

Forthcoming in *British Journal for Philosophy of Science*
Preprint (not copyedited or formatted)

Title: Putting History Back into Mechanisms

Author:

Justin Garson
Department of Philosophy
Hunter College and The Graduate Center, City University of New York

Abstract: Mechanisms, in the prominent biological sense of the term, are historical entities. That is, whether or not something is a mechanism for something depends on its history. Put differently, while your spontaneously-generated molecule-for-molecule double has a heart, and its heart pumps blood around its body, its heart does not have a mechanism for pumping, since it does not have the right history. My argument for this claim is that mechanisms have proper functions; proper functions are historical entities; so, mechanisms are historical entities, too. This thesis runs against the mainstream new mechanist way of thinking about mechanisms, where mechanisms are generally thought of in an ahistorical way. After arguing for this thesis, I draw out some consequences for philosophy of science and metaphysics.

Acknowledgements: I'm very grateful to the participants of the Rutgers-Seton Hall Workshop for Compositional Explanation in Biology and the Neurosciences, held in Newark on Oct. 25-26, 2019, for their valuable feedback and criticism: Ken Aizawa, Holly Anderson, Michael Baumgartner, Mark Couch, Alexander Gebharder, Angela Potochnik, and Mark Povich. I also want to thank Lindley Darden, Kal Kalewold, and Karen Neander for discussion of the paper's ideas. Finally, I thank the journal's reviewers for suggesting valuable points of clarification.

1. Introduction

Here, I'll argue that mechanisms – in the prominent biological sense of the term – are historical entities.¹ That is, whether or not something is a mechanism for blood clotting, or dopamine homeostasis, or spatial memory, depends on its history. Put crudely, your spontaneously generated molecule-for-molecule duplicate would have a heart, and its heart would pump blood around its body just as efficiently as yours does, but its heart wouldn't have a mechanism for pumping blood. It wouldn't have a mechanism for anything, because it wouldn't have the right history. In this way, mechanisms take their rightful place among other notable historical entities in science, such as siblings, adaptations, volcanic mountains, and igneous rocks. (I take it for granted that a mechanism is always a mechanism-for-X; there's no such thing as a bare mechanism.)

Putting the thesis differently: mechanisms are essentially historical. Their essences are, at least in part, historical rather than intrinsic.² Being a mechanism is more like being an original Van Gogh than a chunk of gold. Whether or not something is a chunk of gold depends, at least on the standard philosophers' view, only on present-day facts about its atomic composition. Whether or not the painting hanging in front of you is an original

¹ By using the term 'entity' I don't mean to prejudge, at the outset, the question of whether mechanisms are systems rather than processes; I use 'entity' in a sweeping way for any ontological category, whether systems or processes.

² See Devitt ([2008]) for a recent overview of this distinction.

Van Gogh depends on whether a causal trajectory extends back in time from that painting to Van Gogh's nervous system.

My argument for this thesis is simple. Mechanisms, in the sense that will be clarified shortly, have proper functions. That is, having a proper function is not incidental to being a mechanism; it is constitutive of it. But proper functions are historical entities. So, mechanisms are historical entities, too. My argument has the same structure as: 'All uncles are siblings; whether or not you're a sibling depends on having the right history; so, whether or not you're an uncle depends on having the right history, too.' Establishing my argument is simply a matter of demonstrating the first premise, demonstrating the second premise, and showing that there's no hidden equivocation amongst them.

I want to be clear about what I'm arguing and what I'm not. I think there are at least two main senses of 'mechanism' in science, and probably more (see next section). The first sense of 'mechanism' appears to play an especially prominent role in biology – though it need not be restricted to biology – and it has sometimes, though inappropriately, been likened to a 'machine.' A mechanism in this sense has various parts; these parts do various things, and when these parts and their doings are aligned like *so*, they yield their phenomenon with a certain degree of regularity. But far more important, for my purposes, is that in this sense of 'mechanism,' mechanisms are the sorts of things that can *break, fail, or go awry*. This is the sense at play when we ask after mechanisms for blood circulation, cellular respiration, and wildebeest migration. It also plays into cognitive neuroscience and even evolutionary approaches to psychology, when we ask after

mechanisms for snake detection, parental investment, or male jealousy. This is the sense that philosophers of biology, in the main, have sought to explicate for the last few decades.

There is a second sense of mechanism, however, a quite thin sense, where ‘mechanism’ means nothing more than a mere causal trajectory. What is the mechanism by which the baseball shattered the window? What is the mechanism, that is, the sequence of events linked together by cause and effect, that culminated in the death of Roland Barthes? My claim is that mechanisms in the first of these senses, in the prominent biological sense, are historical.

At least two main consequences stem from contemplating the historical character of mechanisms. The first is epistemic. How do we discover whether or not something is a mechanism, and if so, what it is a mechanism for? Craver and Darden ([2013], Chapter 5) tell us that in order to decide whether something is a mechanism, and what it is a mechanism for, we merely need to poke and prod at it for long enough: observe its behavior, take it apart, put it back together, strategically break it, and so on. In principle, this is all we’d ever have to do. We wouldn’t have to know anything about its history, for example, where it came from or what it was made for. That, I claim, is a mistake.

Whether or not something is a mechanism, and what it is a mechanism for, depend partly on its history. Sometimes, to settle the question of whether something is a mechanism, or what it’s a mechanism for, we must investigate its past: where it came from and how it got there.

Although I'll return to the point later, I want to flesh it out to avoid a potential objection. Just because a property is historical, one might argue, that doesn't mean one must *always* look to history to settle the question, in any given case, of whether it's instantiated.³ (In order to decide that two men are siblings, I don't need to demand their birth certificates; it's enough to look at them carefully.) What it means, rather, is that sometimes, history is unavoidable. Consider adaptations. For something to be an adaptation, it must have been shaped by natural selection for a certain role. Now, suppose you and I disagree about whether a certain trait is an adaptation, or what it's an adaptation for, such as eyespots on butterfly wings, or the inverted mammalian retina. I suppose that, often enough, we could settle our disagreement without having to look at history. For example, to establish that eyespots are adaptations for deflecting attacks away from vital organs, rather than mimicking predators, it would probably be enough to observe butterflies closely in the field, or fiddle with them in the lab, without giving a thought to their history.⁴ But suppose that careful observation revealed that the eyespots did both of those things, to a limited degree, and we couldn't agree about which of the two (if either) is the adaptation and which the byproduct. In that case, we'd have to study its history carefully to settle the issue. We'd need to consider the kinds of selection pressures butterflies confronted in their actual historical setting. And this is, in fact, what biologists often do when faced

³ An anonymous reviewer used a colorful example to make the point: 'I think I know decisively that my colleagues are human without considering their histories, even though to be human requires one to be the descendant of a particular lineage containing other human beings.'

⁴ See Prudic et al. ([2015]). Also, see Kröger and Biehlmaier ([2009]) on the problem of whether the inversion of the retina, so often considered a design flaw, is an adaptation.

with comparable disagreements.⁵ If there's any epistemological upshot to something's being a historical property, it has to be this: there are some circumstances in which we're forced to examine an object's history in order to decide whether or not it has the property in question.

