·38·

Mechanisms are real and local

Phyllis McKay Illari and Jon Williamson

Abstract

Mechanisms have become much-discussed, yet there is still no consensus on how to characterize them. In this chapter, we start with something everyone is agreed on – that mechanisms explain – and investigate what constraints this imposes on our metaphysics of mechanisms. We examine two widely shared premises about how to understand mechanistic explanation: (1) that mechanistic explanation offers a welcome alternative to traditional laws-based explanation and (2) that there are two senses of mechanistic explanation that we call 'epistemic explanation' and 'physical explanation'. We argue that mechanistic explanation requires that mechanisms are both real and local. We then go on to argue that real, local mechanisms require a broadly *active* metaphysics for mechanisms, such as a capacities metaphysics.

38.1 Introduction

Mechanisms have become much-discussed in the current philosophy literature, to begin to match the long-enduring interest in mechanisms in the sciences. Yet there is still no consensus as to the best way to characterize mechanisms. A brief glance at only three major papers attempting to characterize mechanisms illustrates:

> Machamer, Darden and Craver: '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 and Craver, 2000, p. 3)
>
> Glennan: 'A mechanism for a behavior is a complex system that produces that behavior by the interaction of a number of parts, where the interactions between parts can be characterized by direct, invariant, change-relating generalizations.' (Glennan 2002, p. S344)
>
> Bechtel and Abrahamsen: 'A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena.' (Bechtel and Abrahamsen 2005, p. 423)

In this chapter, we will start with one thing everyone – including both philosophers and scientists – agrees on about mechanisms: mechanisms

explain. We will investigate what constraints this imposes on a metaphysics of mechanisms. In Section 38.2 we examine two important premises about mechanistic explanation shared by many in the mechanisms debate: that mechanisms offer a form of explanation distinct from laws-based explanation, and that there two senses of explanation, that we call epistemic explanation and physical explanation. In Section 38.3 we argue that both kinds of explanation require real mechanisms, and in Section 38.4 we argue that scientific explanation using mechanisms requires that mechanisms must be local. In Section 38.5 we argue that real, local mechanisms require what we call a broadly *active* metaphysics, such as a capacities metaphysics, rather than a more passive best-system laws-based metaphysics. In Section 38.6 we deal with two possible objections to our view.

Dialectically, we take our argument to be important because the general trend in characterizing mechanisms in the literature is to use a more passive metaphysics, which our argument shows to be at odds with the basic reasons we examine in Section 38.2 that are often given for developing an account of mechanisms at all. Further, if you think, as many do, that mechanistic explanation is *causal* explanation, then everything argued here transfers to this important species of causal explanation.

38.2 Mechanistic explanation

There are various views about the nature of mechanistic explanation which are sometimes implicit, sometimes explicit in the mechanisms literature. In this section, we will pick out two views which, while not universal, are widely shared by those writing about mechanisms. They are first that mechanistic explanation is a distinct alternative to laws-based explanation, and secondly that there is something which we will call 'physical' explanation, where the mechanism in the world produces the phenomenon of interest. We agree with both of these deep premises of thinking about mechanisms and mechanistic explanation. We will argue that they impose constraints on a plausible metaphysics of mechanisms.

The first view is that mechanistic explanation is a new form of explanation that is distinct from traditional laws-based approaches to explanation – and far more promising as an account of explanation across much of the sciences. As Torres writes: 'The mechanistic model of explanation represents an appealing alternative to classical covering-law (CL) models' (Torres forthcoming, Section 38.1). There are three independent reasons for this, which influence different philosophers in different ways. All three reasons are important.

The first reason is simple: mechanistic explanation as distinct from law-based explanation fits the *practice* of the special sciences in a way that law-based explanation fails quite drastically to do. And the special sciences now

amount to a great deal of science.[1] They proceed by identifying phenomena requiring explanation, and decomposing them. They look for the parts of the mechanism underlying the phenomenon, and start trying to figure out what they do. They work out how the parts go together where, so that what the parts can do in conjunction changes, and ultimately the way the organized behaviour of the whole mechanism produces the phenomenon becomes clear. What such scientists do not do is look for laws and try to build up any kind of explanation resembling a law-based one. Bechtel and Abrahamsen are influenced by this: 'The received view of scientific explanation in philosophy (the deductive-nomological or D-N model) holds that to explain a phenomenon is to subsume it under a law .... However, most actual explanations in the life sciences do not appeal to laws in the manner specified in the D-N model.' (Bechtel and Abrahamsen 2005, p. 421–2)

The second reason for thinking an alternative to traditional laws-based explanation is vital is also simple, and has been extensively argued elsewhere. There are no laws – no exceptionless non-accidental generalisations – in the special sciences and no reason to suppose there ever will be.[2] Those 'laws' that do exist in the special sciences are nothing more than generalizations backed up by an understanding of the mechanisms underlying those laws.[3] This is the direction of explanation between laws and mechanisms that many scientists adhere to. It seems to us that a reductive faith that there will ultimately be discoverable laws everywhere has been seriously empirically undermined by the huge success of science without laws – such as the extraordinary amount now known about the detailed mechanisms of protein synthesis – and should be abandoned. Leuridan mentions this reason for preferring mechanistic explanation: 'If there are no strict laws, there are no D-N explanations. Hence the mechanicist alternative, which states that explanation involves *mechanistic models* (i.e. descriptions of mechanisms) instead of strict laws, might be very welcome.' (Leuridan forthcoming Section 38.1, emphasis in original)

The third reason is that even if fundamental laws were available in the special sciences – exceptionless non-accidental generalizations that we could treat as brute nomological facts about the universe – it is hard to see how they could explain in the way that mechanisms explain. Brute nomological

[1] The special sciences span from chemistry to psychology, economics and social science. Some areas traditionally thought of as physics might still best be thought of as special sciences, as James Ladyman has argued in private communication. He suggests optics, cosmology and solid-state physics. We don't mean to imply that mechanistic explanation is *absent* from physics, just that it *is* ubiquitous in the special sciences.

[2] Cartwright has been very influential here. She even argues that many of the classic universal laws of fundamental physics describe only idealized, closed systems, so are in fact *not* universal laws when applied to the real world. See Cartwright (1983).

[3] Glennan (2002) discusses this extensively. Note that idealised models are a different issue, quite distinct from physical mechanisms – consider that they are *models*. See also Mitchell (2007).

facts themselves call for explanation, which is precisely why scientists do try to explain laws – and they often use mechanisms to try to do this. Glennan (2002) is concerned about this issue, concluding that at least the fundamental laws of physics are not mechanically explicable, or explicable at all. It seems to us that this is acceptable if it is assumed that there is a final, fundamental, lowest, level of physics. But so far we have not discovered such a final level, and we don't have empirical reason to believe there must be one. Mechanisms offer a more accurate characterization of genuine special science explanation. This reason clearly influences Bechtel and Abrahamsen: 'For present purposes we leave laws in place as statements of particularly robust and general phenomena. However, we suggest that explanation is to be found in the mechanisms that account for these laws, not in the laws themselves.' (Bechtel and Abrahamsen 2005, p. 422, footnote 1)

So the first view that mechanistic explanation is a promising and much-needed alternative to laws-based explanation is widely shared and important.[4]

The second view is sometimes only implicit in the literature, but it is clearly present. It is the view that there are two different kinds of mechanistic explanation in line with Salmon's more general distinction between epistemic and 'ontic' explanation (Salmon 1998a, 1998b). The first, epistemic, sense is of explanation as a human practice, aimed at increasing understanding of the world. It often involves the passing of information between human beings. Although it is aimed at understanding the world, it is highly sensitive to the cognitive abilities and background knowledge of human beings.[5] The second, ontological, sense of explanation, particularly important in the scientific practice of explanation using mechanisms, we will call physical explanation. This second sense of explanation is the sense in which mechanisms explain the phenomena they explain by being responsible for them. This happens whether human beings understand what is going on or not.[6] Knowledge of the mechanisms involved in physical explanation might in some cases be beyond our cognitive capacities. There is certainly no a priori reason to be sure we will always be able to know such mechanisms.

