

# An Explanatory Taste For Mechanisms

## Abstract

Mechanistic explanations, according to one prominent account, are derived from objective explanations (Craver 2007, 2014). Mechanistic standards of explanation are in turn pulled from nature, and are thereby insulated from the values of investigators, since explanation is an objectively defined achievement grounded in the causal structure of the world (Craver 2014). This results in the closure of mechanism's explanatory standards—it is insulated from the values, norms and goals of investigators. I raise two problems with this position. First, it relies on several ontological claims which, while plausible, fail to guarantee the objectivity of mechanistic explanatory standards to the degree of certainty required. Second, Craver's position itself introduces a value-laden explanatory standard—the 3M requirement (Kaplan & Craver 2011)—which undermines the closure of explanatory standards. I show how in practice mechanistic explanation is in part guided by explanatory taste, shorthand for background contextual values that influence our standards of explanation. Mechanism often has a particular pragmatically-oriented taste for control, and gerrymanders explanatory standards in order to obtain it. I conclude by arguing that objectivity, rather than being obtained through the right set of explanatory standards, is better thought of as the result of processes of intersubjective criticism, which renders visible the contextual values of communities of investigators and allows them to be controlled for (Longino 1990).

# I Introduction

The goal of separating the explanatory wheat from the merely descriptive chaff is a common feature of accounts of scientific explanation: what must a description do to graduate to the status of explanation? More often than not this task involves the codification of *standards of explanation*—the rules by which we judge explanatoriness. Historically, different accounts have required that explanations subsume observations under laws of nature (Hempel & Oppenheim 1948), unify phenomena (Kitcher 1989), or describe the underlying causal structure of the world that lead to a phenomenon’s occurrence (Salmon 1984). Mechanism, my target in this paper, fits into this last category.

Originating with Craver (2007) one of the more influential strands of Mechanism argues that mechanistic models become explanations when they correctly describe the mechanism responsible for a phenomenon. According to this account—a strain of *ontic mechanism*, sometimes called *mechanistic realism* (Craver & Kaplan 2011)—descriptions become explanations when they correspond to the real mechanisms, the *objective explanations*, that occupy nature (Craver 2007, 2014; Craver & Kaplan 2018)<sup>1</sup>. Explanations already exist in the world in the form of mechanisms; it is the scientist’s job to describe them. Descriptive models therefore become explanations by dint of how closely and completely they represent this mechanistic explanation-in-the-world.

Craver (2014) extends this ontic picture by suggesting that by virtue of its objectivity, Mechanism is *the* guide to a cross-cultural, universal and objective means of causally explaining phenomena, at least within the special sciences (Craver 2014). There is consequently no possibility for explanatory standards to differ across times and places, across sociocultural contexts and with respect to the values and goals of investigators, other than as deviations from the correct mechanistic standards (Craver 2014). I call this position the *closure of explanatory standards*.

This line of thinking appears to motivate Craver’s (2014) declaration of independence:

...my topic is independent of psychological questions about the kinds of explan-

---

<sup>1</sup>Important to note here are the different ways in which Craver (2007, 2014) uses the term *mechanistic explanation*. The term can refer to a text that accurately characterises the real mechanistic explanation-in-the-world, or it can refer to that explanation-in-the-world itself. I will show in the coming sections that for Craver these two senses are nevertheless intimately linked, such that the standards of the former fall out of the objectively-known facts we can glean about the nature of the latter.

ation that human cognitive agents tend to produce or tend to accept. Clearly, people often accept as explanations a great many things that they should reject as such. And people in different cultures might have different criteria for accepting or rejecting explanations. These facts (if they are facts) would be fascinating to anthropologists, psychologists, and sociologists. But they are not relevant to the philosophical problem of stating when a scientific explanation ought to be accepted as such. In the view defended here, scientific explanation is a distinctive kind of achievement that cultures and individuals have to learn to make. Individual explanatory judgments, or cultural trends in such, are not data to be honored by a normative theory that seeks to specify when such judgments go right and when they go wrong. (Craver 2014, pg. 29)

Mechanistic standards of scientific explanation, on this view, are insulated from sociocultural, psychological and other influences traditionally viewed as non-scientific—otherwise known as *contextual values* (Longino 1990). These contextual values are seen by this variant of ontic mechanism as an obstacle to the task of building an account of scientific explanation. In Craver’s (2014) view, while the role of contextual values in science may indicate interesting vectors of research regarding how scientific explanation is done by value-driven investigators, and is potentially “fascinating to anthropologists, psychologists, and sociologists” (Craver 2014, pg. 29) it tells us nothing about what explanations actually are or (crucially) how they *ought* to be done. For Craver, the norms of explanation fall out of an understanding of the objective causal structures of nature, and so allowing contextual values to influence our standards of explanation would result in the incorporation of factors that are “not relevant to the philosophical problem of stating when a scientific explanation ought to be accepted as such” (Craver 2014, pg. 29).

