion of the unity and global micro-reduction of science, and the stratified structure of the world. That is, the notion of levels and global hierarchical organization (especially in the logical empiricists) is it use as a characterization of what the structure of science and its relation to the whole of reality looks or could, in principle, given the unity of science

The reason behind invoking levels in explaining phenomena in biology in the special sciences is often that talk of levels can serve a role as a heuristic, which can aid our understanding (and more satisfactorily explain) specific phenomena.

Just as there are different theories of scientific explanation (e.g. causal, mechanistic, deductive-nomological, evolutionary, etc.) there are different conceptions (and theories) of levels. As Craver points out:

> […] the term “level” has several common uses in contemporary neuroscience. To name a few, there are levels of abstraction, analysis, behavior, complexity, description, explanation, function, generality, organization, science, and theory. Consequently, scientific and philosophical disputes about levels cannot be addressed, let alone resolved, without first sorting out which of the various senses of “level” is under discussion. (Craver 2007:163–64)

---

[52] Or at least a topping-off and bottoming-out level, i.e. the top and bottom levels do not need to be anything more than the highest and lowest level in the relevant domain one studies.

[53] This list about intuitions about hierarchies and levels is not meant to be exhaustive.

This list (and this is only in neuroscience!) shows us that we are dealing with a concept that is frequently used, but seldom defined properly.

As have been indicated above, the concept of levels is fairly common in science and philosophy. But how well defined the concept is and how it is used varies greatly. If it is just assumed that we know what the *relata* (i.e. the things, entities, objects, activities, etc. that are located at the different levels in a hierarchy) in the hierarchy are, or that any notion of level captures the intended use, misunderstandings arise. We should therefore make explcit the intended objects or domain we wish to describe or organize hierarchically and in virtue of what the items in the hierarchy are at the same or different levels.

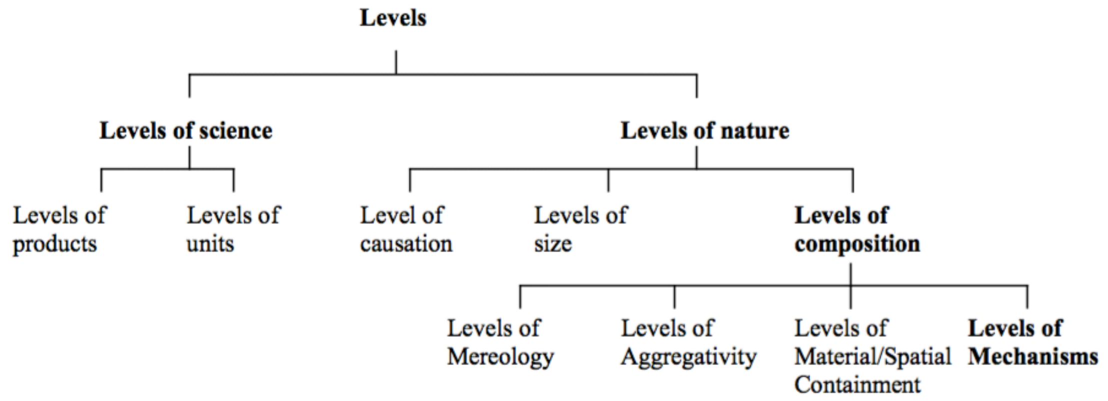

Figure 1: Craver's taxonomy of levels. Here we see that the levels in question are given exceedingly more detailed and explicit inter-level definition (e.g. levels *of aggregativity*, *mereology*, etc.) the further down the taxonomic hierarchy we get. Craver (2007: 171)

Craver (2007: 171–172) proposes three defining questions that help differentiate the different notions of levels into a taxonomic hierarchy.

> (1) What are the relata, or what kinds of entities are placed in a hierarchy of levels?[54]
> (2) What is the inter-level relation, i.e. as a result of what are to items placed at different levels?
> (3) As a result of what are two items placed at the same level?

---

[54] This has been mentioned earlier, since it is perhaps the most general question to answer when one is talking about levels. I include it again for the sake of clarity.

Through answering these questions we can place different notions of levels at difference places in the taxonomic hierarchy (See Figure 1). The first question is supposed to distinguish the first class,[55] where we have an epistemic versus an ontic distinction. Levels of science are defined "by reference to divisions of science rather than by reference to divisions in the structure of the world." (Ibid: 172). The possible relata here are scientific units. The right side of the taxonomic hierarchy (figure 1) is concerned with levels of nature. These levels should be thought of as primarily ontic structures of the world, where scientific units and products are at best approximations (but more generally derivate upon) of these structures. The possible relata for levels of nature are "… entities, activities, properties, and mechanism." (Craver 2007: 177) and distinguishes three different types of interlevel relations, causation, size and composition.

§ 3.5.2 Levels of science and problems with global hierarchies

The possible relata in levels of science are *units*, (e.g. fields, research programs, paradigms, etc.[56]), and *products* (e.g. descriptions, explanations, theories). It is, however, not the case that the scientific products or units are (necessarily) solely epistemic. They may correspond to each other and to divisions in the structure of the world. Oppenheim and Putnam (1958), use levels in this sense.[57] They believe in a stratified and monolithic structure of the world *and* of science, and that these two hierarchies correspond neatly.[58] Each level has its own units such (scientific disciplines in some of the British emergentists; physics, chemistry, biology and psychology, and postulated entities of science in Oppenheim and Putnam; atoms, molecules, cells, multicellular things and social groups) and is governed by the same laws, ceteris paribus laws or generalizations.

The British emergentist proposed such a use of levels of science and nature to account for emergent phenomena or properties, especially in chemistry and biology, and the laws governing them (McLaughlin 1992). Oppenheim and Putnam (1958) use levels to account for

---

[55] The terms used to denote the different places in the taxonomic hierarchy are not to be found in Craver (2007). However, I have chosen to use such term for clarifying the descriptive part of the discussion on the taxonomy of levels.

[56] See, Darden (1991) and Craver and Darden (2013), Lakatos (1977), Kuhn (1962), respectively.

[57] This distinguishes Oppenheim & Putnam's (1958) view on reduction from Nagel (1961: ch. 11). For Nagel the reduction (and consequently hierarchical structure of different) scientific theories is epistemic, while for Oppenheim & Putnam it is epistemic *and* ontic.

[58] It is not clear whether or not the units and its extensions are supposed to be isomorphic (in an ideal explanation). However, if this was supposed to be the case it seems that one needs a 1:1 model (i.e. a qualitatively identical model) is needed. This is obviously a fundamentally misguided view of what explanations are supposed to do.

global micro-reduction and showing the unity of science through the (projected) reduction of higher-level scientific explanations into explanations by lower-level entities and laws.[59] The British emergentists' need for emergent laws became unnecessary, in the field of physics at least, by the reduction of classical thermodynamics to statistical thermodynamics (McLaughlin 1992), and Oppenheim and Putnam's notion of global micro-reduction (especially as *the* means for showing the unity of science) has been largely abandoned in contemporary debates on reductionism (see §3.4 and Sterelny & Griffiths 1999: ch. 7). The main reason that Oppenheim and Putnam's stratified and monolithic view of science has become outdated is that the domain of each science (e.g. biology, sociology, etc.) is no longer believed to be restricted to its own level. There seems to be no (at least *prima facie*) good reason to suppose that all of science is hierarchically organized, stratified and working at different isolated levels. Craver and Darden (2013) and Schaffner (1993) argue (although with a different conception of levels) most biological and biomedical explanations span multiple levels and fields, and to think that the smallest or lowest level[60] is privileged and able to account for every higher-level phenomena only with its own explanatory tools (i.e. the laws and/ or generalizations governing the lowest level, including bridging laws) does not reflect current scientific enterprise. Integration of explanations from different fields in biology (and psychology), such as the (mechanistic) explanation of spatial memory in rats (see §4.2.2 and Craver 2007: ch. 5) shows us that even though phenomena cannot be reduced (in the logical empiricist sense), the explanations at different levels can be integrated.

There is a further problem for the stratified and global view of the hierarchical organization science and the structure of the world. On such a view the products of science and the things in the world must *all* fit into the same supposed hierarchy (of scientific products). Where in the hierarchy should we place stable units composed of the entities postulated by science, such as a network of cells, organs, etc.? The distinction between products and units also brings out a problem. We generally associate the study of our solar system with physics. Does this place the solar system on a lower level than the study of say biological species? If it does, it seems weird that all known biological species are included in (or a part of) something at a lower level than itself.

---

[59] See fn. 57.

[60] Including the field(s) or science(s) that studies the lowest, most fundamental level.

§ 3.3 Local hierarchies

Some philosophers have proposed causation as defining the inter-level relation, perhaps most notably Churchland and Sejnowski (1992). They propose that the causal relations between sequential processes could be differentiated into levels according to their number in the causal sequence from low to high. Such levels of processes make a lot of sense with explanations in (computational) neuroscience and cognitive science. As such, they exemplify a *local* approach to the philosophy, that outside of the systems tradition. For example:

> On this measure, cells in the primary visual area of the neocortex that respond to oriented bars of light are at a higher level than cells in the lateral geniculate nucleus, which in turn are at a higher level than retinal ganglion cells. (Churchland and Sejnowski 1992: 23)

Using levels of processes may work in closed systems, i.e. systems where the causal relations after time t, (say light hits the retinal ganglion cells), are internal to that system. In most biological explanations however, this would be an awkward and unpractical way of distinguishing different phenomena (i.e. processes, activities, entities, mechanisms, etc.) into levels. Think of, for example, all the processes that occur in the formation of two species from one. A geographic isolating barrier splits a population of the same species into two. This event may take a couple of years of, say, bad climate, while the establishment of a gametic isolating barrier[61] may take several generations. However, it would be unnatural to say that the gametic isolating barrier, which relates to all (or most of) the individual organisms in a population, is a higher level than the geographic isolating barrier because it occurred at a later stage.[62] Many explanations in biology, especially evolutionary ones, have several different phenomena with temporal overlapping and are happing at different timescales.