[4] It is not universal. We will show later that careful study of Psillos' views suggests that he believes mechanistic explanation collapses to laws-based explanation. And see Leuridan (forthcoming) for an interesting examination of mechanisms and laws-based explanation using Sandra Mitchell's understanding of 'pragmatic' laws, which are intended to be available in the special sciences.

[5] It is in this sense of explanation that the explanation a biologist gives his small son is very different from the one he will give in a research seminar. It is also the sense in which an explanation which is strictly speaking *false* might often be the best way of explaining something to a particular audience. One might say, for example, 'Mummy is getting fat because she has a baby in her tummy'. Strictly speaking Mummy is not getting fat, nor is the baby in her stomach. Thanks to Julia Tanney for suggesting this useful example.

[6] We will discuss the tricky cases of social and psychological mechanisms in Sections 38.3 and 38.4.

The two senses of explanation are naturally intertwined, of course. If the epistemic sense of explanation is to succeed in increasing understanding of the *world*, rather than merely making up interesting stories about it, the stories had better be describing the mechanisms in the world. In the other direction, as soon as we start trying to *describe* the physical mechanism that produces the phenomenon, we begin to abstract, to prioritise salient details and so on, so that our *description* takes on some epistemic features. But while these two senses of explanation are so intertwined that they can be difficult to separate, they are not the same.[7] This can be clearly seen by realising that they are subject to different constraints. For example, an epistemic mechanistic explanation should be perspicuous to its audience, while the correctness of a description of a physical mechanism producing a phenomenon does not depend on whether its audience understands that description. A second important difference is that the epistemic sense of explanation allows for the imposing of normative desiderata that would be inappropriate to apply to physical mechanisms. For example, one might reasonably attempt to make epistemic mechanistic explanations modular – very crudely, organised so that you can wiggle one bit of the mechanism without affecting the working of other bits.[8] But you couldn't possibly demand that every physical mechanism be modular. They may be or they may not. These differences exist because in the epistemic sense of explanation it is the *description* of the mechanism that explains, while in the physical sense, the *mechanism itself* does the explaining.

This distinction is implicit in much of the literature on mechanisms, and explicit in some papers. In many papers, an explicit distinction is made between mechanisms as they are in the world, and the models, schema or descriptions of mechanisms that we construct to explain the phenomena.[9] In the construction of models or schema we are involved in an explanatory project that is epistemic. This notion is much explored by Bechtel and Abrahamsen, among others, and of course has a very important place in science. But Machamer, Darden and Craver, Bechtel and Abrahamsen, and Glennan all recognize the deeper relation between mechanisms and phenomena. The mechanism as it is in the world is responsible for the phenomena we observe.

[7] Achinstein explores the relation between explaining acts and worldly explainers. See Achinstein (1983).

[8] Woodward does this. His section 5 is all about modularity, and he writes: 'The basic idea that I want to defend is that the components of a mechanism should be independent in the sense that it should be possible in principle to intervene to change or interfere with the behavior of one component without necessarily interfering with the behavior of others.' (Woodward 2002, p. S374).

[9] See Glennan (2002), Machamer, Darden and Craver (2000), and Bechtel and Abrahamsen (2005). Craver goes on to make the distinction thoroughly explicit. He talks of explanations as texts, models or representations used to convey information. He goes on: 'Other times, the term explanation refers to an objective portion of the causal structure of the world, to the set of factors that bring about or sustain a phenomenon.' (Craver 2007, p. 27).

Machamer, Darden and Craver phrase it as the mechanism being 'productive of regular changes'; Glennan says that the mechanism 'produces that behavior'; and Bechtel and Abrahamsen say that mechanisms are 'responsible for one or more phenomena'. They all mean much the same thing.

Bechtel and Abrahamsen don't want to call this kind of thing explanation. We do, following both scientific and common practice. Sometimes a request for explanation is a request to identify the responsible portion of the world. For example, when asking why the Sun looks larger at the horizon, the requested explainer is something about the world, even though it is described in terms of something epistemic – a story about the world. So this physical sense of explanation is a genuine sense of explanation, and it is this physical sense of explanation that we are most concerned with in this paper. It comes in at least two varieties – decompositional and etiological. In the first case, mechanisms explain phenomena by being the lower-level entities and activities that are organized to produce the higher-level observed phenomena. In the etiological case, mechanisms explain events by being the detailed causal history leading up to these events. Although a decompositional mechanistic explanation for a one-off event could be sought, we will focus on the case where mechanisms are the decomposition of a regularly occurring phenomenon. There are interesting relations between these slightly different kinds of mechanistic explanation which we do not have space to explore here.[10]

In conclusion, scientific practice allows for two distinct understandings of mechanistic explanation: epistemic, where the description of the mechanism explains, and physical, where the mechanism itself explains. The literature on mechanisms has followed this, and the view that physical explanation exists, where the mechanism produces the phenomenon of interest, is widely shared. Similarly, the view that mechanistic explanation is a crucial alternative to laws-based explanation is a driving force of the mechanisms literature, widely shared in it.

38.3 Mechanistic explanation requires mechanisms to be real

For now, we put aside the view that mechanistic explanation is distinct from laws-based explanation, although it will become crucial to the overall argument again in Section 38.6. In this section and the next, we focus on teasing out two implications for the metaphysics of mechanisms arising from the use of mechanisms in explanation: mechanisms must be real in this section and local, in Section 38.4.

For mechanisms to yield physical explanations, they must be real. That is, at least some of the mechanisms posited by scientists must exist as worldly

[10] See Glennan (2002) and Craver (2007) for interesting discussion.

entities responsible for the phenomena we take them to explain.[11] This fits scientific practice, matching how scientists treat mechanisms in mechanistic explanation. The mechanism of protein synthesis is a system in the world that produces proteins; natural selection is a complex of worldly processes that produces the adaptation of a population to its environment.[12] It also fits the interesting uses to which mechanisms might be put, in causal inference, for example. Leuridan and Weber (2011, this volume) suggest that identification of an underlying mechanism can be important in addressing the problem of external validity – working out whether a causal relation established in one population will apply in another. If knowing the underlying mechanism is to help, it is because you have identified a worldly entity that you can with reasonable confidence expect to be present/absent or more or less similar in altered circumstances. Cartwright (2006 and forthcoming) suggests that knowing social mechanisms is important to making social policy decisions, which seems to make the same assumption. Physical explanation is the most important kind of mechanistic explanation, because epistemic explanation is parasitic upon it. If epistemic explanations are to explain, rather than merely being stories, there must be real mechanisms to describe. So both kinds of mechanistic explanation require real mechanisms.

So the mechanisms literature is implicitly or explicitly committed to mechanisms being real. The claim is seldom made so baldly: perhaps it has seemed too obvious to comment on. A more detailed understanding of how mechanisms are real is slightly tricky. There are two problems, which we will take in turn: non-'physical' mechanisms, and the functional individuation of mechanisms.

The problem of non-'physical' mechanisms arises because not all mechanisms are completely independent of what people believe of them and how people describe them – the existence of social and psychological mechanisms means that that independence comes in degrees. Whatever social and psychological mechanisms are, they are partly constituted by people, and some may include the beliefs of people as components. For example, enough people believing that the economy is going into recession, and behaving accordingly, might be a necessary step in the economy entering recession. In terms of traditional debates about the existence of the external world, the dependence of such mechanisms on minds, their representational or constructed nature, might be thought to be an unacceptable mind-dependence of such mechanisms, undermining their reality.

[11] The claim that mechanisms are real does *not* commit us to some variety of physical reductionist thesis. We believe there is serious reason to doubt whether science really proceeds by classical reduction. See for example Craver (2007) for an argument that the field of neuroscience has not advanced in ways remotely resembling the reductive paradigm.

[12] Skipper and Millstein argue that natural selection is not a mechanism – at least not one that fits the characterizations of mechanisms offered by MDC and by Glennan. We disagree. See our examination of natural selection in Illari and Williamson (2010).