With Craver’s version of the ontic mechanistic account in mind, I pick out two potential lines of criticism. First, I probe at the view that mechanistic standards of explanation are objective: do the underlying ontic claims support it, and give us good reasons to believe that mechanistic explanations are objective? Second, I test the closure of explanatory standards: do we have good reasons to believe that mechanism’s standards of explanation are thereby insulated

from contextual values?

I argue on both counts there are substantial reasons to think not, at least based on the arguments offered by the ontic mechanist camp so far. Craver's position depends on a sequence of metaphysical claims providing not just plausibility, but a guarantee that nature is mechanistic, such that nature offers us objective mechanistic standards of explanation. While plausible, the assumption that mechanisms are real things underlying phenomena fails to provide the required guarantee. I also argue that the model-to-mechanism-mapping (3M) requirement endorsed by ontic mechanists as an additional explanatory standard (Kaplan & Craver 2011, Kaplan 2015) undermines the closure of explanatory standards, since in advancing 3M in order to disqualify non-mechanistic models from counting as genuinely explanatory, they incorporate an explicitly normative, value-laden standard of explanation into the supposedly value-free ontic mechanismism.

I contend that what this breakdown in the closure of explanatory standards reveals is the value-laden character of mechanistic standards of explanation—not founded on ontological truths about nature, but rather on the contextual values common to the mechanist project. I propose that the means by which investigators often formulate and select successful explanations is entangled with a particular *explanatory taste*. Explanatory taste is shorthand for the contextual values that guide the gerrymandering of standards of explanation in order to serve instrumental ends. The chosen standards allow us to set explanatory end-goals that satisfy the needs of a project: in the case of mechanism, this is at least partly the *control* of phenomena with pragmatic biomedical ends in mind. In accordance with this goal, mechanists elevate descriptions which afford this kind of control to the status of explanations and exclude those sorts of descriptions that do not satisfy this need.

Consequently the view that contextual values have nothing to say about the correct standards of explanation is, I think, misguided. More importantly for actual scientific practice, rejecting examination of the contextual values that influence a scientific project would be a grave error. Objectivity in science is born out of intersubjective criticism and evaluation allows for those values to be brought to the fore, examined, and controlled for (Longino 1990).

A point of clarification: my argument chiefly concerns a particular position on mechanistic

explanation developed most notably in Craver (2007, 2014), Craver & Kaplan (2011), Kaplan & Craver (2011) and Craver & Kaplan (2018). I do not intend to imply that all mechanists adopt or are compelled to adopt this position. I will say more in the following section about the differences between this position and that of other mechanists who either reject the ontic account, or simply do not explicitly endorse Craver's version of ontic mechanism.

## **2 Mechanist explanation and objectivity**

Mechanistic explanation is a popular account of causal explanation, with a particular focus on the cognitive and biological sciences. The basic explanatory unit in a mechanistic explanation is a *mechanism*, and its associated *phenomenon*. A mechanism is defined as "...a structure performing a function in virtue of its component parts, component operations, and their organization." (Bechtel & Abrahamsen 2005, pg. 423). It is this mechanism which produces, underlies or is otherwise responsible for the phenomenon.

Core to mechanistic explanation is the notion that explanations ought to speak to causes; they "must include reference to causal relationships if they are to distinguish good explanations from bad" (Craver 2007, pg. 8). Similarly important (and popular, though not universal) in the contemporary mechanist literature is its appeal to interventionism (Woodward 2003) as a means of locating these causal relations. In brief, if an intervention on the value of X causes a change in the value of Y (all other things remaining equal) then it describes a causal relationship. These variables can stand in for various features of the world: the cell became cancerous because it mutated; the ion channel opened because of a potential difference across the membrane, etc. Describing these causal connections, which form the causal structure of the world, is ultimately what describes a mechanism.