However, Churchland and Sejnowski invokes the hierarchical organization of processes in computational neuroscience, and do not intend it to be applicable anywhere else. Therefore, the counter-example is *moot*; it is an example outside of the intended domain of the hierarchy. A counter-example should come from problems within the right domain. For example, *feedback loops* from higher-level processes to lower-level processes could be an

---

[61] A gametic isolating barrier is a reproductive isolating barrier that inhibits sexual reproduction by making fertilization between two individuals (of what used to be the same species) impossible or making the offspring sterile.

[62] It is not even clear if the processes can be said to happen sequentially, perhaps the two distinct populations where forced to move east and west respectively, and during the years it took to properly make the populations properly allopatric (completely isolated from each other) the populations started diverging in their genetic make-up, making them unable to interbreed. If that is the case, then the two processes were partly overlapping.

obstacle. So if one could construct a counter-example by an appeal to non-sequential or sporadic temporal organization of relevant processes in the brain, it would be difficult to organize the processes into a hierarchy.[63] This is an example of the meta-philosophical claim. We have to be aware of the relevant domain or context the concept is invoked in order to construct useful counter examples. However, when it is unclear what domain or scope the relevant concept is invoked for, misunderstandings arise. So, just as it is the one who attempts to criticize responsibility to restrict his criticism to the relevant domain, it is equally responsibility of the one who invoked the concept to *explicitly* state the intended domain. This is a consequence of the local approach to philosophy of science, and it makes the use of generalized concepts that share their meaning (e.g. mechanism) in sore need of explicit scope-restrictions. When we return to this in §4.x we shall see that some of the mechanists are vague in the explication, something that has provoked philosophers outside of the mechanistic philosophy.

§3.6 Summing up – escape from logical empiricism?

Summing up then, we saw in §2, the logical empiricists global approach to scientific explanation, the *DN*-model of Hempel (Hempel & Oppenheim 1948, Hempel 1965). The main reaction to this was the account of causal explanation as an alternative model of explanation that could adequately deal with the problems of explanatory irrelevance and asymmetry. Second, we looked at the logical empiricists' model for reduction, again the logical demands opened for counter-examples challenging the validity of reduction (i.e. multiple realization). An alternative account of reduction was proposed by the systems tradition, shifting the focus away from the epistemological entities of science and prioritized to the ontological entities that scientist actually study. The discussion on systems approaches reduction and inter-level show us the edifice and advantages of the local approach to the philosophy of science. Lastly, we saw that hierarchical organization is common in science, though the conditions for the organization is seldom spelled out in detail and left implicit. We saw that a global approach is faced with serious difficulties and that it does not conform well to the state of science as it is now or in a foreseeable future. We also saw, through a discussion of an approach to hierarchical organization outside the systems tradition, that a consequence of the local approach to the philosophy of science is the need for awareness of explicit formulation intended scope when invoking concepts or formulating theories. This

---

[63] On such a case it would seem that the hierarchy would consist of a myriad of processes interacting across levels following no particular rules.

restricts the criticism to relevant domain, while it also makes more it difficult to generalize across the science without misunderstanding.

My historical claim is that the mechanists can be seen as a continuation of the system approach, with Salmon's causal-mechanical model as an extension. So a few words on what I take to be the most important contribution from the systems approach and Salmon to the mechanistic philosophy is in order. As we saw in §2.2.3, Salmon has influenced all of the main characters in the mechanistic philosophy, although, perhaps somewhat differently. For Bechtel (2006: 25) Salmon's main contribution was to extend the domain of mechanistic philosophy to escape the clutches of the Newtonian and Cartesian picture of atomistic mechanisms, *just* concerned with the shape and motion of physical particles. However, since Salmon was primarily concerned with physics and left the definition of "mechanism" largely untouched Bechtel (2006) can be seen as providing a definition that extends the scope of mechanistic philosophy to include biological mechanism, especially cell mechanisms. Craver (2007), Glennan (1996, 2002) and Machamer, et al. (2000) also provide a definition of mechanism, but Salmon's influence on them is more extensive. Craver (2007: ch. 4) is heavily influenced by Salmon's notion of constitutive explanation, and he develops a theory of constitutive explanation where the "underlying mechanisms" are defined. This is perhaps the most common approach to mechanistic explanation (that is, constitutive and *not* etiological). Craver (2007: 107–108) points out that mechanistic explanation sometimes are etiological, for example in the explanation of thirst, where dehydration is the underlying mechanism. Glennan (1996) is more influenced by the metaphysical aspects of Salmon. Glennan uses the concept of mechanism to account for the nature of causal interaction. Thus he can be seen as providing an alternative account to causation by utilizing concepts invoked by Salmon. These different influences can also be seen in the diverse motivations and scope that these authors have for mechanistic philosophy

The systems approach has had more similar influence on the authors, as this is what characterizes the common aspects of mechanistic philosophy. They all share the view that mechanisms are complex systems, decomposition and localization is how one individuates the relevant parts, interactions and the organization of the components of a mechanism. They all accept organizational emergence, explanatory and methodological pluralism. Further, they all appeal, to a great extent, to scientific practice as providing criteria for adequate explanation. However, different philosophers in the systems tradition have had different influence on the mechanists. Craver (2007: 109–110), for example, in his account of how to formulate constitutive explanations is greatly influenced by Cummins (1975) *dispositional* theory of functional explanation. Glennan (1996: 56) too explicitly acknowledges the influence Cummins (1975) has had on his decompositional strategy (including others authors of the systems tradition. Bechtel (2006) and particularly Bechtel & Richardson (2010 [1993]: 24–27) are more influenced by Simon (1962, 1996 [1969]) and Wimsatt (1972) in their explication of decomposition and localization. However, the different authors I have appealed to in §3.4 are all included in Craver (2007), Bechtel (2007) and Bechtel & Richardson (2010 [1993]).[64]

In the next chapter we look at the overarching mechanistic philosophy, and its differences. To a great extent, it is continuation of the systems tradition, just with more explcit focus on causality.

---

[64] Most are included in Glennan (1996, 2002) as well, but not all.

§4. The New Mechanistic Philosophy

> At least in biology, most scientists see their work as explaining types of phenomena by discovering mechanisms, rather than explaining theories by deriving them from or reducing them to other theories, and *this* is seen by them as reduction, or as integrally tied to it.
>
> Wimsatt (1976a: 671)

§4.1 Introduction

In this section we shall look in detail at the framework of mechanistic philosophy, especially mechanistic explanation as this is at the heart of the mechanistic framework. There is general agreement on the overarching framework, most of the surface disagreement come from terminological choices and is not very serious. But as we shall see in §4.3, there are some substantial disagreements as well. After I have gone through the mechanistic framework and the disagreements and provided a partial analysis of why there is disagreement (§4.3). This discussion leads us to my to evaluate my meta-philosophical claim in light of the historical account, but especially concerning the mechanistic framework. I argue that in paying attention to the peculiarities of the local approach to the philosophy of science we are faced with other difficulties in evaluating and criticizing the mechanistic framework, than as with the global logical empiricist framework. There are consequences by adopting a local approach to the philosophy of science. I end by sketching how some of these consequences can be dealt with.

§4.2 The mechanistic approach to explanation

Let's turn our attention to the mechanistic approach to explanation. Machamer, Darden & Craver (2000)[65] was the most cited paper in the journal *Philosophy of Science* between 2007 and 2010, and it offers one of the clearest formulations of the mechanistic approach to explanation. I will, however, also turn to Craver (2007, 2015) when I encounter what is intentionally left out of the paper or unsatisfactorily discussed. The overarching differences in the three mechanistic "camps"[66] will be discussed in the §4.3, while minor differences will

---

[65] I will use the abbreviation (MDC) to refer to 'Machamer, Darden & Craver (2000)', for ease of presentation.

[66] Following Levy (2013), the three camps are (1) (MDC) & Craver (2007), (2) Glennan (1996), & (3) Bechtel (2006), Bechtel & Abrahamsen (2005), Bechtel & Richardson (2010 [1993]). This list is not exhaustive for the authors work on the topic, but is meant to pick out the article and books that are most characteristic of the three camps.

only be pointed out if they are relevant for the discussion. As Franklin-Hall (forthcoming) points out, many of the differences are linguistic disagreements, e.g., the intuitions sparked by and the connotations associated with particular terms employed in the explication of mechanisms, so these will receive little attention.

Let's start with the definition of a mechanism given by (MDC):

> 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, Darden & Craver 2000: 3)

To explain a phenomenon mechanistically is to give a description of the relevant entities and activities, and how they are organized such that they regularly produce a phenomenon or a part of a phenomenon. In relation to explanation, then, it picks out the *explanatory relevant* facts for explanation. The explanatory relevant parts of an explanation of a phenomenon (P) consist of (1) *entities*; i.e. the relevant constitutive parts (2) *activities*; i.e., the relevant causal relationships obtaining between the entities, and (3), the *organization* of the relevant entities and activities (MDC: 3). In the language of the logical empiricists we can say that the *explanandum* is (P), while the *explanans* are (1–3). The notion of *production*, *regularity*, and *start* or *set-up* and *finish* or *termination conditions* are the condition and effect of the specific organization of t the *activities*, *entities* that operate within in the mechanism (MDC: 3).