However, this does not mean that social and psychological mechanisms must be merely explanatory schemas or models, thoroughly and essentially dependent on human minds, and so forever restricted to merely *epistemic* explanation. So long as people, and their minds, exist, are part of the ontology of the world, such mechanisms exist and are susceptible to scientific investigation. Such mechanisms are still worldly entities. Further, although models and schemas might under some circumstances themselves be mechanisms, they cannot themselves be part of the mechanisms that they describe. This view also contrasts with the view that such explanatory posits are merely instruments – instruments for making accurate predictions about the behaviour of observables.[13] An instrumentalist view renders science less explanatory. There is no commitment to anything that can explain why the predictions work.

The second problem for the reality of mechanisms is that mechanisms are at least partially functionally individuated. This is a light sense of 'function', where the function of a thing is the role it plays in a system. That role does not require a selective history. Mechanisms are conceived of as the *mechanism for* producing a particular behaviour, where the behaviour is identified as a phenomenon of interest by scientists, when characteristically producing that behaviour is the role of that mechanism in that system.

This partial functional individuation of mechanisms is clear in scientific practice. Take for example the discussions of the mechanism of protein synthesis in three biochemistry textbooks. The notion of function is ubiquitous: all three texts took the function of protein synthesis to be the decoding of the information in DNA to produce proteins, and many lower-level mechanisms involved were functionally described.[14] For example, regulatory mechanisms are understood in terms of ensuring that the right proteins come out. All three texts talk about some kind of repair mechanisms, and they all mean mechanisms whose function is to correct various kinds of malfunction.

This kind of functional individuation of mechanisms is also recognized in the philosophical literature. Both Glennan, and Bechtel and Abrahamsen, are explicit that mechanisms have functions. For Bechtel and Abrahamsen a mechanism is a structure 'performing a function' (2005, p. 423), while for Glennan a mechanism is a mechanism 'for a behaviour' (2002, p. S344). Machamer, Darden and Craver's characterization does not explicitly make a function an essential aspect of a mechanism, but they don't explicitly rule it out either. Independent work by Darden and by Craver makes the link between a

[13] Instrumentalism has a long intellectual history, being traceable to work of Mach and Duhem, among others. See van Fraassen (1980) for a more recent discussion.

[14] See Adams *et al.* (1992), Voet and Voet (2004), and Whitford (2005). For example: 'Uncovering the cellular mechanisms resulting in sequential transfer of information from DNA (our genes) to RNA and then to protein represents one of major achievements of biochemistry in the 20th century.' (Whitford 2005, p. 247) Or consider: 'How do genes function, that is, how do they direct the synthesis of RNA and proteins, and how are they replicated?' (Voet and Voet 2004, p. 92).

mechanism and its explanandum phenomenon explicit (see Darden 2006 and Craver 2007).

The possible problem for the reality of mechanisms is probably clear. A function is required for something to be a mechanism. But the role something plays in a system seems to be a matter of description – the description of the system it is in. The same object individuated structurally, such as the heart, can have different functions according to the description of the system it is in. It might have the function of pumping blood when considered as part of the circulatory system, or the function of making a thump-thump noise when considered as part of a system for comforting a newborn baby. Then functions seem to be not wholly in the world, but set by the description of the system. This is not a spurious unscientific example. In scientific practice, the very same thing individuated structurally can sometimes have one function, sometimes another – and sometimes has no function at all.

Consider an example of this from a discussion of branch migration, the moving of the crossover point at which two molecules of DNA exchange excised strands of DNA. Voet and Voet write: 'such a process moves forward and backward at random and, moreover, is blocked by as little as a single mismatched base pair. In *E.coli*, and most other bacteria, branch migration is an ATP-dependent unidirectional process that is mediated by two proteins . . . ' (Voet and Voet 2004, p. 1189.) In the context of protein synthesis as having the function of decoding DNA to produce proteins, branch migration doesn't have a function. It just happens, being a nuisance that the cell has to have various mechanisms to fix. That is why it is described as a process here.

The same responsiveness to description occurs for other kinds of mechanisms. Take natural selection. When it's the mechanism for adaptation, then it has a function – the bringing about of fit between organism and environment – and so is properly described as a mechanism. But when something else is being described and natural selection is just something happening in the background, then it doesn't have a function, and is described as a process. There are many other examples showing that both biochemists and evolutionary biologists consistently use 'mechanism' to describe only something that has a function, using the word 'process' when there is no function. So mechanisms are mechanisms *for*, while processes are just processes in themselves. Mechanisms, of course, are still what is used in explanation because once in the domain of explanation, you are thinking in terms of a behaviour or phenomenon to be explained. Once you identify that phenomenon, then you are looking for something that has a function – the mechanism that produces that phenomenon, which is thought of as the mechanism *for* that phenomenon.

This looks like a more serious problem for the reality of mechanisms than it really is. It is true that what function a mechanism performs can vary according to our *description* of the system it is in. But this is because many

entities *really do belong to more than one system*. The heart really does circulate blood, and it really does make a thump-thump noise which is comforting to newborns because it is familiar. When we consider different systems we will naturally identify a different role-function for the heart, but this is because of the variation between systems, not because of our *description* of that system. The role an entity plays in any particular system is not a matter of description – it is a worldly fact. Functionally individuated mechanisms are real.

In conclusion, both scientific practice and much of the philosophical literature on mechanisms is committed to real mechanisms, either in physical explanation, or as items to be described in epistemic explanation. That is, mechanisms exist as worldly entities that are responsible for the phenomena they produce. The mechanism itself is different from any model, schema or other description or representation of *that* mechanism, and the mechanism itself is real. Mechanisms are physical explainers, while representations of mechanisms are epistemic explainers. This is the first important constraint on a metaphysics of mechanisms.

38.4 Mechanistic explanation requires mechanisms to be local

The second interesting constraint on a metaphysics of mechanisms comes from a feature of the *kind* of explanation you get when you get a physical mechanistic explanation. The explanation is local to the phenomenon produced. Mechanistic explanations use parts organized to produce the phenomenon, and may also look to containing systems to set the phenomenon, but both parts and systems are local to the phenomenon. This is true of scientific practice, and is also reflected in the philosophical literature on the issue.

Mechanisms, recall, are individuated by their explanandum phenomena. Mechanistic explanation begins by identifying a phenomenon, usually a regularly occurring one, to be explained. What then occurs is the process of examining the area local to the explanandum phenomenon to see how the phenomenon is produced. This is how characteristic mechanistic explanation proceeds.[15] Take protein synthesis as an example. Scientists have a phenomenon that they wish to explain – the production of proteins. What they want to know about is the underlying mechanisms that produce proteins, and the

[15] This understanding of decompositional mechanistic explanation is uncontroversial in the mechanisms literature. See for example Bechtel and Abrahamsen: 'The quest to understand the mechanism responsible for a given phenomenon requires decomposing the responsible system.' (2008, p. 560) They also discuss the importance of looking upwards to higher-level mechanisms. For this chapter, we set aside etiological mechanistic explanation. We suspect that it will exhibit the same kind of locality as decompositional mechanistic explanation, but lack space to establish that here.

conditions in the cell at the time. So scientists examine the details of the cells where proteins are produced. What is going on when proteins are produced? What kinds of molecules and so on are involved, and what do they do? They discover that there is a stage where DNA is copied, a stage where (various kinds of) RNA are made from DNA, and a stage where protein is made from mRNA. In fact, they discover vastly more information than this about all of these stages. But what they are doing certainly involves decomposing into local parts, and investigation of the local cell conditions. They look for the lower level stuff that produces the higher-level stuff, and the conditions under which this happens. In this sense, a great deal of what is described at great length in biochemistry textbooks just *is* protein synthesis.