### **2.1 The ontic claims & the aetiology of objectivity**

Why does a description of a mechanism become an explanation? Some explanatory standards are required to answer that question. To that end, Craver's position aims to present an objectively derived set of explanatory standards. According to this position, Mechanism does not

explain nature merely in the sense that its models are illuminating, enhance understanding or are useful to human endeavours:

“Objective explanations are not texts; they are full-bodied things. They are facts, not representations” (Craver 2007, p. 27)

Good mechanistic explanations are explanatory because they capture objective explanations, those mechanisms or objective explanations-in-the-world (Craver 2007). Explanations are meant to reveal the true mechanistic workings of nature that are waiting out there to be discovered; they shed light on objective ontic structures (Craver 2014). The standards of mechanistic explanation are derived from an understanding of these objective ontic structures, and as such mechanistic explanations inherit this objectivity. The position Craver seeks to outline requires a sequence of supporting claims to advance the notion that mechanistic standards of explanation track objective explanations. We can infer the core claims and their role in supporting Craver’s closure of explanatory standards from his statements around the ontological status of mechanisms and what this status affords us; together they suggest that nature is mechanistic, composed of mechanisms, and our explanations should track this mechanistic reality.<sup>2</sup>

(1) Mechanisms are real

Mechanisms are real entities that really do produce or underlie the phenomena we observe in nature. They are not merely useful fictions constructed as a part of scientific investigation, but “things in the world” (Glennan & Illari 2018a, pg. 2). The view that mechanisms are real, and that this is significant for their explanatory potential, is broadly shared by prominent ontic mechanists, for instance Glennan & Illari (2018b):

“...let us consider whether and in what sense these varieties of mechanisms are real.

Our language has been realist. Mechanisms are things in the world...” (pg. 99).

Glennan’s (2017) account of mechanism presents a similar view, where “...mechanisms and their constituents are things in the world that exist independently of the models we make of

---

<sup>2</sup>The ontic mechanist account is however not divorced from consideration of epistemic virtues: an explanation may be more or less easily grasped, be more or less computable, and be more or less easily communicated to others, and these are worthwhile considerations when producing explanatory texts (Craver & Kaplan 2020). Nevertheless, considerations of these virtues do not bear on the question of what makes for an explanation that satisfies mechanistic explanatory standards.

them” (pg. 10). Talking about mechanisms in cognitive science, Craver & Kaplan (2011) claim that ontic mechanists “expect to ground the taxonomy of cognitive and neural phenomena in objective facts about the mechanistic structure of the brain. By learning which mechanisms are distinct from which others, one can carve the mind–brain at its joints.” (pg. 276).

This claim is important because it functions as a lynchpin of objectivity—it guarantees that mechanisms aren’t merely forms of description, fictions or idealisations, but real things, and when we find them, we find something real. This also points to something of a schism within mechanism between the ontic mechanists, and the epistemic mechanists (Illari 2013). These subdivisions support similar methods and explanatory standards, but provide different justifications for them. Epistemic mechanists (Bechtel & Abrahamsen 2005; Bechtel 2015) are more concerned with how mechanistic explanations enhance understanding of phenomena; representing the ontic side, Craver (2007) is concerned with explaining by revealing the objectively existing causal structure of the world. The ontic mechanists are strict realists about mechanisms, while epistemic mechanists do not consider the ontological status of mechanisms to be so important for articulating an account of mechanistic explanation.

As a consequence epistemic mechanists are criticised precisely because this view allegedly abandons anything except epistemic commitments, including objectivity (Craver 2014; Kaiser 2018). Because epistemic mechanism doesn’t claim to identify real mechanisms, “it is too weak to serve as a guide to the norms that distinguish good explanations from bad and complete explanations from incomplete.” (Craver 2007, pg. 28)

## (2) Mechanisms are explanatorily obligatory

Following from (1), ontic mechanists also accept that mechanisms are not just one kind of explanatorily relevant thing among many that inhabit the world. Mechanisms are (at least within the “special sciences”) precisely what is relevant to causally explaining phenomena in those domains. There may be other things that figure in a mechanistic ontology, but the explanatorily relevant things are mechanisms (Craver 2014). Put another way, mechanisms are necessary for causal explanation, at least as far as cognitive and biological sciences are concerned.