Here is a second, metaphysical, consequence: there has been a trend, in what we can call the metaphysics of science, to rely heavily upon concepts and principles drawn from mechanistic reasoning in biology to think about constitutive explanation in science quite generally, and the constitutive relations to which they refer, for example, the relationship between the solubility of salt and its polar molecular structure.⁶ If I'm right, this move is overly hasty. That is because, whether or not something is a mechanism depends on its history, but the relationships that underlie constitutive explanations do not depend – or do not always depend – on history.⁷ They are, as a rule, ahistorical. Whether or not the polar molecular structure of salt explains salt's solubility doesn't depend on salt's history. It depends only on facts that obtain in the here and now, in conjunction with the timeless laws of physics.⁸ That suggests that we should be cautious about using concepts and explanatory norms drawn from reasoning about mechanisms in biology as a pattern for reasoning about constitutive explanations in science generally.

⁵ Consider Caro et al. ([2014]), which uses historical data on Tsetse fly distributions to show that zebra stripes are probably adaptations for fly-deterrence, not camouflage.

⁶ Craver ([2007], pp. 139-60) is a pivotal source; see Couch ([2011]) and Aizawa ([2016]) for excellent discussions.

⁷ I thank Angela Potochnik for helping me think clearly about this point.

⁸ An anonymous reviewer raised the prospect that at least some constitutive relationships are historical: 'Why do we count the stomach as part of the organism, but the food in the stomach as not part of the organism? Presumably something about persistence, continuing causal role, and proper function.'

Before moving on, a word about the paper's title is in order. Why putting history 'back' into mechanisms? Was there a time when theorists of biology thought of mechanisms in this historical way, and then a time when they stopped thinking of them this way? I cannot speak for biology as a whole, but I am struck by the fact that George Williams ([1966]), in his landmark *Adaptation and Natural Selection*, uses 'mechanism' in an overtly historical way: 'The designation of something as a *means* or *mechanism* for a certain *goal* or *function* or *purpose* will imply that the machinery involved was fashioned by selection for the goal attributed to it (p. 9; his emphases).' For Williams, a mechanism is something that essentially points to a history of 'fashioning by' selection; something is not a mechanism unless it was so fashioned.

Moreover, Williams' comment is no mere slip of the pen. Throughout the book, he asks rhetorical questions about whether such-and-such should be considered a mechanism. For example, he asks rhetorically whether the paws of a fox should be considered a 'mechanism' for making tracks through the snow (p. 13), or whether an earthworm's activities constitute a 'soil improvement mechanism (p. 18).' He thinks it's obvious that the answer is *no*: after all, making tracks through snow isn't what the fox's paws were made for!

This historical use of 'mechanism' is also part and parcel of much of current evolutionary psychology, and evolutionary approaches to the mind generally. When evolutionary psychologists ask, 'is there a mechanism for male jealousy?' they are not asking whether

there is a neural or psychological basis for jealousy in men. They are asking whether there's an evolved adaptation for male jealousy (see Buss [2008], p. 69; Tooby and Cosmides [2006], p. 185). They are asking, in the first place, a question about history: how did we come to be the way that we are?

To clarify, I will not argue that mechanisms are adaptations. I will argue for the weaker claim that whether or not something is a mechanism, and what it is a mechanism for, depend on having the right history. But I am agnostic, here, about what precisely that history must be. As we'll see in Section 4, the sort of history that one must look to in order to decide whether something is a mechanism depends on which theory of proper function is correct.

2. Some Distinctions

Mechanisms have functions; functions are historical entities; so, mechanisms are historical entities, too. That is, roughly, my core argument. But 'mechanism' can mean different things to scientists. 'Function' can mean different things to scientists, too. For example, there are mathematical 'functions;' there are biological 'functions;' there are 'functions' in the sense of mere effects. To get this argument off the ground, I must make some preliminary and, I hope, relatively uncontroversial distinctions. My goal is to fix the reference of my terms clearly enough for my argument to be properly stated and assessed. It is not to resolve the numerous and rather subtle debates about the metaphysics of

mechanisms or functions. Nor is my goal to introduce any novel distinctions. Rather, I purport to summarize an existing body of thinking about functions and mechanisms.

First, function. By ‘function,’ I always mean *proper function*. This is the sense of ‘function’ that is typically at play when biologists say things like, ‘the function of the eyespots on butterfly wings is to deflect attack away from vital organs,’ or, ‘the function of zebra stripes is to deter biting flies,’ or ‘a function of dopamine release by the VTA is to reinforce behaviors associated with an unexpected reward.’ The notion of a proper function is associated with three other ideas, though the exact manner of this association remains controversial (Garson [2016]). First, it is associated with utility. A trait’s function is, as a rule, useful or good for the organism. Second, it is associated with teleology, and specifically, with teleological explanation. At least sometimes, when biologists attribute a proper function to a trait, they are trying to explain why the trait exists, and they do so by citing one of its effects. (When a biologist says, ‘the function of the eyespots on butterfly wings is to deflect attack away from vital organs,’ she is, often enough, trying to explain why butterfly wings have eyespots.) Finally, proper functions are ‘normative.’ All I mean by ‘normative’ is that something with a function can be dysfunctional or malfunctioning. Something like failure of function is possible.

Crucially, my goal here is not to articulate a *theory* of what proper functions are. There are many such theories: the selected effects theory, the fitness-contribution theory, the causal role theory, and others. The point here is to distinguish that sense of ‘function,’ the proper function sense, whatever it ultimately amounts to, from another sense of

‘function.’ According to this other sense of function, ‘function’ just means ‘effect.’ A function of deforestation is climate change. A function of malnutrition is poor academic performance. Neander ([2017], pp. 56-8) recently calls the latter sort of functions ‘minimal functions’ to distinguish them from her ‘normal-proper’ functions. A chief difference between proper functions, and these ‘minimal’ functions, is that proper functions are the sorts of things that can fail. (If a malnourished child ends up getting perfect grades in school, we would not say that there has been some sort of dysfunction or malfunction.) In Section 4, I will make the case that these proper functions are historical entities.

Commentators have long pointed out that the term ‘mechanism,’ like the term ‘function,’ is also ambiguous. One sense of ‘mechanism’ seems to play a very special role in biology; these mechanisms are often likened to machines. But ‘mechanism’ is sometimes used in a much more general, even tepid, sense of the term: it means nothing more than a cause-and-effect sequence. This latter sense of ‘mechanism’ is ubiquitous throughout the natural sciences, including biology.