Precisely because of the functional individuation of mechanisms, the extent of the locality you expect in a mechanism is set by the phenomenon the mechanism produces. Protein synthesis is very constrained, happening within a single cell. This is where the proteins appear, and where we look for the mechanisms that produce them. Compare natural selection, which happens within entire populations – much larger entities than single cells. Nevertheless, if you are interested in how natural selection has produced a particular distribution of trait types in a population, you study the history of the happenings in *that* population. You will also have to study the environment of that population, but that is also local – you don't look anywhere else. Investigating social mechanisms might lead you to look at social relations in an entire country, or even internationally. Investigating emission mechanisms for gravitational waves would lead you to look at something much bigger – the behaviour of binary pulsars, or even the movement of entire galaxies. But in each case the phenomenon sets the extent of appropriate locality. It isn't that mechanisms are always the size of 200 molecules, or three galaxies, it is that they are constrained in locality by where the phenomenon is that they are thought to produce. Other regions of space are not investigated, since they are not considered relevant.

Locality is implicit in much philosophical discussion of mechanisms, and explicit in some cases. Although he doesn't use the word 'locality', Craver, for example, gives the matter some extended discussion. Craver thinks that mechanisms are generally hierarchically nested, and mechanisms at different levels fall in (certain sorts of) part-whole relations. He writes 'The primary difference is that LM levels [levels in the multilevel mechanism for spatial memory] are relationships between a whole and its parts, while levels of processing are relationships between distinct items.' (Craver 2007, p. 178) Clearly, parts of wholes are local to those wholes, so Craver thinks that lower-level mechanisms are local to higher-level ones.

Craver also examines the issue of what, in the local area of the explanandum phenomenon, might be *left out* of the mechanism for the phenomenon. Not everything local is relevant. Writing again of his spatial memory case study, he

claims: 'Components at lower levels are organized to make up the behaviours at higher levels, and lower- and higher-level items stand in relationships of mutual manipulability…' (Craver 2007, p. 170) By mutual manipulability, he means, very crudely, that if you wiggle one, the other wiggles, and vice versa. Not all parts of the cell, for example, are components of the mechanism of protein synthesis in this sense. And this is generally true of mechanistic explanation. If you are investigating social relationships the behaviour of the atoms or molecules of people are irrelevant, and a star collapsing will emit gravitational waves quite independently of the behaviour of any nearby life. While this shows that not everything that is local will be part of the mechanism, it is true that the parts of the mechanism will be local to where the phenomenon they produce exists. Mechanisms come in different sizes because the phenomena they produce come in different sizes. The extent of mechanisms that produce various phenomena is part of what is learned by scientists as the mechanisms for the phenomena are better understood. One of the things that is surprising about various subatomic phenomena in physics, for example, is how non-local their causes and effects might be.[16] Nevertheless scientists in different fields come to have a good idea of how dispersed the mechanisms they are looking for are likely to be. Once that is established, they do not look further afield. It is, if you like, an empirical discovery of any mechanism what extent of local space-time is relevant to the explanation. In all cases, that locality is limited.

Our discussion of the locality of mechanisms sails confusingly close to debates elsewhere in philosophy. It is worth pausing to make clear that the locality claim is a claim about mechanistic explanation in scientific practice and how it is understood in the mechanisms literature. We are not here concerned with metaphysical debates about causality, such as the claim that causal relations are intrinsic.[17] The locality of mechanisms might imply that a precise copy of a mechanism is still a mechanism, which is an interesting parallel of some ways of understanding the intrinsicality claim. But this debate is not what concerns us here. We are also not concerned with alternative *scientific* claims about locality. A familiar locality claim from physics is that there can be no covariation of physical properties that are spacetime separated, where the covariation would demand that the propagation of any causal influence would have to be faster than the speed of light. This might be recast in less scientific terms as a doctrine of no action at a distance, which confusingly enough seems to be what some philosophers mean by locality when they talk about

[16] In this way, quantum non-locality isn't a counterexample to our locality claim, since quantum phenomena are still local to the system. What's surprising is how spread out in space that system can be.

[17] Ned Hall characterises this informally: 'Intrinsicness: The causal structure of a process is determined by its intrinsic, non-causal character (together with the laws).' (Hall 2004, p. 225).

causality.[18] If these are basic laws of physics, then no doubt mechanisms will not violate them. But such laws require support from physics more generally – examining how mechanisms are used in explanation cannot establish that nothing moves faster than the speed of light! Since we are looking for a claim supported by the use of mechanisms in explanation, it should be clear that this is not the locality claim we have in mind.

To reiterate, the claim is that in decompositional mechanistic explanation, the mechanism that produces the phenomenon of interest is looked for and discovered in the area of the phenomenon produced. This makes perfect sense of the idea that the mechanisms discovered are the mechanisms underlying the phenomenon. Other regions of space and time are not considered relevant.

Before ending this section, we will raise three possible problems for the locality of mechanisms: the functional individuation of mechanisms, the existence of non-'physical' mechanisms, and omissions. We will argue that none puts paid to locality, although examining them does yield a better grasp of the locality claim.

The first problem is that mechanisms having functions potentially damages the locality of a mechanism. If functions are relative to a description of a system, and that description, along with the wider system, is not part of the world where the mechanism is, then mechanisms are not wholly local. However, just as for the reality of mechanisms, this problem can be resolved by recognising that functions are more accurately thought of as objective, worldly relations between a mechanism and the (perhaps several) different systems it is in. Consider the heart again, as structurally individuated. Its role as blood-circulator is set by the higher-level system of blood circulation it is a part of, while its role of baby-comforter is set by that different system. But this is not a matter of description, just a relation between the heart and two different (and local) systems it really is a part of. This is not a serious problem for locality any more than it is for reality.

Again, the possibility of non-'physical' mechanisms requires careful thought. We have pointed out already that locality is always relative to the phenomenon explained, so that some non-'physical' mechanisms, such as social mechanisms, might be very large indeed. But it is also true that how thoroughgoing the locality is, how deeply it extends, varies from mechanism to mechanism, affecting effective strategies of decomposition.

Craver recognizes this: 'Localization is one of the most fundamental spatial constraints on interlevel integration (Bechtel and Richardson 1993). Not all mechanisms have easily localized *components*, but when they do, the location

[18] Ned Hall, for example, uses the following sense of locality: 'Locality: Causes are connected to their effects via spatiotemporally continuous sequences of causal intermediates.' (Hall 2004, p. 225).

of different processes can be crucial to understanding a mechanism that incorporates them.' (Craver 2007, p. 261–2) This is a particularly interesting issue for psychological mechanisms, where there is great debate on localization strategies. There is disagreement, for example, on whether specific psychological functions are localized to particular parts of the brain, or processed in a way far more dispersed through the brain.[19] Nevertheless, this becomes an issue once the description of the phenomenon is underway, which is once the locality of the mechanism is already set. In these cases, both sides of the debate still agree that the mechanisms producing psychological phenomena are in the brain. That is local enough.

There is a far trickier problem case.[20] Some decompositional psychological mechanisms cite content-bearing mental states as part of the mechanism. For example, the behaviour of a rat in a maze might routinely require explanation in terms of false beliefs about where food is – where of course the *content* of the false belief is crucial to the explanation. If mental content is external, depending on the history of the interaction of the organism with its environment, then this particular decompositional mechanism is not local. The phenomenon may arise in the brain, but some of its constituents are elsewhere.

This is of course a controversial case. It seems to us there are three options. The first is to say that psychological mechanisms are just quite different from other mechanisms in this way. This is a position one might be driven to, but is ad hoc as a first move. The second option is a denial of externalism about content. All that is necessary for mechanisms is the narrow content of any content-bearing mental state, which supervenes only on the local brain state. Our third option is to take externalism about content more seriously – perhaps as seriously as the extended mind thesis does. Then, we re-construe the appropriate extent of mental phenomena, so that external items are appropriate parts for decomposition of mental events.[21] These are interesting issues worthy of wider consideration, but we lack space to develop them here. We have said enough to indicate the defence of our locality thesis, because on either the second or the third story, the case doesn't violate locality.