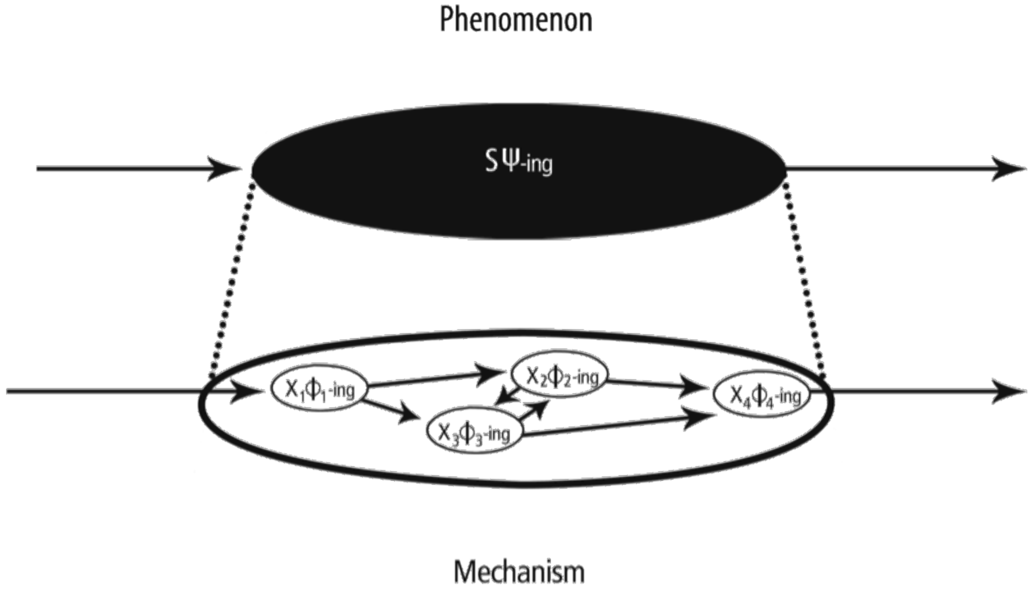

Figure 2: The phenomenon is characterized by a system *S*, preforming an activity or function ψ-*ing*. The mechanism is characterized by the entities X_{1-n}, which are engaged in activities ϕ_{1-n}-*ing*. The dotted lines represent the fact that the mechanism *produces* the behavior of the system *S*. (Craver: 2015: §2)

The activities and entities are the two ontologically fundamental parts of the mechanism, and they are interrelated. (MDC) writes:

> Mechanisms are composed of both *entities* (with their properties) and *activities*. Activities are the producers of change. Entities are the things that engage in activities. Activities usually require that entities have specific properties. (MDC: 3)

The reason behind this dual-ontology is to capture what (MDC) take to be important insight from *both* substantivalist- and process-ontologies. The substantivalist focuses on the disposition or capacities of entities or properties, while the process ontologist focuses on reducing entities to processes (in reifying the activities of these entities) (MDC: 4–5). (MDC) suggest that both perspectives complement each other. A good way to understand what *dispositions* or *capacities* an entity has is by looking at the activities it engages in. And oppositely, a good way to understand why an activity occurs is by looking at the properties (i.e. the disposition or capacities) that enable the entities to engage in process (cf. §3.3.2).[67] "As far as we know, there are no activities in neurobiology and molecular biology that are not activities *of* entities" (MDC: 5). Thus, the ultimate justification for this dual-ontology is by scientific authority.

The *organization* of the activities and entities in a mechanism is what *determines* the how the mechanism operates, i.e., it determines "the ways in which they [activities and entities] produce the phenomenon. The entities are organized in a spatial manner, while the activities are organized in a temporal manner. The entities must be spatially organized in such a way that they can engage in activities with each other, while the activities must be temporally organized in such a way that they ensure the behavior of the entities and ultimately the mechanism (e.g. the temporal order and duration must be right for the entities to be capable of engaging in the relevant activities). The notion of *regularity* is explained by the predictability of the behavior of the mechanism under the same or similar circumstances. So if the activities and entities organized in a particular way, a mechanism should, go from *set-up* to *finish* conditions without the lack of *productive continuity*. Productive continuity is

---

[67] This is a crude and oversimplified presentation of the substantivalist and process approaches to ontology. In §3.4.2, I argue that it is the insight from substantivalist ontology and dispositional approaches to functional explanation (in general the focus on entities, their organization and properties in the system approach), which distinguishes the mechanist approach to explanation from a purely causal explanation. I also argue that such focus on entities is what makes the mechanistic approach so appropriate for explanation in the proximal-physiological parts of biology.

the seamless and complete behavior of the entities and activities that produce the phenomenon (MDC: 3).[68]

The phenomenon is the *behavior* of the mechanism seen at the level of the *system*, i.e., of all the activities and entities and their organization throughout the hierarchical structure of the mechanism and its components. Components, relative to a mechanism, are just a shorthand term for describing (1–3). If we were to look at the composition (i.e. the organization of the activities and entities) of one of these components, we would refer to them as sub-components relative to the mechanism, but if one looks at a component in isolation, we would refer to the component as the mechanism, and the sub-component as the components (see §4.2.1 for an example). Thus, the boundaries of a mechanism are context-relative; they are relative to the mechanism that produces the phenomenon.

The boundaries of the mechanism pick out the relevant *activities* and *entities* that are organized such that the mechanism *produces* the phenomena. Further, to understand what entities and activities (and what organization) that produce the phenomena one need to *decompose* the mechanism into hierarchies. This is process of picking out the *explanatory relevant* parts (entities and activities). In other words, in *decomposition* we individuate the entities and activities of the mechanism. The next step is *localizing* the kinds of entities and the relevant subset of properties for the entities to engage in the activities and their spatial location, and kinds of activities and their temporal location. The specific organization of these relevant activities and entities produce the phenomenon. Decomposition and localization are most commonly referred to in the *discovery* of mechanism (Bechtel & Richardson 2010 [1993]: 23–27), but are essential for the explanation to cite *relevant* facts. It should also be noted that in cases of homogenous reductions, as mentioned in §3.1, the decomposition and localization is "trivial", that is, it reveals little new information about the novelty that appears in virtue of the organization of entities and/or activities. This is why the mechanists are concerned with complex systems, similar to what was discussed in §3.4 (Levy 2013, Glennan 2002).

---

[68] Many have interpreted the regularity condition in the definition of a mechanism in (MDC: 3) as a condition of recurrence or repetitive behavior of mechanism; this is not how it is meant by (MDC). It is enough with predictability (see Craver 2015: §2.1.2).

§4.2.1 The hierarchical organization of mechanisms

As we saw above, a mechanism has components which themselves have sub-components (and the component can be seen as a mechanism with its sub-components as its components). The different levels of mechanisms are therefore levels of composition (see figure 1). However, what distinguishes it from other levels of composition is the *relata*. The mechanisms are at a higher level than its components, and the components are at a higher level than the sub-components, and so on. The inter-level relationship is best characterized as follows: "*X*'s ϕ-ing is at a lower mechanistic level than *S*'s ψ-ing if and only if *X*'s ϕ-ing is a component in the mechanism for *S*'s ψ-ing." (Craver 2007: 189). This can be represented as follows:

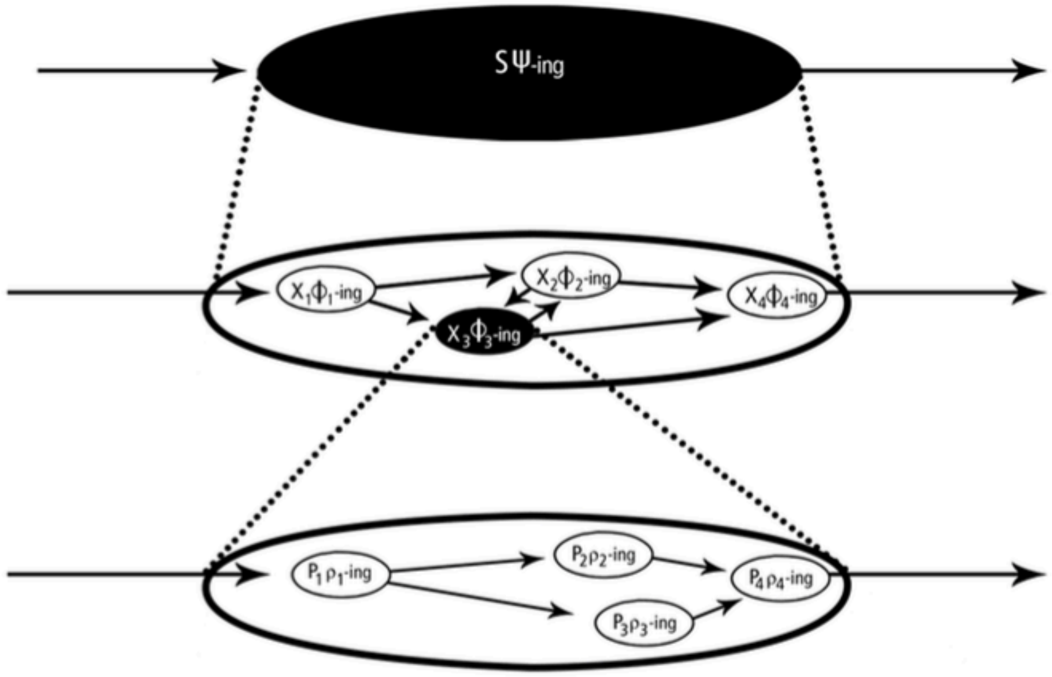

Figure 3: An abstract representation of a mechanistic hierarchy. The highest level, Sψ-ing, is the phenomena. The middle level, $X_{1\text{-}n}\phi_{n\text{-}1}$-ing, describes the components of the mechanism. The lowest level, $P_{1\text{-}n}\rho_{1\text{-}n}$-ing, describes one set of sub-components. (Craver 2007: 189)

The notion of organization play a vital role as well, the lower-level components make up, through their organization, the higher-level components in mechanism (and so the phenomenon). Craver's (2007: 165–70) example is how the explanation of spatial memory in a rat extends many levels of mechanisms. The rat experiences spatial memory (the level of spatial memory, i.e. the phenomenon). This phenomenon is produced by computational properties (i.e. the mechanisms) of neural systems, such as the hippocampus (and other areas in the temporal and frontal cortex) (the level of spatial map formation). The acting entities of the hippocampus are (conjectured to be) long-term potentiation of hippocampal synapses (the cellular-electrophysiological level). The acting entities of the hippocampal synapses are said to be the molecular mechanism that make the chemical and electrical activities at the cellular-electrophysiological level possible. Thus the hierarchy of mechanistic levels in this particular explanation is established.