Consider the first sense, which appears to be special to biology – or, to be more cautious, its paradigmatic exemplars are drawn from the realm of biology, even if its domain is broader. In this sense, we can ask after the mechanisms of blood pumping, or the mechanisms of the silverfish’s wiggling movement, or the mechanism for detecting interaural time disparity. This sense of mechanism has sometimes, inappropriately, been likened to a machine. The problem with the machine metaphor is that it fails to capture

the dynamic, process-like nature of mechanisms.⁹ As Glennan ([2017], pp. 25-7) points out, mechanisms are processes, not systems. To say that a mechanism is a process is to emphasize that it is composed of distinct temporal stages. While there is a mechanism for digestion, that mechanism is not, strictly speaking, identical to the stomach. The mechanism is composed of parts of the stomach whose regular patterns of interaction often result in digestion. Here, I take this idea – that mechanisms are processes, not systems – as foundational.

To be sure, as Glennan notes, it's natural in some contexts to equate *systems* with mechanisms: 'the stomach is a mechanism for digestion.' But this is elliptical. Such ways of speaking are harmless enough, so long as we remember that the stomach isn't a mechanism for digestion, but the mechanism for digestion *acts in* the stomach, or the stomach *implements* this mechanism. Krickel ([2018]) describes mechanisms as 'entity-involving occurrents' to capture their active, process-like nature.

Still, the machine metaphor isn't entirely otiose. It has a kernel of truth, which probably accounts for its continuing attraction. Like working machines, mechanisms in biology have components or parts; these parts each have their special activities; when these parts, with their special activities, are arranged in the right way, when they exhibit just the right sort of spatial, temporal, dynamic, relationships, they yield their proprietary phenomena.

⁹ Darden ([2006], p. 280) notes that there are plenty of machines that are not mechanisms: a stopped clock is a machine, not a mechanism.

In this sense, most working machines, as well as many biological organs, implement one or more mechanisms.

More important for my argument, however, is that in this sense of ‘mechanism,’ mechanisms are the kinds of things that can *break, fail, or go awry*. When we read that Alzheimer’s disease results from ‘potentially irreversible *impairments* of synaptic memory mechanisms’ (Rowan et al. [2003], p. 821; my emphasis), we understand perfectly well what is meant. We might disagree about what it is for a mechanism to ‘break,’ rather than to ‘operate normally,’ but we immediately grasp the distinction between a mechanism that is broken and one that’s not. In fact, a standard method for figuring out how mechanisms work is to break them and see what happens (Bechtel and Richardson [2010/1993], p. xxviii; see also Craver [2001], p. 72; Glennan [2005], p. 448; Craver and Darden [2013], p. 193).

This first sense of ‘mechanism’ appears to be one that philosophers in the ‘new mechanism’ tradition, as a rule, have been intent on clarifying. Consider, for example, Wimsatt ([1976]), an early paper in this area which is routinely cited in the new mechanism literature: ‘*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 (p. 671; my emphasis).’ What is noteworthy about this passage is that Wimsatt is making an observation, first and foremost, about what

biologists take themselves to be doing, rather than physicists or chemists. Biologists are in the business of discovering mechanisms.

Similarly, MDC ([2000]) – that landmark paper in the new mechanism tradition – also suggests that mechanisms have a special role to play in biology. After emphasizing the importance of mechanisms for the philosophy of science, they write: ‘In fact, if one does not think about mechanisms, one cannot understand neurobiology and molecular biology (p. 24).’ This passage suggests, but does not strictly imply, that neurobiology and molecular biology rely on mechanisms in a very special way, in a way that perhaps chemistry or sociology or astrophysics does not. Bechtel and Richardson’s ([2010/1993]) book, *Discovering Complexity*, is also about research heuristics in biology and, to some extent, cognitive neuroscience. Recently, Kaiser and Krickel ([2017], p. 747) noted that philosophical work on ‘constitutive mechanistic explanation’ draws almost *entirely* on biological or neurocognitive examples: ‘the human heart pumping blood...a cell synthesizing proteins...long-term potentiation at synapses of neurons...spatial memory of humans...and cellular respiration.’

What, then, is this *other* sense of ‘mechanism,’ the sense that is somehow analogous to Neander’s ‘minimal functions?’ This sense of ‘mechanism’ is invoked throughout the sciences, not just in biology, and it seems to mean little more than a *mere causal trajectory*. Here, mechanism-talk is an attempt to explicate causation-talk, wherever it happens to arise. (Consider the statement, ‘I don’t believe in the efficacy of intercessory prayer, since I can’t imagine a mechanism for it.’) Moss ([2012], p. 165) calls this sense

of ‘mechanism,’ mechanism as susceptibility to explanation in a general sense: ‘To refer to a biological event or phenomenon in terms of its “mechanism” is also, however loosely, to categorize it as being physically plausible, that is, as being within the ambit of empirical investigation.’ This is not a sense of ‘mechanism’ that is in any way special to biology; it is ubiquitous throughout the natural sciences.

This appears to be the sense of ‘mechanism’ that Glennan, in particular, has taken pains to explicate throughout his career. For example, in his classic 1996 paper, the concept of mechanism is intended to clarify the very idea of causation. One of his choice examples is the relation between turning a key and starting an engine: ‘I am justified in asserting “If I were to turn the key, the car would start” because I know that there is a mechanism which connects key-turning with car-starting’ (p. 63). More recently, Glennan ([2010]) writes of ‘ephemeral mechanisms;’ in this regard, he tells us that, ‘it is not a stretch to use the term “mechanism” in connection with the explanation of singular occurrences,’ and he goes on to use, as an example of such an ephemeral mechanism, the sequence of events leading to the death of French literary critic Roland Barthes in an automotive accident. Most recently, Glennan ([2017]) has elaborated the idea of a minimal mechanism, which is intended to be general enough to cover both ordinary biological mechanisms as well as these ‘ephemeral’ mechanisms. Surely, we can appreciate that Glennan is using the term ‘mechanism’ in a more catholic way than the term is used in, for example, MDC ([2000]).

Moreover, this expansive sense of ‘mechanism’ is the one that Ioannidis and Psillos ([2017], p. 605) have in mind when they identify what they call ‘Truly Minimal Mechanisms,’ an admittedly ‘thin’ sense of mechanism that is identical to a ‘causal pathway.’ And this broad notion is what Wesley Salmon ([1984], p. 297) has in mind, too, when he described causal mechanisms as ‘spatiotemporally continuous causal processes that transmit causal influence from one part of spacetime to another.’ Craver ([2007], p. 8) helpfully calls this latter sort of explanation ‘etiological’ mechanistic explanation, in contrast to ‘constitutive’ mechanistic explanation.

