Mechanisms, ceteris paribus laws and covering-law explanation

Nancy Cartwright, John Pemberton and Sarah Wieten[1]

1. What's in this paper

Mechanisms – stable arrangements of parts that acting together produce novel behaviour (behaviour that is not characteristic of any of the parts singly) are all the rage now in philosophy of science. They are supposed to provide one of modern science's basic explanatory devices. What do they explain? In 1989 Nancy Cartwright introduced the idea of 'nomological machines'[2], which are mechanisms whose repeated operation gives rise to regular behaviours of the kind we record in low-level scientific laws, like Kepler's laws for the motions of the planets. The nomological machine explains the law, which holds 'ceteris paribus' – *relative to* the proper operation of the nomological machine.

In his 2012 paper in the *Journal of Philosophy*, 'Ceteris Paribus Hedges: Causal Voodoo that Works'[3], Michael Strevens adopts the same view, using it to offer a semantics for ceteris paribus (*cp*) laws that defends them from the charge of vacuity. Laws with the phrase *ceteris paribus* in front (like 'Ceteris paribus, printing money causes inflation') have genuine content, he argues, because the clause refers, albeit often without mention, to the mechanism that explains the law. The voodoo consists in the fact that the clause, in referring to the mechanism, provides content to the cp law even though we do not know much about the mechanism referred to, including the features that allow it to generate the behaviour described in the law. As Strevens says, '[W]hat is intriguing is the possibility that, with the help of a familiar Latin expression, we can frame short and simple sentences that entail actual event patterns in all their glorious and gory complexity.'[4] They do just that on the semantics he offers.

All this raises further questions. We shall address three: a question of practical use, an epistemological question and an ontological one. In addition, we shall raise doubts about Strevens' Humean programme.

---

[1] Nancy Cartwright's work for this paper is based upon research supported by the National Science Foundation under grant no. 1632471 and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 667526 K4U), for which she is very grateful. It is acknowledged that the content of this work reflects only the authors' views and that the ERC is not responsible for any use that may be made of the information it contains.

[2] Nancy Cartwright, "Where do laws of nature come from", *Dialectica*, 51, 65-78, 1997.

[3] Michael Strevens, "Ceteris Paribus Hedges: Causal Voodoo that Works", *Journal of Philosophy* 109, 652–675, 2012.

[4] Ibid, page 5.

The question of practical use is 'Of what use are cp laws if we cannot pick out which systems satisfy the cp clause?'. The answer, we shall argue, is in the form of *markers* and *cautions*. We can often use the cp law even though we cannot identify features of the mechanism relevant to its operation because we learn markers that identify systems that afford the relevant behaviour and we also learn cautions about how to treat the system so that it will do so much of the time.

The epistemological question is 'What kind of explanation is involved?'. We shall answer, contrary to what many mechanists argue, 'Old-fashioned covering-law explanation.'

The ontological question is 'What is going on in the world when mechanism M gives rise to/generates/affords the behaviour recorded in law L?'. We shall argue that the *arrangement* of the parts in the mechanism supplies them with features they do not possess separately. M gives rise to behaviour B described in L when B is what it takes for some set of principles that govern features of M's parts all to be instanced in M's operation.

There has been some discussion in the literature about what a mechanism is and just what its boundaries are. James Woodward, for example, employs an 'invariance' account of what a mechanism is according to which 'Mechanisms consist of parts, the behaviour of which conforms to generalisations that are invariant under interventions, and which are modular in the sense that it is possible in principle to change the behaviour of one part independently of the others.'[5] In the medical literature and much of the social science literature,[6] by contrast, 'mechanism' usually refers to the sequence of steps in the causal process by which the cause produces its effect (where each step may itself be accounted for by some underlying powers or causal structure), as in the work of philosopher Daniel Steel who 'use[s] the term mechanism to refer to regularly operating causal relationships'[7]. Jon Elster's sense of the term is different yet again. Focusing on explanation in the social sciences, Elster develops an account of mechanisms designed to fill the explanatory gap between laws and mere description: 'Roughly speaking, mechanisms are *frequently occurring and easily recognizable causal patterns that are triggered under generally unknown conditions or with indeterminate consequences*. They allow us to explain, but not to predict.'[8] Our use

---

[5] James Woodward, "What Is a Mechanism? A Counterfactual Account" *Philosophy of Science* Vol. 69, No. S3, S366-S377, September 2002, abstract.

[6] For a review and criticisms of this 'intervening variables' approach to mechanisms in the social sciences, see Derek Beach and Rasmus Brun Pedersen, *Process-Tracing Methods.* University of Michigan Press, 2016, section 3.3.

[7] Daniel Steel, *Across the Boundaries.* Oxford University Press, 2008, page 40.

[8] Jon Elster, *Explaining Social Behaviour*, Cambridge: Cambridge University Press, 2007, page 36; original emphasis. Elster mechanisms often come in complementary pairs, which is one reason prediction is difficult. For instance, the *endowment effect*, in which 'a memory of a good experience is a good memory' and the 'experience of a bad [memory]

of the term 'mechanism', follows Cartwright's account of nomological machines, according roughly with that of the new mechanists.[9] For purposes of the discussion here, a *mechanism* is characterised by a set *P* of parts, in an arrangement *A* in which the parts display a specific set *Υ* of features and activities. We shall in our final analysis stress the importance of the arrangement for providing the parts with features that fall under general principles and thus allow the mechanism to give rise to the cp law that it does. A mechanism is a *nomological machine* if, when operating 'without interference' (a phrase we shall discuss), it gives rise to stable input-output relations of the kind typically recorded in causal laws... *ceteris paribus* causal laws.[10] Throughout, we shall use the terminology of 'nomological machines' as well as the now more standard terminology of 'mechanism' to underline this role in affording ceteris paribus laws.

To answer the question of practical use, we provide an account of how we learn to recognise and use nomological machines. Strevens touches on this question but does not aim to deal with it. He titles his paper 'Voodoo that works' because the truth conditions he offers for cp laws 'are typically opaque to the very scientists who formulate and test them.'[11] So, as he himself says '.... this power of ceteris paribus hedges may seem to be not only miraculous but useless. What is the practical significance of content in a hypothesis unless the investigators know that it is there?'.[12] We think this is where the real 'voodoo' lies. The reference to mechanisms is opaque but we can still put our cp claims to good use. We do so by learning to recognise markers and cautions. That we can do so is key to much of daily and scientific life.

With respect to ontology, note two problematic notions in the characterisation of nomological machines: 'give rise to' and 'without interference'. Strevens has something to say about both. For 'gives rise to'[13], he begins by remarking that the behaviours described in cp laws are the 'consequences' of the mechanisms. Later he substitutes 'explains' and argues that a good scientific model of the mechanism

---

is a bad memory' suggests opposite outcomes to those of the *contrast effect*, in which memories of past experiences have an opposing effect on present experiences. (Ibid, page 56)

[9] See e.g. Peter Machamer, Lindley Darden and Carl F. Craver. "Thinking about Mechanisms", *Philosophy of science* 67: 1-25, 2000; William Bechtel and Adele Abrahamsen, "Explanation: A Mechanist Alternative", *Studies in the History and Philosophy of the Biological and Biomedical Sciences* 36: 421-441, 2005; Stuart Glennan, "Rethinking Mechanistic Explanation", *Philosophy of Science* 69, S342-353, 2002; Phyllis Illari and Jon Williamson. "What is a Mechanism? Thinking about Mechanisms across the Sciences", *European Journal of Philosophy of Science* 2: 119-135, 2012.

[10] Following Strevens we shall generally confine our attention to cp causal laws although mechanisms can give rise to behaviours described in non-causal laws as well. Also, we shall use 'laws' and 'principles' interchangeably, depending on what is common usage for the ones under discussion; and where there is no danger of confusion, for brevity's sake, we may not always distinguish cp laws from the behaviours they describe.

[11] Strevens (n2), page 1.

[12] Ibid, page 6.

[13] Strevens does use the expression 'gives rise to' but at a different location. He talks of ceteris paribus laws giving rise to regularities.

will explain, in a perfectly usual sense, the ceteris paribus laws that describe what happens when it operates regularly.[14] But this is not good enough to characterise what 'giving rise to' consists in and thus to answer our ontological question. Explanation is a linguistic enterprise. For the ontology, we need to know what is going on in nature between the mechanism and the behaviours, not just what is going on in our science (or in an ideal final science) between a model and the cp law. Nor does Strevens explicitly claim to treat this latter issue -- he just does not take it up. But it is a pressing and, we shall argue, difficult problem that must be addressed. Happily, we think we can offer an account that can do the job. Developing this account in answer to the ontological question is a main aim of this paper.