It should be obvious that levels of mechanisms are local and non-monolithic, and what this entail is that the molecular mechanisms (such as Mg^{2+} ions, Ca^{2+} ions and NDMA receptor) working in a different part of the brain, say the primary motor cortex, are not at the same level as the molecular mechanisms in the hippocampus or the temporal cortex. The relation in the hierarchy of acting entities is based *only* on their relation to each other as parts of the different mechanisms underlying, producing or maintaining a phenomenon.

§4.3 Salient differences between mechanistic approaches

There is agreement between all mechanists that: (a) the phenomenon being explained is the behavior or function of a system, and that a characterization of the mechanism that underlies the phenomenon is *explanatory adequate* for that purpose. (b) Mechanisms are composed of entities with certain dispositional properties or capacities. (c) These entities engage in activities, and (d) the entities and activities are organized in such a way that they produce the phenomenon, i.e. the behavior of the system.

As we noted above, the terminological disagreement are abounds. However, these are "merely" linguistic, and should be seen as relative to the context of explanation. For example, in cases where there is a causal sequence that ends up with a product, say the abnormal cell growth that causes a benign tumor to appear, it gives the right connotation to say that the mechanism *produces* the tumor. While other cases it may be more appropriate to use *underlie*, for example when explaining physiological mechanisms, as the mechanism are responsible for a higher-level function, e.g., eyesight. Lastly, in some cases it might be more appropriate to say that a mechanism *maintains* a phenomenon, for example when we talk about the homeostatic mechanism that holds our body temperature at roughly 37°C (Craver 2015: §2.1.1). These different cases each share the idea that the mechanism should be characterized by (a)–(d), but a shift in linguistic presentation may yield an intuitively easier understanding of the operation of the mechanism. Similarly Bechtel (2006) refers to the *operation* amongst parts in a system, instead of activities of entities in (MDC).

On this Bechtel notes:

> [MDC] employ the term *activity* to draw attention to the fact that components of mechanisms are active. … The term *activity* however, however, does not readily capture the fact that in most operations there is something acted upon. This is the reason I have preferred the term *operation*. Typical of the operations I have in mind are the reactions of chemistry which prototypically involve a catalyst, a reactant, a product, and often a cofactor. In some reactions there is no need for a separate catalyst as the energetic factor are such that the reaction will occur spontaneously. In autocatalytic reactions, which are highly relevant in living systems, the product of the reaction also serves as a condition for more iterations of the reaction. (Bechtel 2006: 30, fn. 7)

So in chemistry, for instance, the notion of activity inadequately captures the *workings* of the entities. However, this is not a substantial issue, and can be resolved by being aware of the typical workings in the relevant system. The appeal to scientific practice, as we have seen is one of the main motivations behind the mechanistic approach, will naturally employ a pluralistic vocabulary to capture the right connotations for the properties of the system. So, having scientific practice as a guiding authority, the mechanist can say that different scientific disciplines or fields will employ different vocabularies, but the abstract structure (i.e. constitution, causation and organization) is still sufficiently similar.

However, there are certain things the mechanists do not agree on which are of substantial concern. The three most salient areas of disagreement, in my opinion, concern the notion of *activity* and *regularity* and the *ontological status* of mechanistic explanation.

§4.3.1 The nature of activities

Until now, the nature of causation and causal terms has largely gone unexamined. The explanatory adequacy of mechanistic explanation does not hinge on what causality really *is*. The mechanistic approach is compatible with a plurality of causal ontologies.[69] However, something has to be said about causation in relation to explanation; we need to know how or

---

[69] I haven't seen a detailed examination of what consequences different ontological views of causation could have on mechanistic explanation. If, for example, some kind of process ontology turns out to be true, I suspect the mechanists (or at least Bechtel) would argue that the *representation* of mechanisms as activities and entities is *epistemically* preferable to a process approach to explanation. See Dupré (2013) for an attack on mechanistic understanding of proximal-biology from a process perspective.

why something is a cause (or effect) and other things are not. Recall that Salmon's (1984) view of causation discussed in §3.2ebma, distinguished causes from non-causes by appeal to mark transmission. However, causation by mark transmission is problematic in the special sciences. Biologists frequently appeal to causal disconnection (e.g. by omission, prevention, or double prevention) as explanatory relevant, and it is not easy to see how these could be causal on Salmon's theory of causation. Even though it has been a source of inspiration for most of the mechanists, as an account of the nature of activities it has largely been abandoned (Craver 2015: §2.3.1). Roughly, three different views of the nature of causation or causal claims are popular amongst the mechanists, the view that causal claims are derivate of the concept of mechanism, that activities are *kinds of causing*, and any attempt to define causality would exclude some kinds of causing, and a counterfactual view of causal relevance.

Glennan (1996) argues for a view that a relationship between to events (variables) is causal in virtue of a *connecting* mechanism.

> A mechanism underlying a behavior is a complex system which *produces* that behavior by the interaction of a number of parts according to *direct causal laws*. (Glennan 1996: 52, italics added)

Glennan offers the concept of mechanism as what underlies or produces the causal relationship. The underlying mechanism is thought to connect the cause with the effect, thus describing Hume's *secret connexion*. The mechanism itself has underlying mechanisms (components) that connect its cause to its effect, and it goes on until it bottoms out in *non-causal* underlying relationships in fundamental physics. The notion of mechanism provided by Glennan (1996) can be seen as metaphysical in nature. Its primary aim is to account for the nature of causation, as opposed to (MDC) where the notion of mechanism primarily has an explanatory or epistemological job.[70]

This view of causation is *acutalist*, *singular* and *reductive*. By acutalist I mean that it takes causation to be productive, that the causal relationship is dependent upon an actual underlying relation. This can be contrasted with a *counterfactual* dependence relation, where the causal relationship is causal if there would be a change in the effect if something different had happened to the cause. By singular (or token) I mean the view that particular or token causal relationships are what make general (or type) causal claims true. For example, (S_S) "Blondie's smoking caused him to develop lung cancer", is a particular instance, i.e. a

---

[70] (MDC) argues for the ontic adequacy of a mechanism, primarily by appealing to scientific practice as authoritative justification (as we saw in §4.2), but the concept of mechanism, on their view, is not supposed to say anything about the nature of causal relationship.

singular causal relationship.[71] While (S_G) “Smoking causes lung cancer”, describes a general causal relationship. When sufficiently many particular instances of “*X*’s smoking caused *X* to develop lung cancer” are found, we are able to generalize to (S_G).

By reductive I mean that is an effort to define what causal relationship *really is*, and not just how they behave or what we can expect from them. This is contrasted with a *non-reductive* view of causation, where the nature of causal relationships is, at best, indirectly described by the different types of causation we can describe. The mechanism view of causation has been charged with circularity, since causal terms (e.g. produce) figures in the definition of the nature of causation. Woodward (2004) argues that a production-based view of causation insufficiently captures the causal notions and terms in science. Many causal relationships in science are “cashed out” by appeal to counterfactual claims about difference making.

The view of activities in mechanism as *kinds of causing* is singular as well. Bogen (2005) and Machamer (2004) argue for such an account. The essential idea is that a definition or theory of causality is unnecessary:

> [T]he problem of causes is not to find a general and adequate ontological or stipulative definition, but a problem of finding out, in any given case, what are the possible, plausible, and actual causes at work in a given mechanism. (Machamer 2004: 27–28)

Following Anscombe, Machamer takes ‘cause’ to be genus term, like ‘organism’. Scientist doesn’t look for a theory of organisms *per se*, but look at different *kinds* of organisms, and in much the same way we should look for kinds of *causing* and not a theory of causation. So the kind of causing will be relative to the context, and in relation to mechanisms, the kinds of causing are activities.[72] Activities are the happenings that produce change, singularly or in concert with other activities.

This view of causation is explicitly non-reductive. However, whether or not it is singular or general, and acutalist or counterfactualist is not entirely clear. It seems to me that intuitively, one would characterize such a view as singular. Since the kinds of causing are relative to a particular context it seems natural that general causal claims consists of a group of sufficiently similar kinds of causing for the generalization to be useful. Further, it seems,

---

[71] At least at the relevant level of description, if we look at a lower level we would find a plurality of singular causes.