Commentators have often observed this (admittedly rough) distinction between senses of ‘mechanism.’ Consider Anderson ([2014a]). Her first sense of mechanism, ‘mechanism₁,’ is ‘tiered,’ ‘anti-reductive,’ and ‘regular,’ and it plays an especially prominent role in ‘so-called higher-level sciences like biology’ (p. 275). Her second sense of mechanism, ‘mechanism₂,’ is more general; it yields a ‘complete account of causation’ (p. 280) and does not require regularity; this sense is consistent with Glennan’s ‘ephemeral mechanisms.’ She notes that Glennan is the main proponent of this second sense of mechanism.¹⁰ Levy ([2013]), too, distinguishes between ‘explanatory mechanism,’ which is prominent in biology and has a ‘relatively concrete geometrico-mechanical flavor’ (p. 102), and ‘causal mechanism,’ which is ‘a view about the nature of causal relations,’ and

¹⁰ Anderson ([2014b]) also articulates three further senses of mechanism that are ‘ontologically flat’ and therefore compatible with a reductive ontology: mechanisms as single equations in structural equations models (mechanism₃), as ‘processes and/or interactions in physical process accounts of causation’ (mechanism₄), and as ‘net constraints on the states open to a system’ (mechanism₅).

which is ‘less central to the philosophy of biology’ (p. 101). Nicholson ([2012], p. 153) also points to the distinction when he separates ‘machine mechanisms,’ which are particularly prominent in biology, and ‘causal mechanisms,’ which are of ‘fundamental importance in scientific practice’ quite generally. As I indicated above, I think there’s a kernel of truth to Nicholson’s distinction, even though it fails to capture that mechanisms are processes, not systems.

My goal here is not to give a fine-grained taxonomy of the different senses of ‘mechanism’ at play in science. It does not matter, for present purposes, whether there are two or twelve distinct senses of the term. Moreover, none of this is to say that the notion of a causal pathway cannot play a fruitful role in the explication of the first, biological sense of mechanism (e.g., Illari and Williamson [2012]; Anderson [2014a], [2014b]). Perhaps the biological sort of mechanism will turn out to be a highly regular causal pathway with specific types of recurring elements. (If so, the relevant concept of a causal pathway would have to have much more structure than the bare notion of a cause-and-effect sequence.) Nor, finally, is my goal to resolve the many complex and subtle debates about the metaphysics of mechanism. What matters here is that we can distinguish, in a rough-and-ready way, between a sense of ‘mechanism’ (or a family of such senses) that is prominent in biology and that, loosely speaking, has the connotation of a machine, and a much more general sense of ‘mechanism’ (or a family of senses) that is ubiquitous in science and means nothing more than a causal trajectory. In the remainder of this paper, when I use the term ‘mechanism,’ I use it only in the first of the two senses.

When it came to functions, we found it convenient to distinguish these two prominent senses of ‘function’ with respect to normativity: proper functions are the sorts of things that can fail, but minimal functions cannot. Is there a similar distinction we can draw with respect to mechanisms? As indicated above, I think there is, and will try to show this in the next section. In this first sense of ‘mechanism,’ the sense that plays a special role in biology, mechanisms are the kinds of things that can break. The mechanism for blood circulation can break and give rise to various diseases. It makes sense to ask questions like, ‘is this mechanism broken, or is it working normally?’ But in the second sense of ‘mechanism,’ mechanism as mere causal trajectory, this normative language seems inappropriate: ‘Fortunately, the mechanism that would have killed Roland Barthes broke, and he was able to live out the rest of his life peacefully in the suburbs of Paris.’

Having set out these preliminary distinctions, I’m now in a position to state, much more rigorously than before, the core argument of this paper: To the extent that a mechanism is the kind of thing that can break, it is because it has the kind of function that can fail. But something has the kind of function that can fail only if it has the right sort of history.

Therefore, a mechanism is the kind of thing that can break only if it has the right sort of history. The next two sections will attempt to demonstrate each of these two premises in turn.

3. Mechanisms have proper functions

The first premise of my argument is that mechanisms have functions. More precisely, mechanisms, in the prominent biological sense of the term, have proper functions.

Having a proper function is not incidental to being a mechanism, but essential to it. In order for something to be a mechanism for blood circulation, it must have the function of blood circulation; in order for something to be a mechanism for triggering action potentials, it must have the function of triggering action potentials. One implication of this view is that, if a state is truly dysfunctional, there is no such thing as a mechanism for it. There is no such thing as a mechanism for cancer, or a mechanism for diabetes, or Alzheimer's disease. Rather, those diseases happen when mechanisms break.

I have argued elsewhere that mechanisms have functions; here I'll summarize my two main arguments for it.¹¹ These are the sociolinguistic argument and the argument from normativity. The first, sociolinguistic, argument, simply points out that, often enough, when biologists and biomedical researchers explain diseases, they do not seek a 'mechanism' for that disease. Rather, they describe the disease as something that happens when a mechanism for a function breaks. Alzheimer's disease, we are told, results from 'potentially irreversible impairments of synaptic memory mechanisms' (Rowan et al. [2003], p. 821). Drug addiction results from the way drug use 'usurps' the 'core synaptic mechanisms' for behavioral reinforcement (Kauer and Malenka [2007], p. 845).

Autoimmune disorders happen when the mechanisms that recognize self molecules do

¹¹ See Garson ([2013], [2017], [2019a], Chapter 10). Piccinini ([2010], p. 286, [2015], Chapter 6), Moghaddam-Taaheri ([2011]), Moss ([2012], p. 166), Craver ([2013], p. 140), Maley and Piccinini ([2018]), and Rosenberg ([2018]), also argue that mechanisms serve functions in some teleological sense of the term.

not work correctly: ‘The adaptive immune system uses multiple mechanisms to avoid damaging responses against self molecules. Occasionally, however, these mechanisms fail...causing autoimmune diseases, which can be fatal’ (Alberts et al. [2012], p. 1539). It is just as if, when biologists and biomedical researchers call something a ‘mechanism,’ they have a shared, and often tacit, understanding that mechanisms have proper functions, and diseases happen when these mechanisms break.

The biologically-informed reader will be quick to point out that biologists do not *always* speak this way. If you do a Google search for ‘mechanism for diabetes,’ you will get many hits, even in prominent science journals. This would seem to be a straightforward contradiction for my view, even an embarrassing one. However, the fact that biologists sometimes speak and write as if there are mechanisms for diseases would only contradict this view if there were no way of accommodating it. But there is: often enough, when biologists speak that way, they are speaking elliptically. When a biomedical researcher claims to have discovered a mechanism for diabetes, she is not saying that there is a mechanism for diabetes in the same sense in which there is a mechanism for blood circulation. Rather, she is saying that she has found a mechanism that is somehow or other implicated in the etiology of diabetes.¹² Consider a recent article that purports to identify a ‘mechanism for cancer metastasis’ (Rolli et al. [2010]). In my view, there is no such thing as a mechanism for cancer metastasis (assuming, of course, that nothing has the function of producing cancer metastasis). A careful perusal of the article, however, shows that the mechanism so identified, and identified as a mechanism, is merely a

¹² Compare Darden et al. ([2018]).

mechanism for cell elasticity. Cell elasticity is a functional and healthy ability which, tragically, can be recruited to help spread cancer.