The third problem is difficult. The problem is that if omissions are part of any mechanism, which they routinely seem to be, then that mechanism will not be local. The mechanism might depend on features thoroughly nonlocal to the phenomenon it produces – perhaps on features that have no location at all. The problem of locating omissions is well known from the vast literature on omissions and *causation*.[22] To illustrate, suppose I, your

[19] See the debate about core realizers. Advocates include Block (2005) and Wilson (2004). These are very interesting issues we intend to examine at length elsewhere.

[20] Thanks to Anthony Everett for raising this point.

[21] See the forthcoming Menary collection for extensive discussion of the extended mind thesis.

[22] See for example Schaffer (2004) and Dowe (2004) for an interesting symposium.

neighbour, promise to water your plants when you are on holiday. I don't, and they die.[23] It seems my omitting to water the plants causes their death. But where is my omission located? Is it in your house where I fail to turn up? Or is it wherever I am, busy not watering plants? There seems to be no satisfactory answer. Yet it is not impossible that omissions are routinely part of some mechanisms. In the context of causation, Schaffer discusses the example of a gunshot through the heart causing death, but it isn't difficult to see the circulatory system and the internal workings of a gun that he describes as mechanisms. Schaffer writes: 'But heart damage only causes death by negative causation: heart damage (c) causes an absence of oxygenated blood flow to the brain ($\sim d$), which causes the cells to starve (e).' Later he considers the gun: 'But trigger pullings only cause bullet firings by negative causation: pulling the trigger (c) causes the removal of the sear from the path of the spring ($\sim d$), which causes the spring to uncoil, thereby compressing the gunpowder and causing an explosion, which causes the bullet to fire (e).' (Schaffer 2004, p. 199) Examples in biochemistry are numerous. Cells routinely alter which enzymes they produce in response to which metabolites are available. A cell stops producing lactase, for example, in response to the *absence* of lactose diffusing into the cell's cytoplasm.

Our view is that omissions are not a problem for the locality of mechanisms in the same way that they are for causation. Causal relations are often subject to absences. Just as my failure to water your plants caused their death, so the Queen's failure to water them caused their death, and so on. There is no location for these omissions. But the kinds of omissions that are routinely part of mechanisms are locatable, and they are local. This is clear in all three examples above. The deoxygenated blood is part of the circulatory system within the body; the removal of the sear from the path of the spring is part of the gun mechanism, and the alteration of gene expression within the cell is in response to differing levels of metabolites within that cell. All of these are in the familiar area where we would attempt a decomposition of the relevant phenomena.

Causation and mechanisms are different here because difference-making is important to causation and not to mechanisms. Often, the conviction that difference-making is crucial to causation is decisive on thinking that absences cause. This is because in the causation case an absence often stands in a difference-making relation to the effect. To illustrate, it is missing the bus that *made the difference* to my lateness, not the empty bus-stop when I got there.[24] That is why it is the *absence* that causes my lateness, rather than the positive

[23] This useful example is originally Helen Beebee's. See for example her (2004).

[24] We take this important point to be a central view of David Lewis' later work. See Lewis (2004). For further discussion of the idea of difference-making and its relation to causation, see Hall (2004). Difference-making itself might be tracked in different ways, such as using simple counterfactual notions, invariance relations, or correlations.

story about where the missing bus actually is. This idea of difference-making is plausibly very important to causation, but it isn't central to a mechanism. So such external omissions are not in the same way relevant to mechanisms, being outside the relevant mechanisms. This means that only local omissions are relevant to mechanisms.

In conclusion, and in spite of some challenging problems, it is a genuine feature of mechanisms that they are local. This is a kind of locality that scientific practice is committed to in how it explains using real mechanisms, and a locality that is, albeit sometimes implicitly, reflected widely in the mechanisms literature. This is the second important constraint on a metaphysics of mechanisms. Mechanisms are both real and really there, and the right metaphysics of mechanisms must respect that.

38.5 Reality and locality require an active metaphysics

The problem for a good metaphysics of mechanisms is to characterize the *interactions* in mechanisms. Recall that all characterizations of mechanisms have two components, with something about the parts of mechanisms, and something about how the parts interact. The metaphysics of the parts of mechanisms has been uncontroversial. They have had different names – 'entities', 'parts' and 'component parts' in the characterizations above – but they have not been much discussed. What is controversial is how to characterize what the parts of mechanisms *do*, the activities or interactions of the parts. It is not surprising that this is controversial, since the interactions are more interesting metaphysically. Unlike the parts, or entities, which are actual, understanding mechanisms also involves understanding what those entities will do in non-actual situations. Scientists seem to know of many mechanisms what they would do when the initial conditions are changed in various ways.

It is in characterizing the interactions of parts of mechanisms that the reality and locality of mechanisms becomes relevant, because not all metaphysical approaches to the interactions allow mechanisms to be real and local. Understanding the metaphysics of mechanisms on this level is now a philosophical problem with no immediate bearing on scientific method, of course. It does, however, bear on our understanding of science. Since both scientific practice and many philosophical treatments of mechanisms are committed to their locality and reality, our argument should be of wide interest.

There are broadly two approaches to characterizing the interactions in mechanisms, and we will argue that all the approaches to metaphysics that allow mechanisms to be real and local lie within only one of these approaches. We call these broad approaches 'passive' and 'active'. Passive approaches characterize interactions using laws or some counterfactual notion or other – either relatively simple counterfactuals, or their more sophisticated cousin

the invariance relation. They then use either a best-system laws grounding for such counterfactual claims, or a modally realist grounding.[25] We call this approach 'passive' because broadly the grounding for counterfactual claims is just patterns of the objects in this or in other worlds.

Passive approaches contrast with active approaches. These latter approaches give an account of interactions in terms of the capacities, powers or activities of entities. So active approaches include Machamer, Darden and Craver's activities approach, where activities are varied things that entities can engage in, like bonding, breaking, pushing and coiling. Also included are Nancy Cartwright's capacities approach, which claims that capacities are properties of objects, and Carl Gillett's powers approach, which follows Shoemaker in individuating properties by their causal powers, so that having a property implies action in certain conditions.[26] We will argue that only active approaches give a local characterization of a mechanism.

We take the points we make about the metaphysical systems to be well-understood aspects of these systems. They also have well-known problems, which we shall not repeat here. Nevertheless the significance of these points for thinking about the metaphysics of mechanisms needs to be spelled out. Dialectically, this is interesting to the mechanisms literature because – outside of traditional metaphysics – passive approaches are generally regarded as metaphysically less problematic than active approaches. This yields a dialectical reason for philosophers working on philosophy of science, but not inclined to do traditional metaphysics, to plump for passive characterizations of interactions in mechanisms. In the core mechanisms literature, only Machamer, Darden and Craver are trying to use an active approach to characterize interactions in mechanisms, but the novelty and apparent extremity of their approach is off-putting to many. Recent work by Gillett using a powers approach is interesting too, but has yet to be picked up extensively in the mechanisms literature. We wish to oppose this dialectical trend towards passive approaches by pointing out that there are well-worked-out metaphysical systems available that do a better job for mechanisms than passive approaches.

Using some counterfactual account or other of the interactions in a mechanism has been popular. This is the approach of Woodward, Psillos and later Glennan. We will raise concerns about the status of this claim later on, but suppose for now that it is a claim about the nature of mechanisms.

[25] These familiar ideas derive from Ramsey, through significant development by David Lewis. Ramsey discusses the possibility of us systematizing our knowledge in a deductive system, suggesting that the laws are the axioms of such a system. See Ramsey (1990, p. 143). We focus on the best-system laws grounding for counterfactuals as it is Stathis Psillos' explicit view (see Psillos 2004), and include modal realism as the major alternative.

[26] See Cartwright (1989), Machamer *et al.* (2000) and Gillett (2006). There is considerable variation within the active tradition, with Machamer, Darden and Craver's essentially dynamic activities the most 'active'. We will come to some of these distinctions later.