[72] Machamer, being one of the authors of (MDC), uses the same terminology as I have used in my explication of mechanistic explanation.

how well one describe the *kind* of causing will point towards a singular view on causation. That is, if you have a meticulous description of two causal relationship that are similar on a less detailed level of description, one would find differences, and if just one difference makes the causal relationship a different kind, then it seems that this would lead to a singular view on causation. I think Machamer (2004) and Bogen (2005) have a singular view of causation, as they argue that one should drop a regularity condition on mechanisms, because some mechanisms are characteristically *non-regular*. Whether or not it is acutalist or counterfactualist is also arguable. It seems that in the quote above Machamer doesn't want to exclude either, nothing that in looking for causes we are looking for "the *possible*, plausible, and *actual* causes at work" (Machamer 2004: 27). It sounds right that to view causation as *kinds of causing* goes hand in hand with a view that there are two different, equally valid, modal conceptions of causation.[73]

The counterfactual view of causation is often associated with Lewis (1973) where, roughly, we describe what it takes for something *C* to be the cause of *E* is a counterfactual dependence of *E* on *C*. In other words, if *C* had not occurred, then *E* would not have occurred. The most influential counterfactual theory of causation on the mechanists is due to Woodward (2003). He characterizes causation by an appeal to *intervention*. Two events, E_1 and E_2 are causally related if when one intervenes on one event there is change in the other. So if E_1 is the cause of E_2, then if one intervenes on E_1 there will be a change in E_2, but not the other way around. This view supposedly captures many of our intuitions of causation and conforms to the methods scientists use in testing causal claims by (Craver 2015: §2.3.4).

The view is *non-reductive*; it uses a causal term, i.e. intervention, in order to pick out causal relationship. Much of the same sentiment as we saw in the kinds of causing account of causation motivates the belief that a non-reductive definition of causal relationship is adequate, however, the counterfactualists (especially the interventionists) appeal to scientific practice as an authoritative justification (e.g. Craver 2007: ch. 3 & Woodward 2003: 18–20). Whether or not this view is singular or general is disputed, Woodward (2003: 74-86) seems to not take a particular stance, although his explication of what counts as causal is primarily framed in general- or type-terms. He argues that in some cases we extrapolate from knowledge of general (deterministic) causal claims to singular, and *vice versa*. Because of the non-reductive character of Woodward's theory of causation, it would not, in my opinion, be

---

[73] See Hall (2004) for the view that production-based (or acutalist) and counterfactual concepts of causation are *not* mutually exclusive alternatives, rather but equally valid.

strange to *not* take a stance on whether singular (token) or general (type) causal relationship is fundamental.

In mechanistic explanation, the interventionist theory of causation is thought to be able to pick out the relevant *activities* by (possible or actual) interventions. Craver (2007: ch. 4) defends an interventionist theory of causation in accounting for the *activities* in a mechanism.

Summing up then, the account of what it takes for something to be an activity is, roughly, three-fold between the mechanists. The reason for the different approaches may be due to other motivations, the mechanism approach to causation is motivated metaphysically, as an account of Hume's *secret connexion*, while the counterfactual, especially interventionist, approach is motivated by the conformity to scientific practice. The kinds of causing approach is motivated, in part, by an appeal to the non-regularity of mechanisms, which we will now turn our attention to.

§4.3.2 What does it take for a mechanism to be regular and does it have to?

We saw in §4.2 that (MDC) saw the regularity of a mechanism as a sort of predictability of the behavior of the mechanism under the same or similar conditions. This notion of regularity does not entail the behavior of the mechanism must be repeated on many occasions, however Leuridan (2010) argues that unless a mechanism exhibit what I will call, *repeated regularity*,[74] then the mechanisms cannot provide a genuine alternative to laws of nature as models for (general) explanation. Leuridan thinks that, even though the special sciences has shown us that universal laws generally fail to be applicable to many of the phenomena they study, some kind of nearly exceptionless generalizations, or pragmatic laws[75] are needed for adequate general explanation. The notion of regularity is such an alternative, so a general mechanistic explanation will be adequate if it behaves regularly or nearly without exception. Andersen (2011) rightly argues that regularity, at least in (MDC) is not meant to capture that mechanisms repeatedly occurs, but that mechanisms *explain* such regularities. The notion of regularity in (MDC) is not repeated regularity, but *organizational regularity*. Thus, when a mechanism regularly produces changes it is not that it has to occur multiple times, but a sort of counterfactual notion of regularity. If the entities and activates of a mechanism is

---

[74] That is, the notion of regularity as the repeated occurrence of the behavior of a mechanism.

[75] Leuridan (2010) argues that mechanisms depends on regularities, as a consequence of this he seem to think that mechanisms, by virtue of their dependence on regularities, also depend on (at least) a pragmatic account of laws. See Mitchell (2009: ch. 3 or Mitchell 1997) for the account of pragmatic laws Leuridan (2010) has in mind.

organized as they are, then the mechanism will always or most times produce that behavior. As such, if there is a case where a system behavior occurs *only once*, but it could be predicted by or counterfactually depended on the organization of the relevant activities and entities of the system, then it could appropriately be called a mechanism.

An organizational approach to regularity has been criticized for not being able to account for stochastic mechanisms (Bogen 2005). He argues that there are mechanisms in neuroscience that do not exhibit organizational regularity:

> [Generalization] Electrical activity of a pre-synaptic neuron in one region (the CA3 area) of the hippocampus initiates electrical activity in a post-synaptic neuron in another area (the CA1 area). … [G]eneralizations whose truth that require *every* CA1 neuron in a properly functioning hippocampus to fire under normal circumstances whenever its pre-synaptic CA3 neurons fire are flatly false. (Bogen 2005: 412)

The CA3 neurons has a stochastic influence on the CA1 neurons, post-synaptic responses to the electrical activity is, at first, infrequent.[76] There are even cases where the statistical relevance between set-up conditions and termination conditions are below 0.5 (i.e. where over 50% of the activation of the mechanism at its set-up condition does not end in its described termination condition). Bogen (2005) goes on to argue that regularities are useful *epistemic tools* that generate generalized knowledge of mechanisms, even if later discovered to be stochastic. I will call this *epistemic regularity*. On such a view, a mechanism does not have to exhibit *organizational regularity*, but in order for there to be possible to use regularity as an epistemic tools for generalization, there has at least to be *repeated regularity* at the system-level or the level of the phenomenon being explained. It seems to me that, on this view, where regularities play an epistemic role, but the mechanisms are (or can be) *ontically* irregular, make the conception of mechanistic explanation more epistemic than ontic, at least in the case of generalized mechanistic explanation.[77] Bogen (2005) would argue that in particular mechanistic explanations, (any kind of) regularity condition is unnecessary for something to be a mechanism. Some mechanists find an epistemic approach attractive, allowing the mechanist to be more modest in the characterization of the relation mechanistic explanation has to the structure of the world.

---

[76] However, as the rate of pre-synaptic electrical activity increases the rate of post-synaptic electrical activity increases, “but the process remains stochastic” (Bogen 2005: 412, fn. 20).

[77] For a related discussion on whether or not natural selection can be analyzed stochastically as the main mechanism underlying evolution, see Skipper & Millstein (2005), Barros (2008), and Levy (2012).

§4.3.3 The ontological status of mechanisms

The epistemic/ ontic distinction is, as we have seen, adopted from Salmon (1984), but it will be useful to go through it in more "modern" terminology. Within the mechanistic approach, the epistemic conception is not logical and inferential, as Salmon's characterization was. The epistemic conception views mechanistic explanation as a *representation* of the mechanism.[78] The mechanistic approach is a (epistemically) valuable strategy for understanding complex systems. The mechanists who defend the epistemic approach usually consider the mechanistic approach *primarily* as a research strategy, and secondarily as a model for scientific explanation.[79] Thus, the mechanistic explanation is the *description* of the mechanism.

The ontic conception, on the other hand, argues that the mechanistic explanations "show how phenomena are situated within the causal structure of the world" (Craver 2007: 21). When an explanation of a phenomenon refers to an objective portion of the casual structure of the world, it refers to the set of factors that produce or sustain the phenomenon. These factors are not representation or descriptions, but "they are full-bodied things" (2007: 27). It is the mechanism itself, the set of factors that produce or sustain a phenomenon that explains.

Bechtel argues that there are certain things we should note about the epistemic, contra the ontic, conception that should motivate to be modest about the realistic character of mechanisms:

> First, mechanisms operative in our cells were operative long before the cell biologists discovered and invoked them to explain cellular phenomena. The mechanisms are not themselves explanations; it is the scientist's discovery and rendering of the aspects of the mechanism that produces what counts as an explanation. Second, the difference between the mechanism and the mechanistic explanation is particularly obvious when considering incorrect mechanistic explanation – in such cases the scientist has still appealed to a mechanism, but not one operative in nature. Such a mechanism exists only in the *representation* offered *by* the *scientist*. It is thus the mechanism as *represented*, not the mechanism itself, that figures in explanation. (Bechtel 2006: 34, italics added)

As we can see, Bechtel doesn't exclude the ontological possibility of mechanisms. He just emphasizes the fact that it is a scientist representation of it that figures in a mechanistic explanation. So, ontologically both Bechtel and Craver can accept mechanisms, but the

---

[78]

[79] Further, see Bechtel (2006), Bechtel & Richardson (2010 [1993]) and Bechtel & Abrahamsen (2005).

human element in explanation pushes Bechtel in the epistemic direction, leaving ontological talk relatively untouched. Illari (2013) observes that Bechtel is (almost) a lone wolf in adopting the epistemic position. He further argues that it is likely because of the influence from Salmon (1984) that most of the mechanists adopts the ontic view.[80] Illari (2013) main charge is that the discussion on epistemic and ontic conception should not be conducted as they were two rivals views, but as two different kinds of constraints that each have different impact on the nature and utility of mechanistic explanation. I agree with this, however I will not attempt to contribute to this here. In the next section I look at the main motivation behind the appeal to mechanism by Glennan, Caver and Bechtel, and appeal to some of the areas of disagreement mentioned above. I argue that, following Levy (2013), three different views on the utility of mechanisms can be extracted from the difference in motivation.