2. Strevens' voodoo

A law claim ought, Strevens assumes, to imply 'its corresponding Humean generalization', where the 'corresponding Humean generalization' is 'a precise statement of the pattern of events that the law would give rise to.'[15] We shall take this for granted. The problem is that many of the generalisations that are really true are hugely complex due to the many 'enabling conditions' that must be present and the many 'interferences' that must be absent, and our law claims are not correspondingly complex. The fix is to preface the claim with 'ceteris paribus', so that 'CP, F's cause G's' expresses the appropriate complex generalisation. It does so because the cp clause narrows the scope of the law claim. It does so in three ways.

First, on Stevens account[16], the cp clause conditions the law claim to a particular mechanism. Strevens says: 'When a causal hypothesis is framed it is supposed to make a claim about a particular contextually determined mechanism: the target mechanism.'[17] Second, the cp clause also secures reference to the enabling conditions that are required for F's to cause G's. Third, it is not sufficient that the target mechanism be present; it must operate properly. The cp clause secures reference to this as well.

So, as the truth condition for 'CP, F's cause G's' Strevens proposes:

---

[14] Cartwright too has claimed that models can be used to explain these ceteris paribus laws; one of the important ways in which models function, she has argued, is as *blueprints* for the nomological machines that give rise to the regularities described by ceteris paribus laws. Nancy Cartwright, "Models: The Blueprints for Laws", *Philosophy of Science*, Vol. 64, S292-S303, Dec. 1997.
[15] Strevens (n2) page 1.
[16] Also on Cartwright's. Note that on neither account is there a suggestion that we must know what the mechanism is – that is the point of Strevens' voodoo. On Stevens' account, there should be *a particular* mechanism, not an indefinite array. Cartwright is more agnostic, merely claiming that generally there is some mechanism or other.
[17] Ibid, page 9.

- When condition O holds, then by way of the target mechanism M, the conditions Z and the property F bring about the property G. [18]

Here Z represents the enabling conditions and 'O is the set of conditions required for the successful operation of M...' To illustrate with one of Cartwright and John Pemberton's standard examples.[19] 'CP, pressing the lever causes flushing of the cistern' expresses this: 'If the toilet mechanism operates properly (O), then by way of the toilet cistern mechanism (M), pressing the lever (F) when the cistern is full (Z) causes flushing (G)'. The conditions that must fill in for O include, as Strevens agrees, all those interfering conditions that might stop the mechanism from affording 'Fs causes Gs'.

Consider now the big advantages of Strevens' semantics: it allows us to move beyond two unemployable conceptions of ceteris paribus hedges, which we call the 'Empty' and the 'Boring' accounts. Strevens' account manages to be both well-formulated and useful and, in so doing, avoids some of the traditional critiques of ceteris paribus hedges. In the Empty account, 'CP, F's cause G's' says 'If Φ were to obtain, F's would cause G's'. Under this account, a cp law is not a claim, rather it is an open formula with a dangling variable, since no content is provided for Φ, without which this expression cannot have a truth value. In the Boring account, the formula is closed, but uninformative. It says 'conditions Φ, such that if Φ were to obtain, F's would cause G's'. This account has the benefit that cp laws state functioning claims. Unfortunately, this work-around weakens the claim until it is Boring. Of course there is some set of conditions for which, should they hold, F's cause G's— if the condition *F's do not fail to cause G's* were to obtain, then 'F's causes G's' would be true. The Boring account gives rise to one of the traditional concerns about cp hedges: they are vacuous and therefore useless; all they really say is 'F's cause G's unless they don't'.

Strevens formulations avoids these problems by referring to a specific mechanism, 'M' – the toilet cistern – with respect to which the cp claim is supposed to hold, thus avoiding the problems of both the Empty and the Boring accounts.[20] It can do so because, Strevens supposes, the context and practices for the use of cp laws are sufficient to secure reference to the intended mechanism.[21] Even though the identification of the mechanism is less specific than if it were identified by a list of the individual properties that make it up, we can still refer to the mechanism, he

[18] We suggest that in Strevens' formulation, reference to M should appear first since O and Z are M relative.
[19] Nancy Cartwright and John Pemberton, "Aristotelian Powers: Without them, What Would Modern Science Do?", *Powers and Capacities in Philosophy: The New Aristotelianism*, J. Greco and R. Groff. (eds), Routledge, 93-112, 2013, Section 2.3.4.
[20] Or better, 'mechanism type'.
[21] Much here depends on being able to pinpoint a target mechanism that all/most interested parties have in mind. The sociological circumstances that allow for this must be quite interesting.

maintains. In contrast with the Empty account then, on Strevens' semantics, the cp claim has no free variables,[22] and in contrast with the boring account, it has real content that could well be – and possibly often is – false. Thus unlike these other two, Strevens' account has genuine content. And, we add, it is *useful* content. So long as it is possible to identify reliably enough when M obtains and when it doesn't, the ceteris paribus law can be used both for predicting and for making changes in the world. The trick then is in our ability genuinely to refer to a mechanism that gives rise to the regular behaviour recorded in the cp law. If we cannot, then Strevens account too will be a Boring one, as Julian Reiss warns:[23] 'There is an M (we know not what), such that if it were to obtain, Fs would cause Gs.' Strevens maintains that we can do just this, by baptism. For the sake of pursuing our three central questions, we propose to simply accept Strevens' claim about this, at least for a great many cases. We can readily point to the toilet in our bathroom and assert that by virtue of this mechanism here, pressing the lever flushes the cistern and also to an acorn we hold in our hands, 'By virtue of the mechanism here, putting this in the ground and providing it with warmth, water and light will causes an oak sapling to grow.'

3. Of what use are opaque claims?

We and Strevens are in agreement that cp laws depend on mechanisms and that the mechanism is often 'opaque': we very often do not know what constitutes the mechanism nor how it operates. We can refer to it but we don't know what makes it up. Yet people can rely on vast numbers of these mechanism-relative cp laws, in scientific practice, in engineering and in our daily lives. How is that possible when the mechanism is opaque, when we don't know what it is that constitutes the mechanism we need if the cp behaviour is to obtain?

Strevens tells us about reference but not about use – which is fine for his purposes. He is, after all, aiming for a semantics in which a law claim, by virtue of having the clause 'ceteris paribus' in front, can make a true claim involving reference to something that is not mentioned in the claim itself. Still, we might look to his discussion of reference for help with the practical problem of use. With respect to fixing reference, Strevens notes three conditions that must hold:

> First, there must be a well-defined "baptismal group" of exemplars. Second, there must be an observer-independent "same mechanism as" relation, that

[22] It seems he must also suppose that in referring to M we also can refer specifically to the M-relative enabling and operation conditions it takes for F's to cause G's by way of M.

[23] In email correspondence September 2018, Reiss also worries that many cp laws have a vast array of different mechanisms.

> is, a criterion for individuating mechanisms that is capable of determining facts of the matter about which [systems] do and do not share a certain mechanism…Third, a single mechanism must in fact cause the behaviour of all or almost all of the members of the baptismal group—it must not be the case that there are several different mechanisms, none statistically dominant.[24]

Unfortunately, these conditions, which are required for the reference of *M* to be fixed, are not very helpful about what it takes for *us* to be able to fix the reference, let alone how we can know when we confront a new system whether it is an M or not, even allowing for a degree of uncertainty about this. We might expect an elaboration of Strevens' second criterion to be of more help. About this he says: 'I propose that two phenomena are brought about by the same causal mechanism just in case they have the same causal explanation. The causal facts that matter for the purposes of mechanism individuation are, in other words, the explanatorily relevant facts.' [25]

The second criterion is supposed to help with the conventional problem generated by the first criterion, that a set of exemplars will always have many, many features in common, perhaps indefinitely many. Not all are meant to be necessary for a new system to fall under the term being introduced. In the context here, *M* is introduced as part of the semantics of a cp law: 'CP, L'. It is supposed to be L-relative. This mirrors Cartwright's answer to the question, 'What is a nomological machine for law L?'. Her answer: 'It is a fixed (enough) arrangement of components, or factors, with stable (enough) capacities that in the right stable (enough) environment will, with repeated operation, give rise to the kind of regular behaviour that we represent in [L].'[26] For both Strevens and Cartwright, the relevant mechanism is picked out by what it takes to get the behaviour described in the law.[27] This may work well enough for Cartwright's task, which was to argue that most of the laws we make practical use of are not 'free-standing' or 'God-given', as we may imagine Newton's or Coulomb's law to be. Rather they hold only on account of special arrangements that afford them. It may also work well enough for Strevens' purposes of fixing reference. But it is peculiarly unhelpful in answering the question of what make it possible to use these cp laws to guide our expectations in real life.