Whether such a view yields real, local mechanisms or not depends on what you take to ground counterfactuals. Take the passive views first. Suppose the truth-conditions for counterfactuals – whether simple or sophisticated – are grounded in a best-system account of laws of nature, where the best-system is judged by the simplicity and strength of laws.[27] If this is so, then mechanisms are not local. Mechanisms depend on two kinds of non-local features. The first is features of many other places and times in this world, those necessary to determine the laws of nature. As we have said in the previous section, this is in tension with the actual practice of mechanistic explanation in the sciences, which examines only local regions of spacetime in constructing mechanistic explanations. The second kind of feature is the simplicity and strength of laws that establish what is the best system of laws. No such features are local, being dependent first on the entire universe, and second on the abstract concepts simplicity and strength, which cannot be located clearly at all. Such mechanisms might also fail to be real – depending on the status of laws. In general, an anti-realist account of laws clearly yields non-real mechanisms; while any passive realist account of laws still has truthmakers widespread in the universe for law, so mechanisms based on such laws will be non-local.[28]

Take the alternative modal realist account of truthmakers for counterfactuals, where the truth of any counterfactual claim depends on what happens to counterparts in nearby possible worlds.[29] Mechanisms involving such counterfactuals might well be real. But the situation regarding locality is worsened. This view makes the truth of counterfactual claims depend not only on what happens elsewhere in this world, but also on what happens in nearby possible worlds. This is the most radically non-local account of the interactions in mechanisms it is possible to have. In general, any passive metaphysical grounding for such counterfactual claims as part of mechanisms will yield non-local, and in some cases also non-real, mechanisms.[30]

[27] This is Psillos' view. He writes: 'The one [story] I favor ties the truth-conditions of counterfactual assertions to *laws of nature*.' He adds later: 'Laws are those regularities that are members of a coherent system of regularities, in particular, a system which can be represented as an ideal deductive axiomatic system striking a good balance between *simplicity* and *strength*.' (Both from Psillos 2004, p. 299).

[28] David Lewis seems committed to this, since he accepts the parallel implication for a laws-based account of causation: 'Like any regularity theory, the best-system analysis says that laws hold in virtue of patterns spread over all of space and time. If laws underlie causation, that means that we are wrong if we think, for instance, that the causal roles of my brain states here and now are an entirely local matter. That's an unpleasant surprise, but I'm prepared to bite the bullet.' (Lewis 1994, p. 479) Note that if the best-system view of laws is as a best system in a world consisting of four-dimensional spacetime, then mechanisms today will also depend on the future.

[29] No one in the mechanisms literature explicitly holds this view, to our knowledge. However, given the history of metaphysical theorizing about counterfactuals, it is too important an alternative to neglect.

[30] We do not discuss Woodward's test-conditions for the truth of counterfactuals because Woodward claims that his view is not a metaphysical one. See Woodward (2003). Nevertheless, if

The only prospect for a real and local metaphysics is some variety of active metaphysics, a metaphysics such as Machamer, Darden and Craver's activities, or Cartwright's capacities. We begin here with Cartwright's capacities view, since it is a more familiar real and local metaphysics. We will raise the issue of how far our arguments transfer to other active approaches later.

On the capacities approach, the interactions between parts of a mechanism are described in terms of the capacities of the entities in the mechanism. Cartwright holds that most general causal claims such as 'aspirins relieve headaches' or 'electromagnetic forces cause motions perpendicular to the line of action', are really ascriptions of capacities – the capacity to relieve headaches ascribed to aspirin, and the capacity to cause motions perpendicular to the line of action ascribed to electromagnetic forces. Capacities are properties, and their instances are real. [31] Cartwright allows for special science 'laws' of the kind we allow for – the rule of thumb, ceteris paribus generalizations that are produced in the special sciences; laws with exceptions that can be explained using a mechanism. But these laws arise out of the reasonably regular reactions of entities with similar capacities: the truthmakers for laws are capacities. She writes: 'It is not the laws which are fundamental, but rather the capacities. . . . Whatever associations occur in nature arise as a consequence of the actions of these more fundamental capacities.' (Cartwright 1989, p. 181)

Using capacities in an account of mechanisms allows what a mechanism is to be local. For Cartwright, the capacities of the entities in a mechanism are properties of the entities, not dependent on anything anywhere else in this world or any other. This nicely fits scientific practice since scientists in many domains spend a lot of time figuring out the capacities of the entities in mechanisms underlying the phenomena that are interesting to them. When scientists point to a mechanism and identify it as the mechanism responsible for certain phenomena, on this view they are pointing to something real and really *there*.

Cartwright ensures that her metaphysics is real and local. An ontology using entities and their powers is structurally similar to that of Cartwright. This approach might also yield real and local mechanisms. See the work of Carl Gillett (2006) and recent as yet unpublished work by Stephen Mumford, both using an active powers metaphysics in their approach to

you did take Woodward's position to be a metaphysical one, his invariance relations would be non-local, albeit more local than either a modal realist or best-system laws view. Invariance relations would still depend on what happened elsewhere and at other times in this world.

[31] Cartwright argues this extensively. See her (1989). She also argues against the possibility of describing away capacities in terms of regularities. She writes: 'One does not just say the acid and the base interact because they behave differently together from the way they behave separately; rather, we understand already a good deal about how the separate capacities work and why they should interfere with each other in just the way they do.' (Cartwright 1989, p. 165).

mechanisms.[32] Machamer, Darden and Craver are also trying to use an active approach as a good metaphysics for mechanisms, but the view remains as yet under-developed. We think it an interesting view worthy of the necessary development.[33]

These views are promising approaches to mechanisms, but note that not all active metaphysical approaches will do. Dispositional approaches are structurally similar to Cartwright's capacities approach and the powers approach, having a basic ontology of entities plus their dispositional properties, rather than entities and their capacities or powers. But the further detail of most dispositional approaches would create problems. First, many accounts of what dispositions are, are non-local. This would apply to either conditional or law-based accounts of dispositions. Second, the local approach to characterizing dispositions is the one which claims that dispositions are intrinsic properties of objects. This makes it an acceptable approach to those wanting a real and local metaphysics and willing to accept metaphysical claims about intrinsic properties, but we are not inclined to accept them. All science needs is clusters of capacities that stick around together for long enough and produce a phenomenon regularly enough for us to get interested in it and look for the mechanism. As Cartwright says of capacities: 'They do indeed endure; on the other hand, their characteristics may evolve naturally through time, and they may be changed in systematic, even predictable, ways as a consequence of other factors in nature with which they interact.' (1989, p. 157)

In conclusion, those wishing for a local and real metaphysics of mechanisms should not use counterfactual notions grounded in laws or other possible worlds in their characterization of mechanisms. There are alternative available metaphysics, along the capacities or active powers lines, which are real and local. These aspects of the capacities or active powers metaphysics are well-known in the core metaphysics literature, of course, but their existence seems to have been largely ignored in the mechanisms literature.

38.6 Objections: Laws, capacities and fundamental explainers

In this section we introduce two major objections to our line of argument. The first objection claims that capacities or powers cannot explain anything at

[32] One might hold that powers are a better prospect for the metaphysics of mechanisms than dispositions, since dispositions can be seen as structural properties of static objects, whereas powers are more dynamic. On this point, see also Machamer (2004). See Illari and Williamson (forthcoming) for further discussion of the dynamic nature of activities.

[33] We argue elsewhere that Machamer, Darden and Craver's activities-entities dualism compares well to Cartwright's entities-capacities ontology, on various criteria for ontology. See Illari and Williamson forthcoming.

all. We argue that they can, and do a job more suited to the special sciences use of mechanisms than laws or counterfactuals. The second objection is an argument attempting to show that on laws or counterfactuals-based stories, mechanisms can still be local. We argue that whether this succeeds depends on precisely what you are claiming in characterizing mechanisms. On one of the cases we identify, there is no serious disagreement, on the other account we argue that mechanisms remain non-local.