I suppose that one could, by dint of effort, dig up some instances in which biomedical researchers discuss a ‘mechanism of disease,’ but in such a way that does not fit this neat attempt to accommodate it. There are about a quarter million articles published annually in medical journals, so I would not be surprised to find some uses of ‘mechanism’ that do not support this thesis.¹³ Keep in mind, however, that the term ‘mechanism’ is ambiguous: there is a prominent biological sense of mechanism, and there is a much more general, even tepid, sense of ‘mechanism,’ where mechanisms are mere causal trajectories. All diseases have mechanisms in the second sense but not the first. So, if one encounters talk of ‘disease mechanisms’ in the literature that cannot be accommodated in the way recommended here, then it is quite possible that the biologist is using the term in the highly general sense of a cause-and-effect sequence.

Let’s take stock. Here is a sociolinguistic fact that is hard to ignore: often enough, when biologists attempt to explain a disease, they do not look for its ‘mechanism;’ instead, they seek to reveal how the disease results from the breakdown of a mechanism for a functional state. This fact demands an explanation. I have presented such an explanation: mechanisms have proper functions. If one rejects that explanation, then it would be good to offer an alternative explanation for that fact. That would strike me as an ordinary and

¹³ See <https://www.nap.edu/read/10866/chapter/28> for a recent attempt to assess the number of medical articles published annually.

appropriate way to proceed. What is not appropriate, however, is to reject this explanation, and then go on as if there is nothing to be explained.

Here is a second argument for the view that mechanisms have proper functions, the argument from normativity. As I said, mechanisms, in the prominent biological sense, are the kinds of things that can break. This is part of ordinary biological and biomedical usage. But here we must go beyond linguistic usage to consider the features of reality that this usage reveals. What is it for a mechanism to break? That is, what is it for an X (an entity in the broadest sense, whether a process or a system) to be a broken-mechanism-for-Y (e.g., a broken mechanism for memory)? It seems to me that two things must be true of it:

- (1) X is a mechanism for Y; and
- (2) X cannot perform Y (for some inner, constitutional reason, and not just because it is in the wrong environment).

The problem, however, is that many of the standard accounts of mechanism in the new mechanism literature, as cited below, do not allow (1) and (2) to be true at the same time. They generally make it a condition of something's being a mechanism for Y that it actually be able to perform Y. If something cannot perform Y, then, by definition, it is not a mechanism for Y; ipso facto, it is not a broken mechanism for Y.

Consider the famous MDC ([2000], p. 3) definition, which says that the mechanism's components must be organized 'such that they are productive of regular changes [i.e., the changes that constitute the mechanism's phenomenon].' Alternately, consider Glennan's ([2017], p. 17) umbrella definition, which holds that the components of mechanisms must be 'organized so as to be responsible for the phenomenon.' In the same vein, Bechtel and Abrahamsen ([2005], p. 423) say, 'A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization.' It would seem that, with this last definition, Bechtel and Abrahamsen recognize that mechanisms serve functions; the problem is that, in other places, Bechtel maintains that functions are merely causal roles, i.e., dispositions (Bechtel and Richardson [2010/1993], p. 18). That would imply that it is not possible for a mechanism to have a function but to have lost the disposition to perform that function. All of these definitions seem to imply that, if a mechanism cannot perform Y, because of a missing or malformed component, or because of some structural or temporal disorganization, then it is not a mechanism for Y at all.

Suppose, however, that mechanisms have proper functions. That is, suppose that in order for something to be a mechanism for Y, it must have Y as its proper function. This assumption lets us see at once what broken mechanisms are. Something is a broken-mechanism-for-Y when it is a mechanism for Y, but it cannot perform its function. In part, what it is for something to be a broken-mechanism-for-Y is to have a proper function that it cannot perform. The reason that the notion of proper function is such a powerful explanatory tool here is that (as we've learned from four decades of the biological functions debate) functions are 'normative.' They are the sorts of things that

can fail. One way of describing the claim that I am making here is this: *mechanisms are the kinds of things that can break because functions are the kinds of things that can fail.*

Put differently, the normativity of mechanism is parasitical on the normativity of function.

One might wonder if there's a different way of understanding how mechanisms break, one that does not invoke lofty conceptions of proper function. Couldn't we just say that a mechanism breaks when it doesn't act the way we expect it to act, or when its behavior departs from an established pattern? Krickel ([2018], p. 47), for example, tells us: 'The normativity of mechanism-talk in the context of pathologies does not stem from their fulfilling a function, but rather from the statistical expectancy arising from the regularity of mechanisms.' The idea here seems to be that the reason we can say of a memory mechanism that it is 'broken' in the case of Alzheimer's disease is that it is behaving in a way that subverts our statistically-grounded expectations.

Unfortunately, I don't think this idea does justice to the notion of a mechanism's breaking, for two reasons. First, there are plenty of instances in which a mechanism's breaking is quite expectable, from a statistical point of view, such as a certain amount of cognitive decline in older people. Second, there are quite a few instances in which something behaves in an unexpected manner, but it's not broken. I recently found out that Old Faithful's eruptions (one of Glennan's examples) aren't quite as faithful as its name would indicate. Intriguingly, scientists can only accurately predict its eruption about 90% of the time. Should we say, of the other 10% of the time, that its mechanism is broken?

Would it be sensible for a Yellowstone geologist to tap his watch and say, ‘Old Faithful should have erupted by now, since it’s been over 94 minutes since the last one, but it hasn’t yet, so its mechanism must be broken?’ It seems to me (not being a geologist) that nobody would talk that way. One would be tempted to say, rather, that Old Faithful’s mechanism is more complicated than we thought, or that some interfering factor adjusted its normal interval. A mechanism’s breaking, and a mechanism’s deviating from a typical pattern, are two quite different things.¹⁴

A final worry should be considered before moving on. An anonymous referee suggested that talk of a mechanism performing, or failing to perform, a function, is inextricably tied to the faulty idea that mechanisms are systems rather than processes. It’s true that in the functions literature, philosophers often attribute functions, in the first place, to systems (‘the heart has the function of pumping blood’) rather than processes (‘the heart’s pumping has the function of circulating blood’). But I don’t attach any deep theoretical significance to this way of speaking. Functions can just as correctly be attributed to processes as to systems.

This section has shown that mechanisms have proper functions. The next section will show that, for something to have a proper function, it must have the right sort of history.

¹⁴ An anonymous reviewer pointed out that in this case, we could say that a pattern has been broken. That sounds right to me. But it still doesn’t sound right to say that the mechanism is broken. Of course, if the periodicity of a system is due to a mechanism that has the proper function of maintaining that periodicity, such as circadian rhythms, then I’d be inclined to say that the mechanism is broken.

This is true, I argue, on any theory of proper function we wish to accept, even those that purport to be ‘ahistorical.’ But if that is right, then mechanisms are historical, too.

4. Proper functions are historical entities

Now, on to the second premise: functions are historical entities. Whether or not something has a function depends on its history. My argument for this claim, which I develop in Garson ([2019b]), is simple: all plausible theories of proper function are tacitly, if not explicitly, historical, and the best explanation for this is that functions are, indeed, historical entities – like adaptations, volcanic mountains, siblings, igneous rocks, and original Van Goghs.