Our answer depends on a nice empirical fact; nice, that is, for us humans. Systems that afford regular behaviours like those recorded in cp laws often come with

---

[24] Strevens (n2), pages 21-22.

[25] Ibid, page 25.

[26] See Nancy Cartwright, "Where Do Laws of Nature Come From?" *Dialectica*, Vol. 51, No. 1, 65-78, 1997; Nancy Cartwright, *The Dappled World: A Study of the Boundaries of Science*, Cambridge University Press, 1999, page 50.

[27] This is explicit in Cartwright's account: a nomological machine is 'an arrangement that…gives rise to the kind of regular behaviour that we represent in our scientific laws.' [Ibid, page50.] She notes that for just that reason many will find her characterisation of a nomological machine unsatisfying if we want to explain the regular behaviour by reference to the machine.

observable, sometimes even fairly precisely measurable, *markers*: characteristics peculiar to them that distinguish them from other, different kinds of systems and from mere heaps of parts that do not behave in any systematic ways. So, we can make use of many of our cp laws because very often there are *markers* that pick out the right kinds of systems to generate the law-like behaviour they prescribe.

We learn that various features are markers, without knowing why they are. We also learn *cautions*: what might damage the systems that give rise to the behaviour we want, or want to avoid, whether they need coddling and how to coddle them, and what can make them better at the job. None of this requires us to be able to say how the system picked out by the marker does what it does.

Markers. We can identify that the mechanism is present and at work without any knowledge of its internal workings. We rely on the changing length of days between summer and winter without need of to understand the movements of the earth that are responsible for them. Many mechanisms we construct come with labels that say what you can rely on them to do. Many do not need labels. Toasters and computers and cars all have a characteristic look. So too with naturally occurring nomological machines. It is easy to distinguish nasturtium seeds from acorns even though one may have no idea why planting a nasturtium seed produces nasturtium seedlings and planting an acorn produces baby oak trees. Moreover, we can use our ability to recognise what a nasturtium seed looks like to grow nasturtiums even if nobody knows how they work.

Consider a hypothetical example from political science. Two countries are in disagreement over a variety of issues, and tensions are mounting. Are they likely to go to war? The 'theory of the democratic peace', also called the 'inter-democracy non-aggression hypothesis', gives reason to answer *no* if they are both democracies.[28] Roughly: democracies don't go to war with other democracies. So: by way of the target mechanism 'democracy pair', even when tensions mount between two counties, disagreements will not lead to war. To the extent that the marker 'democracy pair' is reliable, it can be useful in planning not only military policy but also, for example, international investment policy. There are a variety of accounts of just what systemic features might be responsible for the democratic peace. But understanding the details of the systems that afford it is not necessary for prediction: 'democracy' is a relatively easily accessible marker for when non-aggression is likely. As with much in science, the theory is challenged, and much refinement has occurred over the years. Happily, not all the scientific issues matter

[28] For a philosopher's discussion of the democratic peace and for further references, see Sharon Crasnow, "The Role of Case Studies in Political Science Research", *Philosophy of Science*,79(5), 655-66, 2012.

for purposes of prediction.[29] On the other hand, getting a good enough characterisation of when a country is and is not a democracy is essential if 'democracy' is to provide a policy-useful marker for when to bet against outbreaks of aggression.

For a development-centred example, consider what Angus Deaton and Nancy Cartwright say about Conditional Cash Transfers (CCTs):

> Conditional cash transfers have worked for a variety of different outcomes in different places… Think through the causal chain that is required for CCTs [such as those incentivising child education and vaccination] to be successful: People must like money, they must like (or do not object too much) to their children being educated and vaccinated, there must exist schools and clinics that are close enough and well enough staffed to do their job, and the government or agency that is running the scheme must care about the wellbeing of families and their children.[30]

If this is right, finding places where people have a desire for money and they want their children to be educated and healthy, as well as finding that the government in those places tends to the welfare of its citizens can function as markers for identifying where CCTs are likely to initiate the changes desired.

Cautions. We can learn not only when we have a mechanism of the right kind, we can also learn how to recognise when the mechanism is damaged, what to do to protect it, and what not to do if we want it to keep working -- as when we repeatedly tell teenage children that it's a bad idea to spill Coke on their computer keyboard,[31] when we don't expect the battery that is oozing a bit of gooey liquid to work, or we carry our cell phones outside to hunt reception. Or, we know not to bother planting acorns that float after soaking in water for 24 hours. This too underwrites the usefulness of our mechanism-relative cp laws.

Although we do learn of many mechanisms how to protect them, it is important to underline that nomological machines are often *fragile.* This is a point that Strevens makes as well, using the same example that we often cite: the Phillips curve recording the short-term trade-off between unemployment and inflation. [32] If

---

[29] For instance, for purposes of prediction, in this as in most cases, it does not matter whether the association is causal or merely a correlation.

[30] Angus Deaton and Nancy Cartwright, "Understanding and Misunderstanding Randomized Control Trials", *Social Science and Medicine*, https://doi.org/10.1016/j.socscimed.2017.12.005, 2017, page 10.

[31] To be explicit: 'If the keyboard does not have Coke dropped on it and otherwise also operates properly (O), then by way of the keyboard and computer mechanism, pressing the key marked 'A' when the computer is charged up, causes 'A' to appear on the screen.' The other examples can be reconstructed similarly.

[32] See for instance Nancy Cartwright, "How to do Things with Causes", *Proceedings and addresses of the American Philosophical Association.*, 83 (2), 5-22, 2009.

Chicago School economists are right, this cp law arises from an underlying structure in which economic agents have expectations that match the true probabilities and in which they act to maximise their expected utility. Nobel prize winning Chicago School economist Robert Lucas argues that this structure is fragile, and with it the cp law it gives rise to. As soon as the government tries to use inflation as a handle to affect unemployment, entrepreneurs will recognise inflation for what it is; they will not mistake it for a price rise in their domain and will not be moved to expand their enterprises, opening new jobs. In its efforts to use the cp law, the government breaks the very machine that affords it, or so the Lucas story has it. Once the government acts, we have, as Strevens tells us, a new machine.

4. What 'gives rise to' is and isn't

Recall Strevens' individuation criterion for mechanisms: 'I propose that two phenomena are brought about by the same causal mechanism just in case they have the same causal explanation.' We think the reverse. They have the same causal explanation if they are brought about by the same causal mechanism. So we can't avoid the question: 'What is this "brought about by" relation?'. Nor does Strevens' own terminology uniformly avoid mention of this relation. He talks for instance about states of affairs that a mechanism may or may not *cause,* about the mechanism that *causes* the pattern of variation recorded in a cp law and about the symptoms that a mechanism is *responsible for.* This section discusses what this relation isn't, endorses Carl Craver's suggestion of *constitution* as an answer in some cases and offers some truth conditions for it of our own but closes by pointing out that that does not really remove the puzzle.

We talk about M 'affording' or 'giving rise to' the behaviour B that is described in the cp law. This though is not the only terminology in use. For Peter Machamer, Lindley Darden and Carl Craver (widely referred to as '*MDC*'), M is 'productive' of B.[33] According to Stuart Glennan, M 'produces' B.[34] William Bechtel and Adele Abrahamsen say that the operation of M is 'responsible for' B.[35] And Craver and James Tabery in their *Encyclopedia of Philosophy* article, 'Mechanisms in Science' add 'underlying' and 'maintaining'.[36]

---

[33] Peter Machamer, Lindley Darden and Carl F. Craver. "Thinking about Mechanisms", *Philosophy of science* 67: 1-25, 2000, page 3.

[34] Stuart Glennan, "Rethinking Mechanistic Explanation", *Philosophy of Science* 69, S342-353, 2002, page S344.

[35] William Bechtel and Adele Abrahamsen, "Explanation: A Mechanist Alternative", *Studies in the History and Philosophy of the Biological and Biomedical Sciences* 36: 421-441, 2005, page 423.

[36] Carl Craver and James Tabery, "Mechanisms in Science", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/spr2017/entries/science-mechanisms/,Spring 2017, Section 2.1.1.

Independent of what descriptions are used, what are these descriptions supposed to represent in the world? There are at least three ways available of treating the relation between the mechanism and the causal regularity it gives rise to, and each has problems.