38.6.1 Capacities cannot explain

Recall that the major reason for introducing mechanisms is to explain. The most uncontroversial claim about mechanisms is that they explain, whether in a physical or in an epistemic way. It might be objected to our arguments that positing an active metaphysics such as capacities or powers as explainers, particularly as fundamental explainers, is illegitimate, because capacities and powers do not explain anything. To say A produced B because it has a capacity or power to produce B, or engaged in the activity which brings about B, explains nothing – it might be thought a mere assertion of 'dormitive virtue'.

The first thing to notice in this debate between capacities or powers, and laws or counterfactuals, is that a parallel complaint can be made about laws or counterfactuals. To say that A produced B because there is a law that A produces B might also be thought to explain nothing. The only 'explanation' offered by a law is the recognition that things just tend to happen that way, that things like A tend to produce things like B. Intuitions on whether to prefer something like capacities or something like laws as fundamental explainers do seem to vary.

There is something odd in considering a law or a capacity, alone, as a complete explanation. The explanations we do get tell us so much more. Consider the explanation of how the cell produces proteins. Neither the claim that it has a capacity to do so, nor the claim that cells like it also produce proteins, tell us much. *Mechanisms* explain in terms of lower-level entities and their capacities, powers, activities or some such item of an active metaphysics. Mechanisms as a whole are neither just capacities, nor just laws. Mechanistic explanation generally starts with a regularity: the identification of a phenomenon requiring explanation – usually a regularly occurring phenomenon. In the case of protein synthesis, distinguishing between kinds of proteins produced, and so further dividing the explanandum phenomenon is important. Mechanistic explanation then proceeds by identifying the parts that make up the phenomenon – the production of each protein, and what those parts do, and can do under similar circumstances. To see this as the identification of the entities present, and the capacities or powers that those entities have, seems natural. Lawlike regularities can be useful in describing mechanisms, but as we have explained, this is consistent with an active metaphysics account of mechanisms. On this view lawlike regularities are not fundamental. For those

influenced by the widespread concern of those in the mechanisms literature that there are no special science laws, which we identified as a key reason for turning to consideration of mechanisms, this mechanistic view of explanation is far superior.

38.6.2 Mechanisms using laws or counterfactuals can still be local

There may be a deeper objection to our argument that using passive grounds for counterfactual notions to characterize mechanisms renders them non-local. This is as follows: the status of mechanisms as mechanisms depends only on the natural properties that the mechanisms have. Although these natural properties depend on the laws of nature, or other possible worlds, nevertheless the natural properties are local. Thus, the mechanisms themselves are local. The broad idea seems to be to push the underlying non-local metaphysics into the background, and insist that in general, say, protein synthesis depends only on the natural properties of the molecules in the cell, while natural selection depends only on the natural properties in the population and its local environment. This, the claim would go, yields the required locality in spite of the fact that these natural properties only produce anything or interact with anything in virtue of things widely spread in time and space. Thus mechanisms are constitutively local since what makes a mechanism a mechanism is these natural properties, the laws being merely some kind of inert background conditions.

This is an interesting possibility, which raises the issue of what a characterization of mechanisms is intended to do. Over the course of this chapter we have identified a number of differences in approaches to mechanisms. A genuine possibility here is that some in the debate are not concerned to make any claims about the nature of mechanisms – that is, no claims that really impinge on the metaphysics of mechanisms.[34] This might be a way to defend Woodward, Psillos and Glennan. Perhaps they are merely trying to give an account of mechanisms that will let you *pick them out*, so that you can discriminate mechanisms from non-mechanisms – rather than illuminate what they *actually are*. Perhaps the best characterisation uses counterfactuals or invariance relations because they best let you pick out the mechanisms.

It is possible to read both Woodward and Glennan as merely characterizing how you pick out mechanisms. Woodward's paper is titled, 'What is a mechanism? A counterfactual account,' but his abstract summarizing his argument is less clear. He writes: 'This paper presents a counterfactual account of what a mechanism is. Mechanisms consist of parts, *the behavior of which conforms to generalizations that are invariant under interventions*, and which are modular in the sense that it is possible in principle to change the behavior of one part independently of the others. Each of these features can be captured by

[34] This might be partially due to their primary concern being with epistemic explanation.

the truth of certain counterfactuals.' (Woodward 2002, S366, emphasis added) Perhaps he is only claiming something about what the parts of mechanisms typically do, not what they are. Glennan is also open to this interpretation. He writes: ' "Interaction" is a causal notion that *must be understood in terms of* the truth of certain counterfactuals. The stipulation that these interactions can be characterized by invariant, change-relating generalizations is meant to capture the relevant counterfactual truth claims.' (Glennan 2002, S344, emphasis added) If you read 'must be understood in terms of' fairly lightly, this may not be a claim about a deeper metaphysics.

If this is indeed the aim Woodward, Psillos and Glennan have in mind, then their accounts would fit that aim. Mechanisms do typically exhibit a stability which can very naturally be characterized using various counterfactual notions. If this is the claim then it is perfectly legitimate to say that a counterfactual characterization using a laws or modal realist grounding for counterfactuals is compatible with a local deeper metaphysics. A characterization of this sort says nothing about metaphysics. When serious questions about the nature of mechanisms arise, the non-local metaphysics used merely to pick out mechanisms can fade quietly into the background.

That is one possibility. However, it seems that Woodward, Psillos and Glennan all have at least some intention of arguing for a claim about the nature of mechanisms. In some form or other, they all argue that some counterfactual notion is essential or ineliminable in characterizing mechanisms. This might well be taken as a stronger claim about the nature of mechanisms, beyond any claim about a handy way to pick them out. The quotes from Woodward and Glennan above are certainly open to this stronger reading. Psillos must be read in this stronger way.[35] He argues extensively that counterfactuals are indispensable to a characterization of mechanisms, repeating this kind of claim at several points in his paper. He summarizes the thesis he has argued for towards the end of the paper: 'mechanisms need counterfactuals; but counterfactuals do not need mechanisms. In other words, mechanistic causation requires counterfactual dependence but not conversely. It is in this sense, that the counterfactual approach is more basic than the mechanistic.' (Psillos 2004, p. 315) To claim that the counterfactual approach is more basic than the mechanistic does look like a metaphysical claim in this chapter.

If the claims of any of the three are claims about the basic – ineliminably basic – metaphysics for mechanisms, then it is a claim about the nature of mechanisms. It is not clear how such a claim can fade into the background to allow mechanisms to be considered constitutively local. It is claimed that the nature of mechanisms is to have the natural properties that they do lead to the interactions that they do only in virtue of the laws of nature, or facts about

[35] Note that Psillos could defend himself by retreating into a general anti-realism. But then we think his general anti-realist views will support his views on mechanisms, not vice versa.

other possible worlds. But then what mechanisms *are* is non-local. Depending on the detail of the further claims, mechanisms might also turn out to be non-real.

Another thought is that Psillos, Woodward and Glennan are in various ways concerned with understanding causation, as well as mechanisms. Perhaps the claim that a non-local metaphysics doesn't make *causal* claims non-local is plausible in a way that the parallel claim for mechanisms is not. In considering causal claims, we immediately focus on salient causes, and standardly assume vast amounts of stable background conditions. Perhaps in this context the claim that natural properties are local, treating a passive metaphysics as background, is not unreasonable. But in the context of physical mechanistic explanation – not *epistemic* mechanistic explanation – we are not in the same situation. For mechanisms, the entire structural background is crucial to what a mechanism is. On this metaphysical view, such mechanisms are non-local.

Perhaps the objection discussed in this section could be read as intending merely to deflect the counterintuitiveness of mechanistic explanation turning out to be non-local.[36] If this is so, it seems to fail. Non-local mechanistic explanations of this sort are committed to the core claim that the behaviour of other people like me in this world, or my counterparts in other possible worlds, is in some way relevant to explaining why I bump into lampposts. This is precisely the counterintuitive claim that we deny. A mechanistic explanation of my clumsiness depends only on facts about me. There is a metaphysics available without such a counterintuitive consequence – a metaphysics of capacities, powers or activities, which should be preferred.