This second premise is a little more difficult to establish than the first. That is because it flies in the face of a popular, pluralistic approach to function. This pluralistic approach to function claims that there are two different kinds of proper function at play in biology. The first is the historical kind. Sometimes, when biologists attribute functions to traits, they want to know ‘why the trait is there.’ Almost everyone seems to agree that the selected effects theory of function (which says, roughly, that the function of the trait is whatever it was selected for) does a good job capturing this historical kind of function.

The second is the ahistorical kind. Sometimes it is said that, when biologists attribute functions to traits, they don’t mean to say anything about the trait’s history. They just want to know how the trait currently aids the organism’s survival or reproductive

prospects, or some other interesting feature of the organism. Almost everyone seems to think that the causal role theory of function (according to which the function of a trait consists in its current-day contribution to an interesting system capacity), or something in its neighborhood, captures this ahistorical sort of function.

So, in short, here is the pluralist picture: when biologists are concerned with the historical sort of function, they are implicitly using the selected effects theory, and when they are concerned with the ahistorical sort of function, they are implicitly using the causal role theory.¹⁵

I think this pluralistic view is quite mistaken. My argument is not that the selected effects theory is the only good theory of function and all the others are wrong. My argument is that every plausible theory of function is tacitly, if not overtly, historical. There are no theories of function that are both plausible and genuinely ahistorical. But if there are no plausible ahistorical theories of function, that suggests that functions are historical entities. In other words, if nobody has figured out a good way to define the notion of function without (at least tacitly) invoking history, then there's probably a reason for that: functions are historical entities – they are essentially historical – and the historicity of function manages to impress itself upon everyone's theories.

¹⁵ See, e.g., Godfrey-Smith ([1993], p. 200), Amundson and Lauder ([1994], p. 446), and Maclaurin and Sterelny ([2008], p. 114) for statements of this pluralistic view. Neander ([2017], Chapter 3) and Garson ([2018]) criticize this picture. Also see Cusimano and Sterner ([2019]), and Malaterre ([forthcoming]), for recent approaches to function pluralism.

I will give some examples of theories that purport to be ahistorical, but really are not. Christopher Boorse has, for decades, advanced a ‘goal-contribution’ theory of function. Functions are just contributions to goals. But the organism’s highest goals are survival and reproduction. So, in physiology, the function of a trait is its ‘species-typical’ contribution to survival or reproduction. The function of the mammalian heart is to beat because that is how mammal hearts typically contribute to survival. A trait token (e.g., my liver) is diseased when it cannot make this species-typical contribution to fitness. His major argument for the view is that it tallies with the alleged fact that biologists, as a rule, have ‘no thought of explaining [a trait’s] history’ ([2002], p. 73) when they attribute functions to them. He also thinks it is intuitively obvious that, the very moment a trait begins to benefit an organism, it has the function of doing so, even if it has no history at all.

One of the traditional objections to Boorse’s view is the problem of pandemic disease, raised by Karen Neander ([1991]). Suppose a disease becomes nearly universal. For a thought experiment, suppose a worldwide virus makes everyone blind for a two-day period of time. (What I mean is that everyone is blind at once, not that some people are blind sometimes and others other times.) According to the letter of Boorse’s view, during that two-day window of time, the eye would not have the function of sight, since, during that window, seeing isn’t how it contributes to fitness. (The argument need not rely on clever thought experiments. There are real examples of pandemic disease in nature, such as the one we are currently living through.)

Boorse has discussed, at some length, this problem of pandemic disease. He now says that when we assess what counts as a ‘species-typical’ contribution to fitness, we cannot just look at what the trait does at this very moment. We also have to examine how the trait behaves over a long slice of time that includes the present moment and reaches far back into the past: ‘*Obviously*, some of the species’ history must be included in what is species-typical ([2002], p. 99; my emphasis).’ He says that this time-slice must be longer than ‘a generation or two’ and might include ‘millennia.’ What that means is that, in at least some circumstances, when we’re trying to figure out a trait’s function (such as the eyespots of butterfly wings) we must look to history.

Consider, as a second example, the causal role theory of function. The causal role theory holds, roughly, that the function of a trait token depends on its contribution to some interesting system capacity, where what is ‘interesting’ depends on the interests and goals of the relevant group of researchers. One long-standing problem for this view is that it struggles to make sense of malfunction or dysfunction. That is because, on the causal role view, functions are traditionally understood as dispositions: ‘...to attribute a function to something is, in part, to attribute a disposition to it’ (Cummins [1975], p. 758). But that means that, if a trait token (e.g., my heart’s pumping) loses the disposition to perform the activity that constitutes its function, then it loses the function forthwith – and hence it cannot be dysfunctional, either.

A handful of causal role theorists have both acknowledged the problem and have attempted to sketch an account of dysfunction from the causal role view. A standard

approach is to introduce statistics: put simply, a trait token (e.g., my heart) is dysfunctional when it cannot do what tokens of the type typically do by which they contribute to the aforementioned interesting system capacity (Godfrey-Smith [1993], p. 200; Craver [2001], p. 72). A problem with this otherwise laudable attempt is that it brings us right back to pandemic disease. If the eye's function is what it typically does when it contributes in the right way to the system-level capacity of interest (vision), then if a pandemic disease rendered everyone blind for a month, eyes wouldn't have the function of seeing during that time. An obvious way to get around this problem, and perhaps the only way, is to glance at history: a trait's function would consist in its typical contribution to said interesting capacity, where what's typical is assessed over a long slice of time that reaches far back into the past.

It might not seem very remarkable, in and of itself, that functions are intimately tied to history. Clearly, all biological traits, functional or not, are products of a historical process, and hence 'have a history.' The reason this claim is, in fact, remarkable is that so many theorists have insisted, for so long, that functions are *essentially* ahistorical, that is, that whether or not something has a function, and what function it has, does not depend in any essential way on history. (I realize I'm testing the reader's patience, but here is another thought experiment to clarify the point: the idea is that, even if the world and everything in it were created only five minutes ago in pretty much its present form, things would still have functions – see Bigelow and Pargetter ([1987], p. 188). That's the claim we should deny.) What this discussion reveals is that functions are essentially historical,

like siblings and volcanic mountains, and unlike gold and water. My further claim is that this historical essence manages to well up and infect mechanisms, too.

My point here is not that the causal role theory is wrong, or that Boorse's theory is wrong. I do not know whether they are right or wrong, or whether each one captures a real strand of ordinary biological usage. My point is that they are all historical. But why would this be? Why would it be so hard to come up with a decent theory of proper function that is genuinely ahistorical? An obvious explanation is that functions are historical entities. Put differently, and more precisely, something has the kind of function that can fail only if it has the right sort of history.

With this, the justification of my core argument is complete. Now for some consequences.

5. Some consequences

I have proposed the following argument: To the extent that a mechanism is the kind of thing that can break, it is because it has the kind of function that can fail. But something has the kind of function that can fail only if it has the right sort of history. Therefore, a mechanism is the kind of thing that can break only if it has the right sort of history.