Causes 1. First, we might think in terms of the word that slips into Strevens' discussion and that is suggested by the terminology of production, *causes.* This word takes on different guises in different circumstances. Here we think it is not helpful. In what sense does the operation of a mechanism *cause* the causal processes it gives rise to? How does the nomological machine cause F to cause G; how does, for example, *the operation of the toilet mechanism* cause *pressing the flush lever to cause the toilet to flush?* Generally, causes should proceed their effects. But the operation of the mechanism and the causal process it gives rise to are simultaneous. Most causal processes are continuous in time and thus have intermediate steps. Interrupting these is a conventional strategy for preventing an unwanted effect once its cause has occurred. How does that work when the cause is the operation of a mechanism and the effect is the causing of G by F? Also there should generally be a flow of influence from cause to effect, which is the basic idea behind conserved-quantity-interchange accounts of causal processes. Can we identify some influence that the operation of the mechanism passes to the causing of G by F? And often in ordinary cases of causation, we can mark the putative cause -- in this case, that would be the nomological machine, say the toilet mechanism -- and find the mark later on the effect – in this case, the causing of the toilet to flush by pressing the lever. None of these conventional characteristics of a causal relation are easy to find here. So this does not seem a promising starting idea. As Bechtel and Craver remark in discussing top-down and bottom-up causation: '…the phrase "top-down causation" is often used to describe a perfectly coherent and familiar relationship between the activities of wholes and the behaviours of their components, but the relationship is not a causal relationship. Likewise, the phrase "bottom-up causation" does not, properly speaking, pick out a causal relationship.'[37]

Causes 2. A second strategy that assumes M plays a proper causal role is to insist that the cp law is under specified. A full specification puts the mechanism into the antecedent of the law itself: (In O) M and F and Z cause G. Note that this is not the same as Strevens' more roundabout formulation of the content of the law claim, that *by way of* the mechanism M, in O, F and Z causes G. But it is the approach that Judaea Pearl,[38] among others, advocates in his work on causal Bayes nets. But it has a wealth of problems. If it is the operation of the machine that *M* is supposed to represent, then *F* in the antecedent is redundant: if the machine operates, the toilet

[37] Carl Craver and William Bechtel, "Top-down Causation without Top-down Causes", *Biology and Philosophy*, Volume 22, Issue 4, 547–563, 2006, p.547.
[38] Judea Pearl, *Causality, Models, Reasoning and Inference* (Second edition). Cambridge University Press, 2009.

flushes. If it is the parts and their arrangement that *M* represents, are we to think of the parts and the arrangement as a cause? This is what Pearl seems committed to since on his proposals M figures into the causal graph and into the causal equations in just the same way as F. But if M is a cause, we could expect it fairly regularly to have the kinds of characteristic of causes we just described, involving temporal priority to the effect, existence of spatial and temporal intermediaries between cause and effect and flow of influence. But again, it would take some fancy footwork to maintain they are there or explain away the need for them. The use of this proposal in Bayes nets faces the additional difficulty that the nodes in a Bayes net are supposed to be random variables. That means they have a range of allowed values with a probability distribution over them. But what are the allowed values in our toilet example? Any structure, dreamt or undreamt of, that has a lever? And where can the probabilities over these come from?

Constitution. This is advocated by Craver. MDC suppose that mechanisms are made up of organised *entities* and *activities*: 'Mechanisms are entities and activities organised such that they are productive of regular changes from start or set-up to finish or termination conditions.'[39] According to Craver, when the operation of a mechanism explains a phenomenon, this 'variety of explanation is *constitutive* (or componential)'[40]. Craver uses the diagram in Figure 1 to represent the connection between a phenomenon and its mechanism, stating that:

> S's ψ-ing is explained by the organization of entities $\{X_1, X_2, \ldots, X_M\}$ and activities $\{\phi_1, \phi_2, \ldots, \phi_M\}$[41]

Where: S is the mechanism as a whole, ψ is the behaviour of S 'as a whole', the X_i's are the component entities of S, and ϕ_i are the component activities of S (where ϕ_i is the activity of X_i).

---

[39] Ibid, page 3.
[40] Carl Craver, *Explaining the Brain*, Clarendon Press Oxford, 2007, page 8. Original italics.
[41] Ibid, page 7.

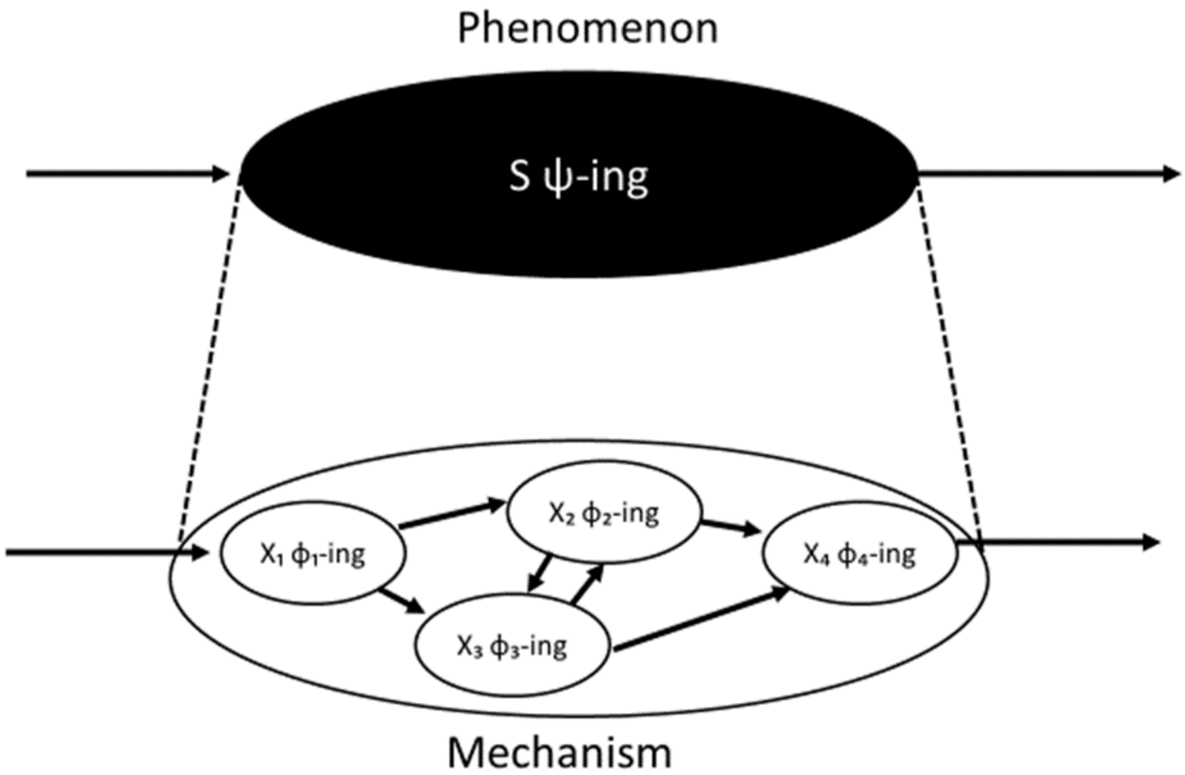

Figure 1: Craver's diagram of a phenomenon and its mechanism[42]

Following Craver, we could suppose that the machines parts doing what they do in consort *constitutes* the behaviour described in the cp law. This constitution account certainly avoids the difficulties facing causation and it seems that it may work for the kinds of cases that Craver often focuses on. In these cases, the phenomenon to be explained is the ψ-ing of a system S, and the explanation is the organised activities (the ϕ_i-ings) of the parts of S at each stage of S's ψ-ing. The canonical example is the neuron transmitting a signal, which on Craver's account just *is*, or *is constituted by*, the organised parts (components) of the neuron and their activities, especially its membrane and gates and the potassium and sodium ions.

But what about the kinds of cases we have been discussing: 'CP, Fs cause Gs' where F and perhaps G are frequently not features of the mechanism but instead are features of inputs and outputs to it. For instance, putting five quarters in the machine and pressing C3 causes a can of Coke to drop out (which is true of a vending machine but not of a parking meter). Or, putting bread in the machine and pressing the lever causes the bread to toast (which is true of a toaster but not a toilet).

---

[42]Ibid, page 7, Figure 1.1.

It should be no surprise, though, that Craver's account does not fit these cases well since what is to be explained is different. Craver's examples are ones where *S* and *M* refer to the same thing and where, as Craver and Tabery put it in the *Stanford Encyclopedia*, 'The phenomenon [to be explained] is the behaviour of the mechanism as a whole.'[43] MDC, Bechtel and other mechanists use the details of the structure and activities of a system to explain how *that system* – e.g., the neuron – does what it does. Cartwright and Strevens are engaged in a different enterprise. They are concerned with the truth of cp laws and with what underwrites their truth. In this case the mechanism is employed to explain how something that is not the mechanism but a single feature, and often not even a feature of the mechanism – an 'F' – causes a different feature – a 'G' – that may also not be a feature of the mechanism.