This position allows non–mechanistic models to be utilised towards this end of describing mechanisms, since they are still ultimately means of getting at mechanisms (Kaplan 2015); on

the other hand they fail at being explanatory if they are not mapped or grafted to a mechanism in some way. Consequently I take ontic mechanists to hold the position that mechanisms are explanatorily obligatory: descriptions of mechanisms, as well as supporting non-mechanistic models (like mathematical models) that help elucidate features of mechanisms, are necessary and sufficient for our objective explanatory purposes.

(3) Objective explanatory standards can be derived from nature.

If mechanisms really do produce or underlie phenomena, and descriptions thereof are explanatorily adequate, this then affords mechanists a straightforward way of obtaining standards of explanation. Mechanistic standards of explanation must be geared towards revealing these explanations in the world:

The complete constitutive explanation for a given explanandum, in this ontic sense, includes everything relevant (that is, everything that makes a difference) to the precise phenomenon in question. (Craver & Kaplan 2018, pg. 14)

Complete neuroscientific explanations are distinguished from incomplete explanations...by the fact that complete explanations capture all of the relevant causal relations among the components in a mechanism. (Craver 2007 pg. 61–62)

A key mechanistic standard of explanation is *completeness*. Completeness, according to ontic mechanists, is not a subjective or pragmatic criterion but rather a reference to a state of affairs in nature—what counts as complete is whatever correctly represents the objective mechanism comprehensively enough. Whatever allows us to get at the objective explanation waiting in the world, describe the relevant causal and constitutive relations, and exclude irrelevant relations, are consequently the objective standards of explanation:

Good mechanistic explanatory texts...are good in part because they correctly represent objective explanations. Complete explanatory texts are complete because they represent all and only the relevant portions of the causal structure of the world. (Craver 2007, pg. 27)



This condition of relevance also demonstrates the connection between completeness and the aforementioned notion of realism about mechanisms that ontic mechanists subscribe to. When a feature of the world is ruled to be relevant to an explanation, this means it really is a part of the mechanism producing the phenomenon of interest.

Hence completeness is the definitive mechanistic explanatory standard. A text can be in the process of attaining completeness, or presently be complete. In either scenario it meets this standard. Interventionism makes this standard achievable, since it permits investigators to find what features are relevant—and hence required for a complete description of a mechanism—and which features of the world are irrelevant. Completely describing (accurately) the causal structure of the world—a mechanism—underlying the explanandum phenomenon is what graduates a description to an explanation.

(4) Mechanistic explanations are objective explanations.

If nature really is full of mechanisms, mechanisms are what we need to describe in order to explain, and our ways of explaining them are derived from the objective causal structure of nature, then mechanistic explanations ought to explain objectively.

I group (1)—(4) together as the *ontic claims*.

## **2.2 Closure of explanatory standards, and a declaration of independence**

One consequence of this account is the *closure of explanatory standards*. If mechanistic explanatory standards are objectively derived from nature, then there is no room for other factors to determine those standards. The contextual values—goals, biases, background beliefs of communities of investigators and so on—simply have nowhere to get any purchase. Explanatory standards are already determined by objective, universally available facts about nature, whether we like it or not (Craver 2014).

Having established closure, Craver (2014) makes his declaration of independence, according to which the question of what explanations are (and what explanatory standards support them) is insulated from context. Contextual values are, in Craver’s view, not relevant to how

modes of explanations are formulated. Mechanisms are really out there producing phenomena and provide us with objective standards of explanation for those phenomena. People may mistakenly believe that context—the goals and values they want an account of explanation to support, for instance—can be decisive to the shape explanations should take, but if Craver is to be believed, they are objectively wrong.

To make the position even starker, in Craver’s view no features of an investigator or community of investigators can have any bearing standards of explanation, since mechanistic explanations and explanatory standards are stable across contexts. Anyone interested in explaining nature (at least where nature is composed of mechanisms) must, to do the job properly, locate certain timeless, universal truths about mechanistic explanation in order to explain. Whether I am investigating the world in prehistory before the written word, or tapping on a keyboard in the 21<sup>st</sup> century the same standards apply; the only proper standards for genuine explanation are mechanistic ones. Including anything else inserted into our explanatory standards serves only to muddy the waters.