There are two main consequences of recognizing that mechanisms are historical entities, both of which I've alluded to in the introduction.

The first is epistemic. It implies that, at least in some contexts, in order to know whether or not something is a mechanism, and what it is a mechanism for, we must consider its history. It is not enough to poke and prod at it in a lab or in the field: take it apart, strategically break it, put it back together, in short, to carry out the various operations that Craver ([2007], p. 141) summarizes under the label of ‘mutual manipulability.’ Additionally, we would have to know something about its history.

There is an obvious analogy to adaptations. Most philosophers of biology recognize that adaptations are historical entities. In order for the Venus flytrap’s sticky leaves to be adaptations for catching bugs, they must have an actual, historical, track record of catching bugs, and they must exist now, in part, because of this track record. That means that one good way to determine that the flytrap’s sticky leaves are an adaptation – and sometimes, the only way – is to investigate its history. I am making an analogous claim about mechanisms.

There is a big difference, however, between mechanisms and adaptations. For I do not claim, here, that mechanisms are adaptations, that is, that in order for something to be a mechanism for X, it must have been shaped by natural selection for X. My claim is that it must have the proper function of doing X, and therefore it must have the right kind of history. But what kind of history is the right kind? It is whatever kind of history the correct theory of proper function says it must have. If one accepts the selected effects theory, then one is committed to the claim that in order for something to have a proper function X, it must have been shaped by selection for X. If one accepts Boorse’s goal-

contribution account, one is merely committed to the claim that in order for something to have a function, it must have contributed to survival or reproduction over a long period of time that includes the present moment and stretches back into the past. (Something can contribute to survival and reproduction and not undergo selection, for example, if there is no variation in the population for it, or if it is not inherited. So, Boorse's view does not collapse into the selected effects view.)

Here is a second, metaphysical, consequence. There is a tendency among philosophers to use concepts drawn from the new mechanism tradition as a template for thinking about constitutive explanations in science quite generally.¹⁶ For example, we say that salt's solubility is explained by salt's polar molecular structure. This polar molecular structure, however, does not cause salt's solubility, in the before-and-after sense. It somehow constitutes this very solubility. What, then, is this mysterious relation of constitution – if there is, in fact, a single relation at hand? What, exactly, is the relationship between the polar molecular structure of salt, and its solubility, such that the former can play into a constitutive explanation of the latter? Craver, as noted above, suggests that we use the notion of mutual manipulability to articulate this relationship.

It strikes me as incautious to use concepts drawn from the study of mechanisms in biology to understand, in a philosophically rigorous manner, constitutive explanations in science quite generally.¹⁷ That is because mechanisms are historical, but constitutive

¹⁶ See Aizawa ([2016]) for discussion.

¹⁷ Aizawa ([2016]) and Couch ([2012]) concur, but for different reasons than I do. Note that the question at issue here is whether we should use, e.g., mutual manipulability, to

relations, as a rule, are not. There are plenty of constitutive relations in science (for example, the relation between the polar molecular structure of salt and its solubility) that are purely ahistorical. They depend merely on chemical composition and the immutable laws of physics. My concern is that we may be borrowing concepts that are well-suited, and even well-designed, for understanding the biological world, and precipitously carrying them over into the study of the physical world as a whole. As I noted in the introduction, I cannot prove that this is a mistake. The point is that, once we recognize that there are no mechanisms (in the prominent biological sense) in the realm of physics and chemistry, we can see that there's no prior justification for importing the language and concepts associated with them into that level of reality.

articulate philosophically the nature of constitutive relations in science. It is not a question of using mutual manipulability as a technique or practice for discovering constitutive relations in, say, physics or chemistry. I don't have a problem with the latter; my problem is with the former.

References

Aizawa, K. [2016]: 'Compositional Explanation: Dimensioned Realization, New Mechanism, and Ground', in K. Aizawa, and C. Gillett (eds), *Scientific Composition and Metaphysical Ground*, London: Palgrave Macmillan, pp. 75-90.

Alberts, B., et al. [2012]: *Molecular Biology of the Cell: 5th Ed*, New York: Garland.

Amundson, R., and Lauder, G. V. [1994]: 'Function Without Purpose: The Uses of Causal Role Function in Evolutionary Biology', *Biology and Philosophy*, **9**, pp. 443-69.

Anderson, H. [2014a]: 'A Field Guide to Mechanisms: Part I', *Philosophy Compass*, **9**, pp. 274-83.

Anderson, H. [2014b]: 'A Field Guide to Mechanisms: Part II', *Philosophy Compass*, **9**, pp. 284-93.

Bechtel, W., & Abrahamsen, A. [2005]: 'Explanation: A Mechanist Alternative', *Studies in the History and Philosophy of Biological and Biomedical Sciences*, **36**, pp. 412-41.

Bechtel, W., and Richardson, R. C. [2010/1993]: *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*, Princeton, NJ: Princeton University Press.

Forthcoming in *British Journal for Philosophy of Science*
Preprint (not copyedited or formatted)

Bigelow, J., and Pargetter, R. [1987]: 'Functions', *Journal of Philosophy*, **84**, pp. 181-96.

Boorse, C. [2002]: 'A Rebuttal on Functions', in A. Ariew, R. Cummins, and M. Perlman (eds), *Functions: New Essays in the Philosophy of Psychology and Biology*, Oxford: Oxford University Press, pp. 63-112.

Buss, D. M. [2008]: *Evolutionary Psychology: The New Science of the Mind. 3rd ed*, Boston: Pearson.

Caro, T., et al. [2014]: 'The Function of Zebra Stripes', *Nature Communications*, **5**, 3535.

Couch, M. [2011]: 'Mechanisms and Constitutive Relevance', *Synthese* **183**, pp. 375-88.

Craver, C. [2001]: 'Role Functions, Mechanisms, and Hierarchy', *Philosophy of Science*, **68**, pp. 53-74.

Craver, C. F. [2007]: *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*, Oxford: Clarendon Press.

Craver, C. [2013]: 'Functions and Mechanisms: A Perspectivalist View', in P. Huneman (ed), *Function: Selection and Mechanisms*, Dordrecht: Springer, pp. 133-58.

Forthcoming in *British Journal for Philosophy of Science*
Preprint (not copyedited or formatted)

Craver, C. F., and Darden, L. [2013]: *In Search of Mechanisms: Discoveries Across the Life Sciences*, Chicago: University of Chicago Press.

Cummins, R. [1975]: 'Functional Analysis', *Journal of Philosophy*, **72**, pp. 741-765.

Cusimano, S., and Sterner, B. [2019]: 'Integrative Pluralism for Biological Function', *Biology and Philosophy*, **34**, 55.

Darden, L. [2006]: *Reasoning in Biological Discoveries*, Cambridge: Cambridge University Press.

Darden, L., et al. [2018]: 'Harnessing Formal Concepts of Biological Mechanism to Analyze Human Disease', *PLoS Computation Biology*,
<https://doi.org/10.1371/journal.pcbi.1006540>

Devitt, M. [2008]: 'Resurrecting Biological Essentialism', *Philosophy of Science*, **75**, pp. 344-82.