Consider a typical, well-understood case where a mechanism gives rise to regular behaviour of the kind that can be recorded in a cp causal law: Millikan's famous oil drop experiment to measure the charge of the electron. In Millikan's apparatus, a negatively charged oil droplet hovers between two charged plates, pulled down by gravity and up by electric attraction. Due to air resistance, it also feels a drag force proportional to its velocity. Millikan measured the charge q on the droplet by adjusting the potential difference between the plates till the droplet was at rest, so he could calculate $\mathbf{F}_{electric} = q\mathbf{E} = \mathbf{F}_{earth} \oplus \mathbf{F}_{drag}$. The charge q is due to free electrons on the oil drop, all of which have the same charge q_e. Though the drops differ in charge, for each drop, $q = nq_e$; so q_e can be estimated by measuring q for a number of drops. So we have here a well-attested cp law:

> *Millikan*: CP, adjusting the potential difference in the right way causes the oil drop to be stationary.

The behaviour in the Millikan cp law is given rise to/generated by/afforded by the operation of Millikan's apparatus: 'If the Millikan apparatus operates properly (O), then by way of this apparatus (M), the behaviour recorded in *Millikan* will occur. The adjusted potential difference pulls up on the drop as does the drag of the air and these two forces in the same direction together balance the force of the earth pulling the oil drop in the opposite direction. So the oil drop is subject to no force. Since $\mathbf{F} = m\mathbf{a}$, the oil drop is motionless.

Constitution makes sense in the neuron case and may well be an adequate account for many, perhaps even all, of the cases that Craver and others in philosophy of biology have in view. Typical synonyms for 'constitutes' are 'amount to', 'adds up to', 'makes up', 'composes' and 'comprises'. It seems true that the neuron's parts

---

43 Craver and Tabery (n24), Section 2.1.

doing what they do amounts to/adds up to/makes up/composes/comprises its transmission of a potential difference. It is difficult to see, though, how what is recorded in a good description of the parts of Millikan's apparatus and what they do amounts to/adds up to/etc. the potential difference's causing the oil drop to be stationary. So we look for different account that can cover cases like this and that may be even be found more informative than constitution even where constitution seems to fit.

Here is a suggestion for some truth conditions for our kinds of case that involve constitution. Note though that these do not turn the 'gives rise to/generates/affords' into the constitution relation. We call this the *I-MP-O* account: Input (e.g. F(t)) – Mechanism Process – Output (e.g. G(t')).

> *I-MP-O*. M gives rise to/generates/affords 'By way of the target mechanism M, given O, the conditions Z and F(t) cause G(t')', where F is a feature of a system outside M and G may be so as well, if and only if 1) Given Z and O, F(t) causes a 'starting state' in M, and 2) Given Z and O, that starting state initiates a continuous process in M in which each state is caused by previous ones until a final state is reached which causes G at t' (or in which G is instantiated at t').

This can also double, with simple amendments, for cases where F and G are features of parts of the mechanism itself – so long as one is careful about the demand in clause 1 that F(t) *cause* what gets labelled 'the starting state'. We can use a well-known example of Wesley Salmon[44] to illustrate why clause 1 matters. Consider a mechanism composed of a rotating beacon in the centre of a stadium with a high circular wall. The beacon light can be switched on and off at t. When it is on, a white spot sweeps around the wall. Shall we say, *CP, A white spot at p causes a white spot at p' later?* Surely not. But without clause 1), this cp law is afforded by the operation of the beacon mechanism.

Perhaps our quarrel with Craver's account is just a quibble and our offer of I-MP-O is superfluous. Perhaps we should after all accept that the activities of M constitute the causing of Gs by Fs. The puzzle seems hardly solved however, neither by Craver's proposal nor by *I-MP-O*. In the case of I-MP-O, we still want to know what it is about the mechanism that allows a trigger of F(t) to initiate a process of change in the states of the mechanism that ends in a state in which G(t') obtains or which causes G(t). And in the case of constitution, we still lack an account of *why*. Whenever it is true that x constitutes y, there is a reason that it does so. The kind of reason can

---

[44] Wesley Salmon, *Scientific Explanation and the Causal Structure of the World,* Princeton University Press, 1984, page 141.

vary from case to case. What matters is that it is not arbitrary what constitutes what, or what kinds of things constitute what other kinds.

At the Board of Examiners meeting, the Chair takes role and announces, 'We constitute a quorum.' Why do *we* constitute a quorum? Because 'we' includes the Chair of the Board of Examiners, the Secretary, all three external examiners and five internal members of the Board. That is what the University's *Learning and Teaching Handbook* says it takes to make a quorum. Later at the meeting you raise your hand after a proposal has been discussed. Raising your hand constitutes voting 'yes' to the proposal. It does so because that's the convention at the Examiners' meeting.

Or consider the case of the 17-year-old whose neck was broken when a rugby scrum collapsed. Was he right in claiming that the referee's failure to police the scrum constituted a breach of the referee's duty of care? That's debatable, and indeed it was debated in the British Courts. The young man won the case because, the judge found, the referee had not enforced the safety requirements set out in the Laws of the Game, which contained special provisions about players under nineteen years old, and in particular required front rows to engage in a 'crouch-touch-pause-engage' sequence. The point is that the young man could not just claim that the referee's behaviour constituted a breach of his duty of care; there was, rather, a reason that it constituted a breach: the referee is supposed to enforce those safety requirements and he did not.[45]

The reason of course need not be something written in a rule book. It can, for example, depend on the kind of thing that is to be constituted and what that thing is supposed to do. Why can't a heap of bricks constitute a fence? Because a fence is meant to enclose an area, and a heap of bricks does not do that. 'David Hume' does not constitute a correct answer to 'Who wrote *The Wealth of Nations*?' because Adam Smith, not Hume, wrote *Wealth of Nations*. The parts in the arrangement pictured in the diagram in the design specifications constitute the toaster because that's what makes it up and allows it to do its job.

So, even if the relationship between behaviours in the mechanism and those described in the cp law is taken to be constitution, this still leaves a big unanswered ontological question, parallel to the one we ask. We ask, 'What is it for M's operating to give rise to/generate/afford Fs causing Gs?'. If the answer is that M's operating constitutes Fs causing Gs, what is the reason for that? *Why* do the joint activities of the parts of the mechanism in this particular arrangement constitute this particular behaviour? Our last example is the kind we need to think about for understanding how mechanisms give rise to cp laws. There the answer seems to be that the parts of the toaster behaving as they are designed to constitutes toasting of

---

[45] For a description of this case, see https://www.lawteacher.net/cases/negligence-duty-cases.php.

the bread because when they do what they are supposed to, the bread is toasted, and there is no more to getting it toasted than letting the parts do their job. This answer seems to be a good reason to count the actions of the parts as constituting toasting, but the reason still seems incomplete. *Why* when the parts behave as they are supposed to does the bread get toasted?

The answer we shall propose is that what is true of M is that the parts and their arrangements call into play different general laws at once and make them combine in novel ways they otherwise could not. Although our concern is to discover what the relationship between the mechanism and the behaviour described in the cp laws is in the world -- that is, relations in the material mode -- we propose to begin the hunt in the formal mode, looking at cases where models of mechanisms are used to *explain* cp laws. So we shall next address the epistemological question: 'In what sense does M and its operation explain the behaviours it gives rise to?'.

5. *Explaining 'explains' – the epistemic question answered*

Strevens focuses on opaque mechanisms, ones whose workings we do not understand and which we may only be able to identify by pointing, since his principal aim is argue that this opacity does not make the related cp claim false, meaningless, trivially true, or useless. If, on the other hand, we want to find the relationships between the mechanism and the cp laws it gives rise to, it is best to focus on cases where we know the mechanism and how it operates.

Mechanists tend to see mechanistic explanation as very different from covering-law explanation. For example, Antti Revonsuo, writing under the title 'On the Nature of Explanation in the Neurosciences', claims, 'Explanation in basic neuroscience is a prime example of causal-mechanical explanation *rather than* explanation in terms of universal laws and principles.'[46] Or consider Craver and Tabery, who title the section on explanation in their *Stanford Encyclopedia* article 'Mechanisms in Science: From Formal Analyses to Material Structures', where the formal analysis in question is the covering-law account. They write: 'According to [the covering-law model], explanations are arguments showing that the event to be explained … was to have been expected on the basis of laws of nature and the antecedent and boundary conditions …. Mechanists, *in contrast*, insist explanation is a matter of elucidating the causal structures that produce, underlie, or maintain the phenomenon of interest.' [47] They go on to note a number of concerns expressed by mechanists about covering-law explanation including:

---

[46] Peter Machamer, Peter McLaughlin and Rick Grush (eds.), *Theory & Method in the Neurosciences*, University of Pittsburgh Press, 2000, page 47. Emphasis added.
[47] Craver and Tabery (n24), Section 3 (first para). Emphasis added.