§4.3.4 What are mechanisms used for?

What I picked out as the three different "camps" in the new mechanistic philosophy in all have the same overarching view on the concept of mechanism. However, as we saw in §4.3 there are some substantial differences. Building on these differences, I suggest that the different motivations behind the use of the concept of mechanism in these camps, can partially explain them. By explicating the different motivations we get a clearer picture of the different utility the mechanists believe the concept of mechanism has.

In §4.3.1, we touched briefly upon Glennan's (1996) view, he intends to use the concept of mechanism as the *secret connexion* that relates a cause to its effect. Continuing Salmon's (1984) attempt to clarify the nature of causal interactions by providing an explcit account of what the underlying mechanisms are (see §3.2).[81] Thus, Glennan utilizes the concept of mechanism as a metaphysical concept, accounting for the nature of causation. Let us call this the *metaphysical* motivation.

(MDC), Craver (2007), and Bechtel sees the concept of mechanism as providing facts about an *explanandum*-phenomenon that are explanatory relevant. As we saw in §4.2, the entities, activities and their organization that produce or underlie the *explanandum*-phenomenon are the *explanans*, i.e., the explanatory relevant and adequate facts for a

---

[80] (MDC), as we have discussed clearly have ontic views of explanation. Machamer (2004) disagrees on the nature of activities with Craver and Darden (which he co-authored (MDC) with), but seems to maintain an ontic conception of mechanistic explanation. Glennan (2002, 2005) explicitly adopts the ontic conception. A notable ally of Bechtel is Wright (2012).

[81] Glennan notes: "Curiously, nowhere in his book does Salmon offer an explicit definition of mechanisms" (Glennan 2002: 343). Seeing Salmon wanted to account for the nature of causation to vindicate his view on causal-mechanical explanation, it seems natural to view Glennan as attempting to do just this.

successful explanation.[82] Thus, the utility of the concept of mechanism is to provide an account of how scientists explain phenomena (at least in the proximal disciplines of biology). Let us call this the *explanatory* motivation.

Lastly, there is the view that the concept of mechanism and mechanistic modeling are a valuable utility for understanding and discovering complexity and complex systems. Bechtel & Richardson (2010 [1993]) states the following in the introduction on mechanistic explanation:

> Our aim is to develop a cognitive model of the dynamics of scientific theorizing that is grounded in actual scientific practice. Our focus is on one kind of explanation, one involved in understanding the behavior of complex system in biology and psychology. (Bechtel & Richardson 2010 [1993]: 17)

As we can see, Bechtel & Richardson conceives mechanistic explanation as crucial for scientific theorizing about complex systems. Thus, the concept of mechanism plays a strategic role, and the explanations are thought to be a part of this strategic role. Let us call this the *strategic* motivation.[83]

The motivations overlap, especially the strategic and explanatory motivation. However, I believe that an attention to these different motivations clarifies why there are some substantial differences between the mechanists. For example, it is not very surprising that Bechtel adopts an epistemic conception of mechanism, since his main motivation is to provide a *cognitive model* for scientific *theorizing*. Similarly, that Glennan opts for a reductive view of causation is also not very surprising, given that his main motivation is metaphysical. These different motivations are useful to demarcate the different views, and one should be aware of them when we inspect the different mechanistic aspects of the mechanistic framework. If we, for example, return to case of whether mechanism that are stochastically organized are to be counted as mechanisms we could approach the question with a strategic motivation and claim, yes, if looking at a stochastically organized mechanism yields greater understanding of the complexity of neural interaction, the we should view them as mechanism. On the other hand, if we approach with the explanatory motivation in mind the answer could be less positive. One could say that since we cannot account for the actual activities between the entities involved, the explanation of neural interaction still has the

---

[82] Bechtel (2006) and Bechtel & Abrahamsen (2005) would say that the description of the mechanism adequately *represents* the relevant factors that produce the phenomenon, because of their epistemic view of mechanistic explanation.

[83] I am following Levy (2013) in using ‘strategic’ and ‘explanatory’.

character of a mechanism-sketch, where there are still “black-boxes” that need to be opened.[84] However, even though the three camps in mechanistic philosophy have different motivation behind the use of the concept of mechanisms, there is one motivation that they all share, the notion that mechanisms should resemble how it is used in scientific practice. This is partly a motivation, but for some of these authors it works as a justificatory authority as well.[85]

§4.4 Consequences of the local approach

It is time to turn our full attention towards the meta-philosophical claim of my thesis and make it more explicit. I have already illustrated with some examples that if we keep in mind the restrictions that usually follow with a local approach to the philosophy of science, *especially pluralism*, we can get a clearer picture on intended scope of arguments for- and counter-arguments against the relevant theories, concepts or frameworks under discussion.

However, it is not always easy to keep the restrictions that follow with the local approach in mind, since it is not ‘follow’ in a strictly logical sense. It is in more ordinary “Wayne enters, Garth follows” sense. There is no necessity that Garth shows up whenever Wayne does, it is just the case that very often when Wayne shows up, Garth is there too. The point being, because the restrictions which usually follow from a local approach to the philosophy of science not *necessarily* follow, it is important that they are explicitly stated. If they’re not, we risk confusing the intended scope of, the relevant domain, the motivation or even methodology behind the view in question. This is the burden of pluralism; one cannot have the cake and eat it too. However, the burden is mutual between author and critic. In order to have a constructive debate, one has to be able to talk about the same things and there has to be an agreement on the scope of the domain of discourse.

One of the virtues of the logical empiricist doctrine to philosophy is that it is “easy” to criticize. What I mean by this is since the framework is global, the domain is already set. “All explanations can be formulated as arguments”, one might say, or “all true theories in the special sciences should reduce to physical theories in the long run” (to paraphrase Fodor 1974: 97). Such claims have already given us the domain, namely “true theories” and “explanations”. So, it is easy to criticize, in the sense that one doesn’t have to pay particular

---

[84] By “black-boxes” I mean a suggested mechanism that produces a phenomenon, where the components (i.e. entities, activities and organization) is unknown. See Craver & Darden (2013: ch. 5–9) for extensive discussion on how to open “black-boxes”.

[85] As I have already mentioned (MDC) appeals to scientific authority for their dual-ontology in neurobiology and molecular biology.

attention to the domain, motivations, or scope, it is already given. This, and the strictly nature of the logical of the logical empiricists framework, make them an easy target. A single counter-example could potentially be fatal. As we have seen, this was the case for reduction with multiple realizability.[86] When a framework is intended to be global counter-examples abound, and they are usually not very complicated. Take two of the counter-examples to the *DN*-model we have encountered already, "male pregnancy" and "the length of a flagpole". These are intuitively easy counter-examples to understand. The ideas they bring forth, however, are sometimes vital to the progress of the philosophy of science. *Explanatory relevance* and the *asymmetry of explanation* are ideas that the mechanists (amongst others) *model* their explanatory framework on (Craver 2007: ch. 2). So, the value of global approaches may prove to be of *negative* value, it brings out cases of particular importance.

Most of the counter-examples to the logical empiricists' concepts are due to the fact that the domain of discourse was not restricted beforehand. Restricting may of course be a perfectly reasonable thing, it may increase the chances of success, but it also serves to limit the number new possible discoveries.[87] Let's look at an example of such restriction:

> (3M) In successful explanatory models in cognitive and systems neuroscience (a) the variables in the model correspond to components, activities, properties, and organizational features of the target mechanism that produces, maintains, or underlies the phenomenon, and (b) the (perhaps mathematical) dependencies posited among these variables in the model correspond to the (perhaps quantifiable) causal relations among the components of the target mechanism. This principle is restricted to cognitive and systems neuroscience and so allows that there are legitimate nonmechanistic forms of explanation (Craver & Kaplan 2011: 611)

Such a restriction tells us what domain the authors are interested in having discussion in. However, it limits it audience, because, presumably you should know a great deal about cognitive and system neuroscience to evaluate the (3M), since the authors have explicitly stated it as the relevant domain of discourse. Lets call this the specialization consequence. I do not wish to judge the value of this consequence however; some may find it attractive and

---

[86] I do not mean that it is impossible for theory reduction to amend for the possibility of multiple realizability, but it is difficult without extensive provision or revision. In this case, it seems easier to find problems with multiple realizability itself, rather than with ways of circumventing the invalidity it produces. See Couch (2009) for some examples of the restricted possibility of using multiple realizability.

[87] I have to explicate what I mean here to avoid misunderstanding. By restricting one can discover new things at a *faster rate*, but one excludes possibilities. In other words, *not* restricting can yield the same discoveries as restricting (and several more), but it would generally be at a *slower rate*. There would be bigger number of "trails and errors" by not taking provisions. So there is no win-win situation here, I just note that this is a consequence of taking provisions.

necessary for progress in the philosophy of science, while some may see it as a serious obstacle for the availability of the philosophy of science literature for the audience. It is just important to note that it is a consequence for the local approach to philosophy of science.

To briefly sum up, the local approach usually encourages pluralism and explanatory autonomy, it is of significance importance that one explicitly states the relevant restrictions this puts on the debate. The critic should also respect these restrictions, but if they are implicitly or unclearly stated, or even left out, frustration and misunderstanding arise. Such restrictions also limit the availability of the literature to people with knowledge of the domain in question.