### **3 Critique of mechanistic objectivity**

This ontic claims raise some questions. First of all, Craver’s (2014) claim about the closure of explanatory standards is phrased as a sort of guarantee. What underpins this guarantee of closure are the ontic claims. So, to be certain about the closure of explanatory standards, we need to be certain of the undergirding ontic claims. If the ontic claims are taken to really reflect *what is*, distinct from human pragmatic concerns, there must be some good reasons to believe the constituent claims.

It is important here to reiterate the direction of the justification for closure. The ontic mechanist claim is not that the explanatory successes of mechanism should make us confident of this mechanist ontology. The reasoning is the reverse: because we are already so certain about the ontic claims, we can derive from them a set of explanatory standards that guarantee objectivity; “the norms of scientific explanation fall out of a prior commitment on the part of scientific investigators to describe the relevant ontic structures in the world.” (Craver 2014, pg. 41) The ontic claims are the bedrock for mechanistic explanatory standards.

The immediate problem with this position is that it requires assumptions—the ontic claims—to work as a guarantee. Assumptions, I take it, are not guarantees, and so this “prior commitment” to a mechanistic ontology fails to provide the necessary guarantee of objectivity. But perhaps the ontic mechanistic account has the resources to offer further reasons to believe that a mechanistic ontology falls out of a mechanistic science.

Imagine the scenario that (1) were not known with certainty. If that were the case, then it seems that the mechanist conclusion—that mechanistic standards of explanation track objective explanations—would be far less certain. Unless we have good reasons to believe that nature is indeed made up of mechanisms, then there remains the plausible case where nature is not made up of mechanisms. In other words, if (1) were brought into doubt, the set of ontic claims undergirding the objectivity of mechanistic explanatory standards would also be in trouble. This is because if nature isn’t really full of mechanisms producing phenomena, then there is no guarantee that striving for complete descriptions of mechanisms is necessary for explanation, nor that this would lead us to an objective explanation. To be clear, these claims may still be of epistemic value if either were shown to be uncertain on the ontic side of things.

I labour the point in order to make the stakes quite clear. The position being put forth by Craver puts a large stock in its objectivity, upon which its norms are based. Most importantly, (though I leave this point for a later section) such a failure would undermine the closure of explanatory standards.

With these stakes in mind, I will review the evidence for these supporting claims, focusing mainly on (1). Let us also bear in mind the fairly high standards that have been set here by Craver’s position. It is not just supposed to be likely, possible, or even just plausible that (for instance) mechanisms are real and explanatorily adequate. Claims (1) through (4) are meant to be such a certainty that mechanism can lock out all other considerations that might influence the correct explanatory standards.

### **3.1 Testing the mechanist toolkit**

Craver (2007) holds to a straightforward realist position in service of the ontic claims, specifically (1)—if we can observe entities engaged in causal relationships via ideal interventions, then

we can take those things to be real. For us to have reason to think mechanistic parts are real, “they should be robustly detectable, they should be able to be used for the purposes of intervention, and... they should be physiologically plausible” (pg. 161-2). The obverse of this claim is that if we cannot empirically observe something through an intervention, it does not exist, and it is a *how-possibly fiction*:

There is no evidence that souls or entelechies exist...We cannot intervene with predictable outcomes to change souls and entelechies...we are justifiably suspicious of claims that such things exist. But none of these reasonable criteria fails for higher-level items in neuroscience. Molecules, neurons, brain regions, and brain systems all clearly satisfy these standards.(Craver 2007, pg. 15)

As Craver indicates here, interventionism provides confidence about the existence of mechanisms, since it gives us fairly clear criteria for when we have or have not observed entities engaged in a causal relationship. Being able to describe nature in terms of fine-grained relationships between causally- and constitutively-connected variables ought to reveal to us the true underlying ontic structures in nature, which will turn out to be mechanistic.

So, when paired up with (1), interventions seem to give us a good idea of how to go about establishing what mechanistic entities are really out there. However, my aim here is to see if we have good reasons to believe (1) independent of the assumption that (1) is the case. Could the mechanist toolkit of ideal interventions (and I bundle this together here with Craver’s additional constraining criteria of stability, robustness and physiological plausibility) provide a good reason to believe, assumptions aside, that mechanists do indeed have the objective facts about nature, out of which fall the objective mechanistic standards of explanation?