1. its inability to deal with causal /etiological explanation;
2. its inability to distinguish re-descriptions of the phenomenon in general terms from explanations that reveal the mechanism that produces it;
3. its possible lack of depth (subsuming a phenomenon under any true law will count as a complete explanation so that the level of detail may be insufficient for satisfactory explanation);
4. its requirement for laws which may often be unavailable in the biological and special sciences;[48]

Practicing social scientists are also prone to contrast covering-law and mechanistic explanation. For instance, in their classic text *Case Studies and Theory Development in the Social Sciences*[49], Alexander George and Andrew Bennett briefly review the standard philosophy of science literature and argue that mechanistic explanation can solve two problems faced by the deductive-nomological (D-N) account, which along with the Inductive-Statistical (I-S) account is the standard formulation of covering-law explanation. The first is the problem of distinguishing 'between causal and spurious regularities'[50]. This problem is akin to 1. above. The 'second problem with the D-N model is that its predictions must be rendered with perfect certainty', a problem which, they argue following Wesley Salmon, I-S version does not successfully solve. This adds another to the list of concerns about covering-law vis-à-vis mechanistic explanation:

5. For covering-law explanation, outcomes are supposed to be fixed.

We do not see such a contrast between mechanistic and covering-law explanation. We urge rather that the best model for the explanations there, including those for nomological machines, is the old covering-law model.[51] Standard mechanistic explanations are not separate from covering-law explanations but are, rather, a subset of them. And a good number of exemplary covering-law explanations are equally exemplary mechanistic ones. Kepler's laws are deduced from Newton's laws, including the general principle that **F**=m**a** and the bridge principle that an object of mass m located **r** from another mass M experiences a force GMm/**r**^{2}; the cp law

---

[48] Ibid, Section 3.1.

[49] Alexander George and Andrew Bennett, *Case Studies and Theory Development in the Social Science*, MIT Press, 2004. See also Derek Beach and Rasmus Brun Pedersen, *Process Tracing Methods Foundations and Guidelines*, The University of Michigan Press, Ann Arbor, 2013, chapters 3 and 4.

[50] Ibid, page 132.

[51] Note that Cartwright (Nancy Cartwright, "Models: The Blueprints for Laws", *Philosophy of Science*, Vol. 64, S292-S303, Dec.1997) too, like Strevens, claims that we use models to explain cp laws. As I.F. da Cunha (2018) explains, 'Cartwright differentiates nomological machines from models, which she conceives of as the blueprints of the nomological machines. That is, models specify the ceteris paribus conditions that nomological machines must present in order to display the behaviour predicted by the laws.' (Ivan da Cunha, "Constructing Dystopian Experience: A Neurath-Cartwrightian Approach to the Philosophy of Social Technology", *Studies in History and Philosophy of Science*, https://doi.org/10.1016/j.shpsa.2018.05.012, 2018, page 2.)

called 'the Phillip's curve' (that in the short term rising inflation reduces unemployment) is deduced, as we noted, by Chicago School economists in a 'rational expectations' model from the general principle that agents act to maximise their expected utility and the bridge principle that an entrepreneur's utility in the setting modelled is constituted by the firm's profits; the cp law that the rooster flaps it's wings, spreads it's feathers, and crows while the three kings above bow to the Virgin Mary and Child is deduced in models of the great Three Kings Clock of Strasbourg from the law of gravity, the laws of simple machines, and bridge principles that link location on the earth's surface to being subject to the pull of gravity, rigid rods resting on fulcrums to levers, inelastic cables passing over free-turning low-friction wheels to pulleys and so forth.

Perhaps the feeling of contrast results from focusing on a certain simple subspecies of covering-law explanations that do not, at least on the face of them, invoke mechanisms. For instance: 'Why does this neuron transmit messages?'. Because: 'CP, all neurons transmit messages'. Perhaps the role for the parts and their arrangements is not transparent in the description of covering-law explanations, since these get lumped under the expression 'antecedent and boundary conditions'. Or perhaps these boundary conditions are conceived of too simply. For instance: to explain Kepler's 1st law, that the planets travel in elliptical orbits with the sun as one of the foci, we use Newton's $\mathbf{F} = m\dot{\boldsymbol{v}}$. The boundary conditions include the value of m and an initial value of **v**. Given these we can solve the differential equation to get the elliptical orbit. That does not look like a mechanical explanation. But of course, far more is necessary. We have to fill in **F**. For that we need to know the structure of the mechanism. It is common in presenting this Newtonian explanation to begin with a diagram like Figure 2, which pictures a simple nomological machine made of just two parts, a large object and a small object, arranged some distance apart. Their relevant features are the masses of the two objects, M and m, their separation **r,** and the relative velocity of the small mass with respect to the larger, which has a component $\dot{\boldsymbol{r}}$ along **r** and $r\dot{\theta}$ perpendicular to **r**. Because of the features of the objects and their arrangement, the larger one pulls on the smaller with a force $GMm/\mathbf{r}^2$. Now we can construct a proper, filled-in differential equation. Of course, if we start our explanation with that filled-in equation, the role of the parts and their features, arrangements and activities will not be apparent.

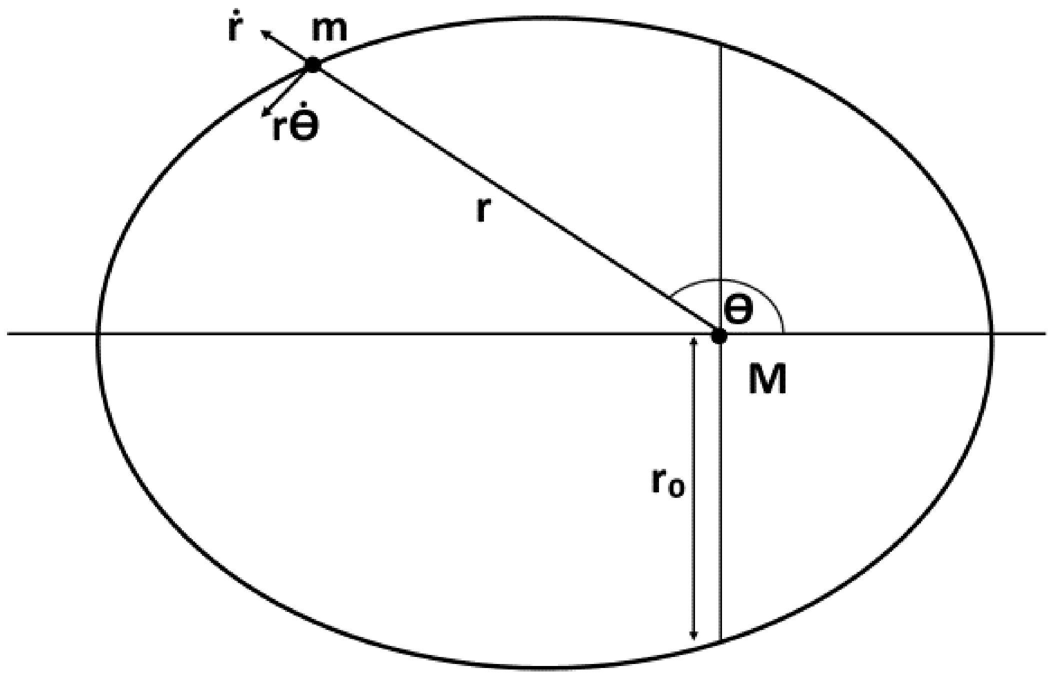

Figure 2: Elliptic orbit of small mass around large mass

Objections 1,2 and 3 adumbrated by Craver and Tabery are thus not relevant to our thesis. We do not claim that any derivation that satisfies the general covering-law demands can do the jobs they call for. But rather, if mechanistic explanations can do these jobs, as they argue, then so can covering-law explanations since, we claim, standard mechanistic explanations are generally a subspecies of covering-law explanations. Nor does objection 5 bear on our claims here for we do not suppose that covering laws are all either 'deterministic' or statistical. Rather many of the central covering laws used in mechanistic explanations are 'tendency laws' that tell what a cause contributes to the effect, not what overall effect actually happens, as in the law of gravity, Coulomb's law and the law describing the drag of the air in the Millikan experiment described above.[52] This leaves objection 4.

Another reason that others perceive a contrast where, we argue, the correct relation is species/sub-species may be due to a doctrine that has often sat alongside the covering-law model: that the business of science is discovering general laws. That is decidedly not the business that much of biology is in, or so argue Bechtel and

[52] For more on tendency laws see, for example, John Stuart Mill, "On the Definition of Political Economy", The Collected Works of John Stuart Mill, Volume IV - Essays on Economics and Society Part I, ed John Robson, Toronto: University of Toronto Press, 309-40, 1967; or Nancy Cartwright, *Nature's Capacities and Their Measurement*, Oxford University Press, 1989, chapter 4.