§4.4.1 The burden of pluralism and mechanistic philosophy

The restrictions of pluralism and burden of explicating these is, I believe, a source of conflict and frustration in the discussions on mechanistic philosophy. This is partly because the mechanists sometimes are vague (in some cases seemingly *intentionally* so), but also because critics refuse to acknowledge or are unaware of such restrictions. For example, Craver and Kaplan's restriction on the scope of their principle in cognitive and systems neuroscience is immediately followed up with "we see no reason to exempt all of cognitive science from the explanatory demands [that follows from the principle]" (Craver & Kaplan 2011: 611, fn. 10). Such a move makes it difficult to understand what domain they want the discussion to be in. Do they want to discuss the generality of their principle or do want to discuss their principle in cognitive and systems *neuroscience*? Presumably both, however, it is a move that doesn't encourage failure. If one argues back, "the explanatory demands do not conform well with approaches in cognitive psychology" they could just respond; "oh well, our paper is about cognitive and systems neuroscience, and that still works". The point here is that if the explanatory demands of cognitive psychology are different, then *why fix something that isn't broke?* A more charitable interpretation of Craver and Kaplan's move is that they are merely pointing out the possible new areas of mechanistic research. I believe such intentions lie behind this move. However, non-mechanistic philosophers may see this as the mechanists hiding behind a wall, not wanting negative contribution. Or that they just want to participate in debates where they are on the winning team.

To once more contrast this with the global approach, the restrictions follow from a local approach picks out the relevant domain for counter-examples. When one says that the restrictions can *in principle* be violated one opens up for wider domain of counter-examples. In such cases, it looks similar to the global approach, however with one "drawback"; if you find a counter-example the mechanists can immediately pull back behind the safeguard of their restrictions. Making such moves, that is conjecturing generalizations out from particular instances, without any commitment can make it unappealing for the rival of the mechanist to pursue.

I think such vague claims on the generality of the mechanistic approach also encourage disagreement within the mechanistic framework. Lets look at three different possible mechanists and their views on how to deal with the restrictions a local approach to the philosophy of science places on them. Especially how they view generality.

(M1) The first mechanist claims that the scope should be limited to cases that *perfectly* fit the mechanistic schema e.g., no stochastic mechanisms or "loose" organization. [88] He will presumably be uninterested in cases with stochastic mechanism and mechanism that are "loosely" organized, because they don't fit the schema. As such he can claim that sciences that deal with such phenomena are not mechanistic. He takes no interests in the generality of the mechanistic framework over and above the scientific explanations *actually* fit the schema.

(M2) The second mechanist claims that the different conditions are not that important. "We can be the rules without breaking them". He is interested in cases of stochastic mechanism and "loose" organization precisely because they are cases for extending the generality of the mechanistic framework to include a bigger number of phenomena.

(M3) The third mechanist is a sort of middleman, the best cases are those that are a perfect fit, but we should be open for revisions. He could be interested in stochastic mechanism or "loose" organization because they are cases where scientists don't know enough yet (cf. *epistemic regularity*), or because they may reveal possible advances on the mechanist framework. He might also be inclined to conjecture generality as Craver and Kaplan (2011).

However, if the three different mechanists don't explicitly state their interest or intentions, discussions on the possible generality will go across purposes. Such discussion could also lead to fruitful meta-philosophical and normative discussions on the mechanistic framework. Should it as general as possible? Should we "bend the rules"? Or should we stick

---

[88] By "loose" organization I mean systems that are organized, but poorly. Think of a football team that doesn't have coach. They are sufficiently organized to play football, but it is random who does what tasks. If one of the players finds himself on the one end of the pitch he's a striker, but if he is at the other he is a defender. In this way there is no telling where or what entities (players) of the system (team) will do what activities (task), but they are organized (they work together against the other tema). As such, it is case where mechanistic explanation, with all conditions met, will be difficult.

with the framework, *as it is*, in order to not loose the explanatory power inherent in the framework. I do not intend to solve such question, just pointing out that discussions on them are sorely missed from the perspective of a scientifically inclined, but not specialized, philosopher.

§4.5 Conclusion

I have argued that that the mechanistic framework can be seen as a continuation of the local approach to the philosophy of science. The global approach of the logical empiricists was the instigator of the shift towards the local approach. We have seen that the advances in scientific explanation, particularly with Salmon's causal-mechanical model are not easily characterized as global or local. However, the systems approach to inter-level explanation, which in conjunction with Salmon's causal-mechanical model constitutes the mechanistic approach explanation, is. A consequence of the local approach is that accounts of concepts are usually relative to the discipline or context. This makes the scope of concepts restricted. This is not a problem, but without explicit attention we can easily lose track and talk across purposes. This is equally true within mechanistic approach as outside it. If we are explicit about the intended scope of the concepts we can get fruitful discussion. One direction is in specialized local restriction, where *actual* scientific practice is the final judge. A second direction is in evaluating the potential generalizability of the mechanistic approach. If one tries to steer in between, one may easily get trapped. This can cause confusion and talk across purposes; even worse one could be interpreted as just wanting to be on the winning team, no matter the cost. Explicating the difference of these directions could also open for an audience without the specialized training in the special sciences.

References

- Achinstein, P. (1983), *The Nature of Explanation*, New York, NY: Oxford University Press.
- Andersen, H. (2011), “Mechanisms, Laws, and Rgularities”, *Philosophy of Science*, 78 (2): 325–331
- Andersen, H. (2012), “The Case for Regularities in Mechanistic Causal Explanation”, *Synthese*, 189: 415–432.
- Alexander, S. (1920), *Space, Time and Deity*, 2. Volumes. London: Macmillan.
- Armstrong, D. (1983), *What is a Law of Nature?*, Cambridge, MA: Cambridge University Press.
- Bechtel, W. (2006), *Discovering Cell Mechanisms: The Creation of Modern Cell Biology*, Cambridge Studies in Philosophy of Biology, Cambridge, New York, NY. Cambridge University Press.
- Becthel, W.
- Bechtel, W. & Abrahamsen, A. (2005), “Explanation: A Mechanistic Alternative”, *Studies in the History and Philosophy of Biological nad Biomedical Sciences*, 36: 421–441.
- Bechtel, W. & Richardson (2010 [1993]), *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*, Second Edition. Cambridge, MA. MIT Press/ Bradford Books,.
- Broad, C. D. (1925), *Mind and its Place in Nature*. New York: Harcourt, Brace & Company, Inc.
- Bromberger, S. (1966), “Why Questions”, in *Mind and Cosmos: Essays in Contemporary Science and Philosophy*, Colodny, R. (ed.), Pittsburg: University of Pittsburg Press.
- Bogen, J. (2005), “Regularities and Causality; Generalizations and Causal Explanations”, *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2): 397–420.
- Campbell, D. T. (1974), “Downward Causation”, *Studies in the Philosophy of Biology*, Ayala, F. & Dobzhansky, T. (eds.), University of California Press: 179–186

- Caroll, J. W. (2010), “Laws of Nature”, *The Stanford Encyclopedia of Philosophy* (Sping 2016), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2016/entries/laws-of-nature/>.
- Craver, C. F. (2001), “Role Functions, Mechanisms, and Hierarchy”, *Philosophy of Science*, 68 (1): 53–74
- Craver, C. F. (2006), “When Mechanistic Models Explain”, *Synthese*, 153 (3): 355–376.
- Craver, C. F. (2007), *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Oxford University Press. .
- Craver, C. F. & Darden, L. (2001), “Discovering Mechanisms in Cell Biology: The Case of Spatial Memory”, *Theory and Method in Neuroscience*, P. K. Machamer, Rick Grush, Peter Mclaughlin (eds.), Pittsburgh: University of Pitt Press, pg. 112–137.
- Craver, C. F. & Darden, L. (2013), *In Search of Mechanisms: Discoveries across the Life Sciences*. Chicago, IL: The University of Chicago Press.
- Craver, C. F. & Tabery, J. (2015), “Mechanisms in Science”, *The Stanford Encyclopedia of Philosophy*, (Spring 2016 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2016/entries/science-mechanisms/>.
- Churchland, P. S. & Sejnowski, T. J. (1992), *The Computational Brain,* Cambridge, MA: The MIT Press.
- Couch, M. (2009), “Multiple Realization in Comparative Perspective”, *Biology and Philosophy*, 24 (4): 505–519.
- Cummins, R. C. (1975), “Functional Analysis”, *Journal of Philosophy*, 72: 741–764
- Cummins, R. C. (2000), ““How Does It Work” Versus “What Are The Laws”: Two Conceptions of Psychological Explanantion.” *Explanation and Cognition,* In F. Keil, F. & Wilson, R. A (eds.)*,* MIT Press: *117-145*
- Darden, L. (1991), *Theory Change in Science: Strategies from Mendelian Genetics*, New York: Oxford University Press.
- Darden, L. (2006), *Reasoning in Biological Discoveries: Essays on Mechanisms, Interfield Relations, and Anomaly Resolution*, Cambridge Studies in Philosophy and Biology, Cambridge University Press.
- Darwin, C. (1859), *On the Origin of Species by Means of Natural Selection: or the Preservation of Favoured Races in the Struggle of Life*. Sixth Edition, with all

Additions and Corrections. [Published in 2008 by Arcturus Publishing Limited, London]