In conclusion, capacities can explain and mechanisms are still non-local on any claim that counterfactual notions are part of their nature, their metaphysics. The active metaphysics of capacities or powers clearly comes out better, thoroughly satisfying both locality and reality.

38.7 Conclusion

We have argued extensively that both scientific practice and much of the mechanisms literature is committed to mechanisms being both real and local. We further argued that if mechanisms are to be real and local, so that they can be used in physical explanations of phenomena, in a form distinct from laws-based explanation, they require an active metaphysics such as Cartwright's capacities approach, a powers approach, or an activities approach.

We have framed all our arguments here about mechanistic explanation. We believe that exploring mechanistic explanation will be illuminating to meta-

[36] We thank an anonymous reviewer for raising this possibility.

physical debates about causation, but not necessarily in a simple way, meaning that it is best to understand mechanistic explanation thoroughly, and then go on explicitly to consider its relation to causation and causal explanation. For those interested in scientific methodology more than metaphysical debates, it is probably sufficient to note that scientists involved in mechanistic explanation of the sort we describe see themselves as straightforwardly involved in causal explanation. See for example the work of an evolutionary biologist: 'The main purpose of evolutionary biology is to provide a rational explanation for the extraordinarily complex and intricate organization of living things. To explain means *to identify a mechanism that causes evolution and to demonstrate the consequences of its operation.*' (Bell 1997, p. 1, emphasis added.) Or consider the view of a biochemist: 'Uncovering the cellular mechanisms *resulting in sequential transfer* of information from DNA (our genes) to RNA and then to protein represents one of major achievements of biochemistry in the 20th century.' (Whitford 2005, p. 247, emphasis added) Both clearly see themselves as investigating causes. From this point of view, all our arguments apply straightforwardly to this important variety of causal explanation.

Acknowledgements

We would like to thank the Leverhulme Trust for funding this work on mechanisms. We are also grateful to three anonymous referees for extensive comments leading to improvement of the work. We would also like to thank numerous colleagues at Kent, Bristol, and internationally for discussion of ideas used here. Remaining errors are, of course, our own.

References

Peter Achinstein (1983). *The Nature of Explanation*, OUP, Oxford.

Roger L. Adams, John T. Knowler and David P. Leader (1992). *The Biochemistry of the Nucleic Acids* (11th edn), Chapman and Hall, London.

William Bechtel and Adele Abrahamsen (2005). Explanation: A mechanist alternative, in *Studies in the History and Philosophy of the Biological and Biomedical Sciences*, 36, 421–41.

William Bechtel and Adele Abrahamsen (2008). From reduction back to higher levels, In B.C. Love, K. McRae and V.M. Sloutsky (eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 559–64), Austin, TX, Cognitive Science Society.

Helen Beebee (2004). Causing and nothingness, in L.A. Paul, E.J. Hall and J. Collins (eds.) *Causation and Counterfactuals*, MIT Press, Cambridge, MA, pp. 291–308.

Graham Bell (1997). *Selection: The Mechanism of Evolution*, Chapman and Hall, New York.

Ned Block (2005). Two neural correlates of consciousness, in *Trends in Cognitive Sciences*, Vol. 9, 2, 46–52.

Nancy D. Cartwright (1983). *How the Laws of Physics Lie*, Clarendon Press, Oxford.

Nancy D. Cartwright (1989). *Nature's Capacities and their Measurement*, Clarendon Press, Oxford.

Nancy D. Cartwright (2006). Evidence-based policy: Where is our theory of evidence?, in *Selected Proceedings of the German Analytic Philosophy Conference*, Berlin, September 2006.

Nancy D. Cartwright (2009). 'Causality, invariance and policy', in Harold Kincaid and Don Ross (eds.) *The Oxford Handbook of Philosophy of Economics*. Oxford, OUP, pp. 410–23.

Carl Craver (2007). *Explaining the Brain*, Clarendon Press, Oxford.

Lindley Darden (2006). *Reasoning in Biological Discoveries*, CUP, Cambridge.

Phil Dowe (2004). Causes are physically connected to their effects: Why preventers and omissions are not causes, in C. Hitchcock (ed.) *Contemporary debates in Philosophy of Science*, Blackwell, Oxford.

Bas van Fraassen (1980). *The Scientific Image*, Clarendon Press, Oxford.

Carl Gillett (2006). The metaphysics of mechanisms and the challenge of the new reductionism, in M. Schouten & H.L. de Jong (eds.) *The Matter of the Mind*, Blackwell, Oxford.

Stuart Glennan (2002). Rethinking mechanistic explanation, in *Philosophy of Science*, 69 (September 2002), S342–S353.

Ned Hall (2004). Two concepts of causation, in L.A. Paul, E.J. Hall and J. Collins (eds.) *Causation and Counterfactuals*, MIT Press, Cambridge, MA, pp. 225–76.

Phyllis McKay Illari and Jon Williamson (forthcoming) 'In defense of activities'.

Phyllis McKay Illari and Jon Williamson (2010). 'Function and organization: Comparing the mechanisms of protein synthesis and natural selection' in *Studies in the History and Philosophy of the Biological and Biomedical Sciences*, 41, 279–291.

Bert Leuridan (2010). Can mechanisms really replace laws of nature?, in *Philosophy of Science* 77, pp. 317–340.

Bert Leuridan and Erik Weber (2011). The IARC, and mechanistic evidence, in Phyllis McKay Illari, Federica Russo and Jon Williamson (eds.) *Causality in the Sciences*; Oxford, Oxford University Press.

David Lewis (1994). Humean supervenience debugged, *Mind*, 412, 471–90.

David Lewis (2004). Void and object, in L.A. Paul, E.J. Hall and J. Collins (eds.) *Causation and Counterfactuals*, MIT Press, Cambridge, MA, pp. 227–90.

Peter Machamer, Lindley Darden and Carl Craver (2000). Thinking about mechanisms, *Philosophy of Science* 67 (March 2000) 1–25.

Peter Machamer (2004). Activities and causation: The metaphysics and epistemology of mechanisms, *International Studies in the Philosophy of Science*, 18:1 (March 2004) 27–39.

Richard Menary (2010). *The Extended Mind*, Ashgate MIT Press, Cambridge, Massachusetts.

Sandra D. Mitchell (1997). Pragmatic laws, *Philosophy of Science*, 64(4), supplement, S468–S479.

Stathis Psillos (2004). A glimpse of the secret connexion: Harmonizing mechanisms with counterfactuals, *Perspectives on Science*, 12(3), 288–319.

Frank P. Ramsey (1990). Law and causality, in D.H. Mellor (ed.), *FP Ramsey: Philosophical Papers*, CUP, Cambridge.

Mark Ridley (2004). *Evolution* (3rd edn). Blackwell, Oxford.

Wesley C. Salmon (1998a). Comets, pollen, and dreams: Some reflections on scientific explanation, in his *Causality and Explanation*, OUP, Oxford, pp. 50–67.

Wesley C. Salmon (1998b). Scientific explanation: Three basic conceptions, in his *Causality and Explanation*, OUP, Oxford, pp. 320–32.

Jonathan Schaffer (2004). Causes need not be physically connected to their effects: The case for negative causation, in C. Hitchcock (ed.), *Contemporary debates in Philosophy of Science*, Blackwell, Oxford.

Robert A. Skipper (Jr.) and Roberta L. Millstein (2005). Thinking about evolutionary mechanisms: Natural selection, *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36, 327–347.

Phillip J. Torres (2009). A modified conception of mechanisms, in *Erkenntnis* DOI 10.1007/s10670-008-9125-y. 71, 233–51.

Donald Voet and Judith G. Voet (2004). *Biochemistry*, John Wiley, New York.

David Whitford (2005). *Proteins: Structure and Function*, John Wiley, Sussex.

Robert A. Wilson (2004). *Boundaries of the Mind*, CUP, Cambridge.

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

James Woodward (2003). *Making Things Happen: A Theory of Causal Explanation*, OUP, Oxford.