Garson, J. [2013]: 'The Functional Sense of Mechanism', *Philosophy of Science*, **80**, pp. 317-33.

Garson, J. [2016]: *A Critical Overview of Biological Functions*, Dordrecht: Springer.

Forthcoming in *British Journal for Philosophy of Science*
Preprint (not copyedited or formatted)

Garson, J. [2017]: 'Mechanisms, Phenomena, and Functions', in S. Glennan, and P. Illari (eds), *Routledge Handbook of Philosophy of Mechanisms*, London: Routledge, pp. 104-15.

Garson, J. [2018]: 'How to be a Function Pluralist', *British Journal for the Philosophy of Science*, **69**, pp. 1101-22.

Garson, J. [2019a]: *What Biological Functions Are and Why They Matter*, Cambridge: Cambridge University Press.

Garson, J. [2019b]: 'There are no Ahistorical Theories of Function', *Philosophy of Science*, **86**, pp. 1146-56.

Glennan, S. [1996]: 'Mechanisms and the Nature of Causation', *Erkenntnis*, **44**, pp. 49-71.

Glennan, S. [2005]: 'Modelling Mechanisms', *Studies in the History and Philosophy of Biological and Biomedical Sciences*, **36**, pp. 443-64.

Glennan, S. [2010]: 'Ephemeral Mechanisms and Historical Explanation', *Erkenntnis*, **72**, pp. 251-66.

Glennan, S. [2017]: *The New Mechanical Philosophy*, Oxford: Oxford University Press.

Forthcoming in *British Journal for Philosophy of Science*
Preprint (not copyedited or formatted)

Godfrey-Smith, P. [1993]: 'Functions: Consensus without Unity', *Pacific Philosophical Quarterly*, **74**, pp. 196-208.

Illari, P. M., and Williamson, J. [2012]: 'What is a Mechanism? Thinking About Mechanisms Across the Sciences', *European Journal for Philosophy of Science*, **2**, pp. 119-35.

Ioannidis, S., and Psillos, S. [2017]: 'In Defense of Methodological Mechanism: The Case of Apoptosis', *Axiomathes*, **27**, pp. 601-19.

Kaiser, M. I., and Krickel, B. [2017]: 'The Metaphysics of Constitutive Mechanistic Phenomena', *British Journal for the Philosophy of Science*, **68**, pp. 745-79.

Kauer, Julie A., and Robert C. Malenka. [2007]: 'Synaptic Plasticity and Addiction', *Nature Reviews Neuroscience*, **8**, pp. 844-58. .

Krickel, B. [2018]: *The Mechanical World: The Metaphysical Commitments of the New Mechanistic Approach*, Cham, Switzerland: Springer.

Kröger, R. H. H. and Oliver Biehlmaier. [2009]: 'Space-Saving Advantage of an Inverted Retina', *Vision Research*, **49**, pp. 2318-21.

Forthcoming in *British Journal for Philosophy of Science*
Preprint (not copyedited or formatted)

Levy, A. [2013]: 'Three Kinds of New Mechanism', *Biology and Philosophy*, **28**, pp. 99-114.

Machamer, P., Darden, L., and Craver, C. F. [2000]: 'Thinking about Mechanisms', *Philosophy of Science*, **67**, pp. 1-25.

Maclaurin, J., and Sterelny, K. [2008]: *What is Biodiversity?*, Chicago: University of Chicago Press.

Malaterre, C. [Forthcoming]: 'What are Ribozymes For? Arguing for Function Pluralism', in J. Gayon, A. De Ricqlès, and A. Dussault (eds), *Functions: From Organisms to Artefacts*, Dordrecht: Springer.

Maley, C. J., and Piccinini, G. [2018]: 'A Unified Mechanistic Account of Teleological Functions for Psychology and Neuroscience', in D. Kaplan (ed), *Integrating Psychology and Neuroscience: Prospects and Problems*, Oxford: Oxford University Press, pp. 236-56.

Moghaddam-Taaheri, S. [2011]: 'Understanding Pathology in the Context of Physiological Mechanisms: The Practicality of a Broken-Normal View', *Biology and Philosophy*, **26**, pp. 603-11.

Moss, L. [2012]: 'Is the Philosophy of Mechanism Philosophy Enough?', *Studies in the*

Forthcoming in *British Journal for Philosophy of Science*
Preprint (not copyedited or formatted)

History and Philosophy of Biological and Biomedical Sciences, **43**, pp. 164-72.

Neander, K. [1991]: 'Functions as Selected Effects: The Conceptual Analyst's Defense'
Philosophy of Science, **58**, 168-84.

Neander, K. [2017]: *A Mark of the Mental: In Defense of Informational Teleosemantics*,
Cambridge, MA: MIT Press.

Nicholson, D. J. [2012]: 'The Concept of Mechanism in Biology', *Studies in the History
and Philosophy of Biological and Biomedical Sciences*, **43**, pp. 152-63.

Piccinini, G. [2010]: 'The Mind as Neural Software? Understanding Functionalism,
Computationalism, and Functional Computationalism', *Philosophy and
Phenomenological Research*, **81**, pp. 269-311.

Piccinini, G. [2015]: *Physical Computation: A Mechanistic Account*, Oxford: Oxford
University Press.

Prudic, K. L., et al. [2015]: 'Eyespots Deflect Predator Attack Increasing Fitness and
Promoting the Evolution of Phenotypic Plasticity', *Proceedings of the Royal Society B*,
282, 201415.

Rolli, C. G., et al. [2010]: 'Impact of Tumor Cell Cytoskeleton Organization on

Forthcoming in *British Journal for Philosophy of Science*
Preprint (not copyedited or formatted)

Invasiveness and Migration: A Microchannel-Based Approach', *PLoS One*, **5**, e8726.

Rosenberg, A. [2018]: 'Making Mechanism Interesting', *Synthese*, **195**, pp. 11-33.

Rowan, M. J., Klyubin, I., Cullen, W. K., and Anwyl, R. [2003]: 'Synaptic Plasticity in Animal Models of Early Alzheimer's Disease', *Philosophical Transactions of the Royal Society of London B*, **358**, pp. 821-8.

Salmon, W. [1984]: 'Scientific Explanation: Three Basic Conceptions', *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, **1984**, pp. 293-305.

Tooby, J., and Cosmides, L. [2006]: 'Toward Mapping the Evolved Functional Organization of Mind and Brain', in E. Sober (*ed*), *Conceptual Issues in Evolutionary Biology*, 3rd ed., Cambridge, MA: MIT Press, pp. 175-95.

Williams, G. C. [1966]: *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*, Princeton: Princeton University Press.

Wimsatt, W.C. [1976]: 'Reductive Explanation: A Functional Account', in R. S. Cohen, and A. Michalos (*eds*), *Proceedings of the 1974 meeting of the Philosophy of Science Association*, Dordrecht: D. Reidel, pp. 671-710.