- Dennett, D. (1994), “Cognitive Science as Reverse Engineering: Several Meanigs of ‘Top Down’ and ‘Bottom Up’, in *International Congress of Logic, Methodology and Philosophy of Science*, Prawtiz, D. & Westerstahl, D. (eds.), Dodrecht: Kluwer, International Congress of Logic, Methodology and Philosophy of Science (9th; 1991).
- Dowe, P. (1992), “Process Causality and Asymmetry”, *Erkenntnis*, 32(2): 179–196.
- Dupré, J. (2013), “Living Causes”, *Proceedings of the Aristotelian Society*, 87: 19–35
- Fetzer, J. (2014), “Carl Hempel”, *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition), Edward N. Zalta(ed.), URL = <http://plato.stanford.edu/archives/fall2015/entries/hempel/>.
- Fodor, J. A. (1974), “Special Sciences (or: The Disunity of Science as a Working Hypothesis”, *Synthese*, 28: 97-115.
- Franklin-Hall, L., (forthcoming), “New Mechanistic Explanation and the Need for Explanatory Constraints”, in *Scientific Composition and Metaphysical Ground*, Aizawa, K. & Gillette, C. (eds.).
- Friedman, M. (1974), “Explanation and Scientific Understanding”, Journal of Philosophy, 71: 5–19
- Futuyma, D. J. (2005), *Evolution*, Sunderland, MA: Sinauer Associates, Inc.
- Godfrey-Smith, P. (1993), “Functions: Consensus Without Unity, *Pacific Philosophical Quarterly,* 74 (3): 196–208.
- Glennan, S. S. (1996), “Mechanisms and the Nature of Causation”, *Erkenntnis*, 44: 49–71.
- Glennan, S. S. (2002), “Rethinking Mechanistic Explanation”, *Philosophy of Science*, 69: S342–S353.
- Hedström, P. & Ylikoski, P. (2010), “Causal Mechanisms in the Social Sciences”, *Annual Review of Sociology*, 36: 47–67.
- Hempel, C. G. (1965), *Aspects of Scientific Explanation; and Other Essays in the Philosophy of Science*, The Free Press, New York, and Collier and McMillian Ltd., London.
- Hempel, C. G. & Oppenheim, P. (1948), “Studies in the Logic of Explanation”, *Philosophy of Science*, Vol. 15: 135–175.

- Hull, D. (1972), “Reduction in Genetics—Biology or Philosophy?”, *Philosophy of Science*, 39 (4): 491–499.
- Illari, P., (2013), Mechanistic Explanation: Integrating the Ontic and Epistemic, *Erkenntnis*, 78 (2): 237-255
- Kaufmann, S. A., (1971), “Articulation of Parts Explanation in Biology and the Rational Search for Them”, in *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 8, Boston: MA: 257-272
- Kim, J. (1992), “The Multiple Realization and the Metaphysics of Reduction”, *Philosophy and Phenomenological Research*, 52 (1): 1-26
- Kim, J. (2002), “The Layered Model”, *Philosophical Explorations*, 5 (1): 2–20.
- Kitcher, P. (1989), “Explanatory Unification and the Causal Structure of the World”, *Scientific Explanation*, Philip Kitcher & Wesley C. Salmon (eds.), Minnesota Studies in the Philosophy of Science, Vol: 13, Minneapolis: University of Minnesota Press, pg. 410–505.
- Kuhn, T. (1962) *The Structure of Scientific Revolutions*, Chicago, IL: University of Chicago Press. [50th anniversary edition 2012].
- Lange, M. (2007), *Philosophy of Science: An Anthology*, Malden, MA: Blackwell Publishing.
- Lakatos, I. (1977), *The Methodology of Scientific Research Programs: Philosophical Papers Volume 1.* Cambridge: Cambridge University Press.
- Leuridan, B. (2010), “Can Mechanisms Really Replace Laws of Nature?”, *Philosophy of Science*, 77: 317–340.
- Levy, A. (2013), “Three Concepts of New Mechanisms”, *Biology and Philosophy*, 28 (1): 99-114.
- Lewis, D. (1973), *Counterfactuals*, Cambridge, MA: Harvard University Press.
- Lewis, D. (1986), *Philosophical Papers*, Volume II, New York, NY: Oxford University Press.
- Lycan, W. G. (1987), *Consciousness*, First MIT Press paperback edition 1994, MIT Press.
- Machamer, P. (2004), “Activities and Causation: The Metaphysics and Epistemology of Mechanisms”, *International Studies in the Philosophy of Science*, 18: 27–39.

- Machamer, P., Darden, L., & Craver, C. F. (2000), "Thinking About Mechanisms, *Philosophy of Science*, Vol. 67 (1): 1–25.
- McLaughlin, B. P. (1992), "The Rise and Fall of British Emergentism", *Emergence or Reduction: Essays on the Prospects of Nonreductive Physicalism*. Edited by A. Beckerman, H. Flohr & J. Kim. New York: de Gruyter, 1992: 49–93.
- McLaughlin, P. (2001), *What Functions Explain: Functional Explanation and Self-Reproducing Systems*, Cambridge Studies in Biology and Philosophy. Cambridge University Press.
- Mayr, E. (1942), *Systematics and the Origins of Species*. New York: Columbia University Press.
- Mill, J. S. (1843), *A System of Logic*. New York: Harper & Brothers, Publishers, Franklin Square [1882, eight edition].
- Mitchell, S. D. (1997), "Pragmatic Laws", *Philosophy of Science*, 64: S468–S479.
- Mitchell, S. D. (2009), *Unsimple Truths: Science, Complexity, and Policy*, Chicago: University of Chicago Press.
- Nagel, E. (1961), *The Structure of Science: Problems in the Logic of Scientific Explanation*, New York: Harcourt, Brace, and World, Inc.
- Nicholson, D. J. (2012), "The Concept of Mechanism in Biology", Studies in History and Philosophy of Biological and Biomedical Sciences: Vol: 43, 152–163.
- O'Connor, T. & Wong, H. Y. (2012), "Emergent Properties", *The Stanford Encyclopedia of Philosophy*, (2012 spring edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2012/entries/properties-emergent/>.
- Oppenheim, P. and Putnam, H. (1958), "Unity of Science as a Working Hypothesis", *Concepts, Theories, and the Mind-Body Problem, Minnesota Studies in the Philosophy of Science II*, edited by H. Feigl, M. Scriven, & G. Maxwell. Minneapolis: University of Minnesota Press, 3–36.
- Putnam, H. (1967), "Psychological Predicates", in *Art, Mind, and Religion*, Capitan, W.H. & Merrill, D.D. (eds.), Pittsburgh: University of Pittsburgh Press: 37–48.

- Simon, H. A. (1962), “The Architecture of Complexity”, *Proceedings in the American Philosophical Society.* Vol. 106 (6).
- Simon, H. A. (1996 [1969]), *The Sciences of the Artificial*, third edition, Cambridge, MA: The MIT Press.
- Salmon, W. C. (1971), “Statistical Explanation”, in *Statistical Explanation and Statistical Relevance*, Salmon, W. C. (ed.), Pittsburg: University of Pittsburg Press: 29–87.
- Salmon, W. C. (1984), *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, Princeton, NJ.
- Salmon, W. C. (1989), “Four Decades of Scientific Explanation”, *Scientific Explanation,* Philip Kitcher & Wesley C. Salmon (eds.), Minnesota Studies in the Philosophy of Science, Vol. 13, Minneapolis: University of Minnesota Press: 3–219.
- Salmon, W. C. (1994), “Causality Without Counterfactuals”, *Philosophy of Science*, 61: 297–312.
- Salmon, W. C. (1997), “Causality and Explanation: A Reply to Two Critiques.”, *Philosophy of Science*, 67: 285–300.
- Salthe, S. N. (2012), “Hierarchical Structures”, *Axiomathes*, 22(3): 335–383.
- Schaffner, K. F. (1993), *Discovery and Explanation in Biology and Medicine.* Chicago, IL: University of Chicago Press.
- Sterelny, K. & Griffiths, P. E. (1999), *Sex and Death: An Introduction to Philosophy of Biology*, Chicago, IL & London: The University of Chicago Press.
- Tooley, M. (1977), “The Nature of Laws”, *Canadian Journal of Philosophy*, 7: 667–698.
- van Fraassen, B. (1980), *The Scientific Image*, Oxford: Oxford University Press.
- van Fraassen, B. (1989), *Laws and Symmetry*, Oxford: Clarendon Press.
- Wilson, F. (2007), “John Stuart Mill”, *Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), Edward N. Salta (ed.), URL = <http://plato.stanford.edu/archives/spr2016/entries/mill/>.
- Wimsatt, W. C. (1972), “Complexity and Organization”, *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Schaffner, K. F. & Cohen, R. S. (eds.), Dordrecht: Reidel, 1972: 67–86.

- Wimsatt, W. C. (1976a), “Reductionism, Levels of Organization, and the Mind-Body Problem”, in G. Globus, I. Savodnik & G. Maxwell (eds.), *Consciousness and the Brain*. New York: Plenum Press: 199–267.
- Wimsatt, W. C. (1976b), “Reductive Explanation: A Functional Account.” *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Michalos, A. C., Hooker, C. A., Pearce, G., & Cohen, R. S. (eds.), Dordretch: Reidel, 1974: 671–710.
- Wimsatt, W. C. (1986), “Forms of Aggregativity.” In *Human Nature and Natural Knowledge*, ed. Donagan, A., Perovich, N., & Wedin, M., Dordrecht: Reidel: 259–293.
- Wimsatt, W. C. (2007), *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Cambridge, MA: Harvard University Press.
- Wilson, F. (2007), “John Stuart Mill”. *The Stanford Encyclopedia of Philosophy* (Spring 2016 edition), Edward N. Salta (ed.), URL = <http://plato.staford.ed/archives/spr2016/entries/mill/>.
- Woodward, J. (2003), *Making Things Happen: A Theory of Causal Explanation*, Oxford University Press, New York, NY.
- Woodward, J. (2014), “Scientific Explanation”, *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2014/entries/scientific-explanation/>.