Robert Richardson [53], who are among the founding fathers of the mechanistic account of explanation. Revonsuo provides a neat summary of their view:

> Additional support for this view of the nature of biological explanation comes from Bechtel (1994), who argues that biological knowledge is not primarily represented in universal laws or linguistic structures. Biologists typically first identify an interesting system at one level of organization in nature and then try to figure out what the components of this system are, how they interact, and how they produce the effects that can be observed at the level of the whole system. When they go about this task, they try to take the system apart or visualize it better with the help of various research instruments in order to figure out what the components and microstructures of the system are like. From these data biologists attempt to build an idealized model of the system, the purpose of which is to show the general structure and function of the system. The model may be only partially (if at all) clothed in linguistic representations; instead; all kinds of diagrams and figures can often best depict the component structures of, and their mutual interactions with, the biological system in question…[54]

We can readily agree with Bechtel and Richardson that the advances in biology they note have little to do with the discovery of new general laws and almost everything to do with uncovering the structure of systems they have identified as biologically interesting. That has no bearing on whether or not general laws play a central role in the models biologists construct of how those systems operate to do what they do. However the model is presented, with diagrams and figures (analogous perhaps to Figure 2), or with equations, or narratives, or whatever, why should we believe that structures that match the model can do what they are supposed to? Why is it true that the model that pictures just those components in those arrangements explains 'the effects that can be observed at the level of the whole system'?

The answer, we propose, is that these effects are just what is to be expected given the features of the parts in that arrangement and the covering laws in which these features figure. The effects are just what is to be expected because that is what must happen if all those features act as they should under the general laws that govern them. The trajectory of the oil drop in Millikan's experiment is different from that of the earth going around the sun. But in both cases the masses act in accord with the law of gravity. The other features too – like the charge or the resistance of the air –also act in accord with the general laws that govern them. Their joint

---

[53] William Bechtel and Robert Richardson, "Emergent Phenomena and Complex Systems", *Emergence or Reduction? Essays in the Prospects of Non-reductive Physicalism*, Ansgar Beckermann, Hans Flohr and Jaegwon Kim (eds.), 257-88, de Gruyter, 1992.

[54] Antti Revonsuo, "On the Nature of Explanation in the Neurosciences", *Theory and Method in the Neurosciences*, Peter Machamer, Peter McLaughlin & Rick Grush (eds.), University of Pittsburgh Press, 45-69, 2001, page 48.

actions, in accord with all these general laws at once, explain why the system does just what it does.

It may be thought, 'Yes, but physics examples are the easy ones. It is no surprise that physics, with its rich tool-kit of general principles, uses covering laws in its mechanistic explanations. What about elsewhere?'. We maintain that physics is not special here. Examples of mechanistic covering-law explanations in the socio-economic sciences, described as such, can be found in Cartwright's 1995.[55] These include a money multiplier and a debt-generating mechanism. Here we shall look at one of the mechanists' own favourite examples from biology: signal transmission in the neuron. Figure 3 illustrates a few of the parts and stages involved in such transmission.

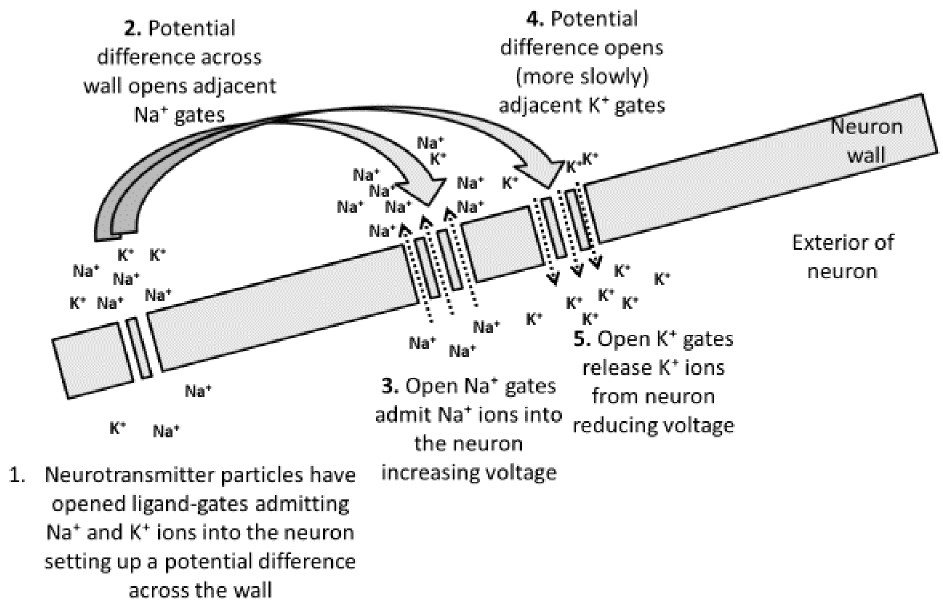

Figure 3: Some steps in signal transmission in a neuron

Beginning students are typically told, 'Signals *within* neurons are transmitted electrically, however signals *between* neurons are transmitted chemically across the synapse'[56]. We take it that this means that these signals are transmitted in accord with well-known laws of physics and chemistry. Here is the basic mechanistic explanation for the electromagnetic transmission in the neuron.

> *Basic NT explanation*: On both the outside and the inside of the neuron sit positively charged sodium and potassium ions. In the rest state there is somewhat more positive charge on the outside of its membrane than inside,

[55] Nancy Cartwright, "Ceteris Paribus Laws and Socio-economic Machines", *The Monist*, 78 (3):276-294, 1995.
[56] Cara Flannigan, Dave Berry, Matt Jarvis and Rob Liddle, *AQA Psychology*, Cheltenham, Glouc: Illuminate Publishing, 2015, page 117.

> so that the voltage measured from the inside is slightly negative. Key features of the neuron are sodium-selective and potassium-selective gates in its wall which open or close in response to certain stimuli. Arriving neurotransmitter particles dock with receptors opening sodium and potassium gates (ligand-gates), allowing ions to enter the neuron, thus increasing the positive charge in the cell and the local voltage across the cell wall. If the voltage passes a given threshold, that stimulates an adjacent sodium gate to open; sodium ions grouped outside the neuron flood through the open gate due to the electro-chemical gradient. The change of charge distribution increases the local potential difference across the wall, opening the next adjacent gates. In the meantime, the open sodium gates close quickly stopping the rise in voltage. The stimulus also opens a slower-to-open potassium gate and positively charged potassium ions flow out because the membrane is now more negatively charged on the outside than on the inside, so that the voltage drops. The potassium gate then closes, and sodium and potassium pumps then restore the rest state, so that the process can be repeated with the advent of a new stimulus.

There is of course more to be said. For instance, concerning the diffusion forces affecting the flow of the sodium and potassium ions. But these too behave as they should, according to standard diffusion equations. Or, how do the gates open and close? A helical protein string embedded in a pore in the wall of the neuron features an uneven charge distribution and is contorted by this potential difference in a way that leads to the opening of the passage. Again, this is in accord with what is to be expected given the basic laws of electromagnetics – which is not surprising since signal transmission across the neuron is modelled *electrically*. Beyond that, one might next explain how the protein is structured that allows it to contort as it does. That is likely not to use electromagnetic principles. But for satisfactory explanation it should use general principles that hold not just in the proteins in neuron gates, but elsewhere as well.

Central to the basic NT explanation is Coulomb's law of electromagnetic attraction and repulsion, which played a pivotal role in the Millikan apparatus that we described earlier: like charges repel each other and opposites attract. Coulomb's law is even made use of in the same way in the two mechanisms. Millikan shifts the ratio of positive to negative charge on the two plates in his apparatus to create a voltage difference between them. In the neuron the open gates allow the ratio of positive to negative charge on the outside and inside of the membrane to shift thus adjusting the voltage difference across it.

But, one might object, this similarity does not dispel the contrast between mechanical and covering-law explanation since the explanation of how the oil drop comes to a stop in Millikan's apparatus is itself a mechanistic explanation. We agree. It is both mechanical *and* covering-law. It is the fact that it is covering-law that makes it undoubtedly an *explanation* and not a mere description of what happens. We understand why the explanandum behaviour occurs given the structure of the mechanism because that is the behaviour that must occur if all the laws we cite, which we take to be true, are to be true – and thus not violated in this mechanism.

We do not want to be dictatorial about the term 'explanation' though. There may be cases both in biology and elsewhere where the features displayed by the parts of a mechanism do not obey the general principles inside the mechanism that they do outside it. That is one sensible thing one could mean by the claim that the behaviour that the mechanism gives rise to is *emergent.* We do not want to deny that there may be emergent behaviour in this sense. Describing what is going on in the mechanism when this behaviour occurs is certainly a contribution to knowledge; perhaps it is reasonably called 'explanation'. What we want to stress is that, by far and away, most of the satisfying mechanistic explanations available in both natural and social science are covering-law explanations. It is true that some covering-law explanations are not mechanistic explanations. The Hodgkin–Huxley formalism often cited by mechanists or the differential equations describing transitions between neuron states in a Markovian scheme are good examples. But, to repeat, that some covering-law explanations are not mechanistic does not show that most mechanistic explanations are not covering law.

Supposing we are right that the covering-law model is generally in play in good mechanistic explanations. That answers our *epistemic question*. Does that get us any further forward with our *ontological* one? What is happening, not in our representations but in the world? Coulomb's law helps explain both why adjusting the ratio of charges between the two plates causes the oil drops in Millikan's experiment to come to a halt and why the arrival of neurotransmitter particles at the front end of a neuron causes the release of neurotransmitter particles at the other end. Causes explain their effects, but surely we do not want to claim that Newton's laws cause Kepler's to be true nor that Coulomb's causes the cp behaviours that the Millikan and neuron mechanisms give rise to. But if not that, then what? Recall, if we opt for I-MP-O, we still need to know what it is about M that means that F(t) starts off a process of changes of state in M that ends in G(t'). Or: If you decide the answer should be 'constitution', what is the reason that those parts acting in just those ways constitute these behaviours?

6. The ontological question answered

We can almost read the ontological answer from the epistemic one, and we can see it as involving constitution, but not in the way that Craver pictures it. Work a good while ago by Adolf Grunbaum[57] gives a clue as to how. Newton’s laws explain Kepler’s because Kepler's laws are what Newton's *amount to* in the context of the planetary system. In the language of Section 5, the behaviour described in Kepler's laws *constitutes* the obtaining of Newton's laws given the arrangement of the planets and the sun. Travelling in the elliptical orbit prescribed by Kepler's laws just is what it is for a planet to do what Newton's laws dictate in the presence of the sun. So we suggest this condition, which covers not only the behaviours described in cp causal laws, which have been our main focus, but can apply for non-causal regularities as well:

> *Affording.* Suppose that behaviour B (e.g. Fs cause Gs) occurs in conditions Z and O if mechanism M (characterised by parts P, arrangement A, and features ϒ) operates. M = <P,A,ϒ > gives rise to/generates/affords B if, for some ϒ ’⊆ ϒ and general principle G(ϒ’) governing features in ϒ’, all the principles in G(ϒ’) are instantiated in B’s occurring in Z and O.

Pemberton argues that mechanists tend to pay insufficient attention to arrangements.[58] Arrangements matter crucially here because they confine how general principles are instanced. For example, Towfic Shomar[59] models an arrangement in which two charges attract each other yet the one moves away from the other, and in part on account of that attraction; if it did not, Coulomb’s law would be violated in that arrangement. The arrangements play two roles in doing so.

First, arrangements introduce new features that parts do not have by themselves. A good strong branch, or a shovel, balanced over a rock or a log becomes a lever, which obeys the law of the lever, as levers do wheresoever a lever is found, whether with its end wedged under a wheel to heave a car out of the mud or functioning as a seesaw in a park. When there is excess charge on one plate compared to the other in Millikan’s experiment, there is a potential drop across the plates. A similar but different arrangement with some positive charge removed from one plate to the other will exhibit a different voltage, sometimes one different in sign, which is what happens when excess sodium ions are located inside the neuron membrane rather than outside. So, by virtue of the arrangement, new features obtain and new laws are called into play, and perhaps others become irrelevant because the features they govern disappear from the parts when they are so arranged.

---

[57] Adolf Grunbaum, philosophy of science undergraduate lectures, University of Pittsburgh circa 1963.

[58] John Pemberton, *Integrating Mechanist and Nomological Machine Ontologies to make Sense of What-How-That Evidence*, http://personal.lse.ac.uk/pemberto, 2011.

[59] Nancy Cartwright, *The Dappled World: A Study of the Boundaries of Science*, Cambridge University Press, 1999, Figure 3.1a on page 60 and Figure 3.1b on page 61.

Second, arrangements fix which activities happen when: which happen together and which after which. Millikin's calculation supposes that the pull of the earth, the electromagnetic attraction, and the drag of the air all happen at once, which is right because the drop is falling through the air, in close vicinity to the earth just while the voltage is being adjusted. In the neuron, the gates nearest the incoming impulse open first, then the next ones along, so that the gates open sequentially along the length of the axon, which is typically long and thin. One could imagine a different shaped membrane with a more symmetric arrangement, which would give rise to very different behaviour in the mechanism.

Together this means that the arrangements are crucial to what general laws are instantiated in the mechanism and what the behaviour will be when they are all instantiated in the same process. This is why arrangements play such a central role in mechanistic explanation. But they do so precisely because mechanist explanations rely on covering laws.

Return now to the issue of constitution. Recall Section 5. We suggested that when M gives rise to '*CP F(t) causes G(t)*', where F is a feature of a system outside M and G may be so as well, what is happening is that F(t) causes a 'starting state' in M that initiates a continuous process in M in which each state is caused by previous ones until a final state is reached which causes G at t' (or in which G is instantiated at t'). To this we now add that for these relations to hold among the states of M is what it is for the general laws that apply to M's features and their actions all to be instanced in this causal process.

We suppose that the general laws relevant to the features of the mechanism determine the behaviour of its parts (and the parts of the parts, etc.) within each arbitrarily short time period. These laws may be expressible by differential equations (such as the force laws of physics; the pushing of one object on another, e.g. gas pressure); or laws of heating, compressing (e.g. laws concerning coefficients of restitution), stretching, distorting, retarding (e.g. laws concerning friction), dissolving, diffusing. Other relevant laws may be expressible in qualitative terms, e.g. laws governing the cutting of a knife. When the mechanism operates normally, these laws obtaining simultaneously for all the parts (and parts of parts) together in their given configuration determine the behaviour of the salient parts of the mechanism in each arbitrarily short time period and hence the continuous behaviour of the mechanism through time. Together they are the reason that the initial state causes the final state and hence that the stimulus F that causes the starting state can truly be said to cause G later.

Consider again the example of neuron transmission. Here we may take F(t) to be the arrival at t of neurotransmitter particles at the head of the neuron and G(t′) to be the triggering of the release of neurotransmitter particles from the synaptic vesicles at the end of the neuron. The arrival of neurotransmitter particles (F(t)) causes ligand-gated ion channels in the neuron to open. This is part of the starting state of the neuron. Other important aspects of the starting state are that the voltage-gated channels are closed and there are more potassium ions inside the neuron than outside and conversely with sodium ions. The neuron then exhibits a continuous, orderly sequence of states over time, sparked by F(t), crucial among them being ones which exhibit a potential difference above the threshold, which will cause the first sodium-gated channel to open, which causes later states in which others are open, in turn producing a neuron state in which there is a large potential difference. So the action potential travels down the neuron's axon to the presynaptic terminal at the end. That in turn causes G(t′).

Although in Section 6 we focused on Coulomb's law, in this process we see activities in which a number of different well-established general laws are instantiated together, for instance:

a) A cloud of particles contained by a wall in which there is a gate which is open (closed) can enter (not enter) the gate and cross the wall.
b) A (net) force on a free-moving particle accelerates/moves it in the direction of that force.
c) A distribution of charges gives rise, via the Coulomb law that we have focused on, to forces on local charged particles.
d) A flexible object subject to differential forces on differing parts contorts.

These general laws are derived from our broader empirical experience, not from our study of the neuron. Although these laws are familiar and unremarkable, they are central to the operation of the neuron, as to many other mechanisms. These general laws apply to the parts of the neuron as follows:

- The sodium gate of the neuron allows (prohibits) the passage of sodium ions (law a).
- The being open of the gate is an instance of the contorting of a flexible object (law d) (here the *flexible object* is the helical protein embedded in the sodium gate).
- The sodium ions are (i) the *local charged particles* subject to a force (law c), (ii) the *free-moving particles* (law b), (iii) the members of cloud of particles that enter (do not enter) the gate (law a) and (iv) components of the *distribution of charges* (law c).

We see that, during each arbitrarily short period of time, the pertinent general laws being true together of the neuron's parts in their given arrangement at that stage determines the behaviour of the parts and hence the mechanism as a whole at that stage and thereby the obtaining of an orderly sequence of states over time.