### **3.2 Underdetermined ontology**

One initial concern is the plausibility of interventions alone delivering a mechanistic ontology. One could imagine interventions as a net that can be cast out into the world, which will bring in neutral, objective (and mechanistic) facts about nature, and which subsequently reinforce the idea that the ontic structure of the world is mechanistic (Craver’s statement about souls

and entelechies seems to veer in this direction). But the interpretation of that data gained via interventions as confirming or disconfirming a particular ontology, be it of neurons and brain regions or souls and entelechies, is decidedly not neutral in a way that would establish mechanistic realism.

Craver & Kaplan (2011) themselves raise this substantial problem with a hypothetical intervention-generated ontology within the domain of cognitive science. They note that pre-existing theory and assumptions are decisive to how investigators interpret incoming data. For instance, the interpretation of interventions in the form of double dissociation experiments has permitted cognitive neuroscientists to localise particular brain functions within distinct brain regions. However, these findings depend on several assumptions, e.g. that the brain is functionally modular, and of the validity of the standard taxonomy of mental functions. Craver & Kaplan (2011) concede this potentially leads to a kind of circularity where “our ability to carve the brain into distinct mechanisms requires some idea of what those mechanisms do, and this requires some commitment about which capacities require explanations.” (pg. 76)

This acknowledges a basic problem one would encounter in trying to argue for mechanistic realism, which resembles Quine’s (1951, 1975) problem of underdetermination: data alone is not enough to objectively determine which of many possible interpretations of that data is correct, and we are only really able to arrive at an interpretation of data by accepting a body of preliminary assumptions. While hypothesising underlying mechanisms corresponding to the causal relations observed is certainly a plausible interpretation of said data, one could provide other equally plausible non-mechanistic interpretations by simply holding to different background assumptions.

For instance, Craver’s account assumes a close connection between interventions and mechanisms. For anyone occupying this position, it follows that an intervention revealing a causal relationship also reveals (a portion of) a mechanism. But while interventionism is amenable to describing causal mechanisms, there is no necessary connection between the two (see for instance Woodward’s (2013) discussion of non-mechanistic explanations)—what connects them and what thereby does much of the heavy ontological lifting, is the background assumption that causal relations revealed by interventions correspond to mechanisms. As Craver &

Kaplan’s (2011) double dissociation example demonstrates, pre-existing assumptions are what binds together interventions and mechanisms.

Hence without already assuming (1), interventions are insufficient to confirm or disconfirm the existence of particular entities in an objective manner. Since we are always entering into investigations of nature with hypotheses and theories that propose a certain state of affairs for testing (for instance, that nature is full of causal mechanisms, or that the brain is modular), the interpretation of interventions is always done through such a prism of background assumptions.

For example, one can find causes without mechanisms, so long as one’s interpretation of intervention data is non-mechanistic. Dynamical models, for instance, can provide more abstract, non-mechanistic descriptions of the behaviour of natural systems that can be intervened upon to establish causal relations (Meyer 2020a, 2020b). If the sole criterion of Craver’s position for establishing the existence of something is that a part of the world can be intervened upon, then the dynamics of a system ought to be just as much a part of our ontology. This would mean that a non-mechanistic part of the causal structure of the world can do explanatory work, and would suggest that part holds an equally plausible claim to be real ontic as more traditional mechanistic entities.

Take for instance the oft-cited case of the Haken–Kelso–Bunz (HKB) model of bimanual coordination (Haken et al 1985). This dynamical model describes and predicts with great accuracy how bimanual coordination (in this instance, wagging the index fingers on both hands simultaneously) falls into in-phase and anti-phase patterns depending on the frequency of the oscillations. The model is described by the following differential equation:

$$\frac{d\phi}{dt} = -a \sin \phi - 2b \sin 2 \phi$$

Where  $\phi$  is relative phase, and the ratio  $b/a$  represents the inverse of the frequency of bimanual oscillations. Intervening on  $b/a$  causes changes in  $\phi$  and hence the HKB model describes a causal dependency between these variables (Meyer 2020a). Consequently, this causal relationship forms part of a causal, dynamical explanation of bimanual coordination. Here we seem to meet Craver’s realist requirements: since intervention data is all that is required to confirm that

an entity belongs in our ontology, then these dynamical variables should be granted the same ontic status as, say, the opening or closing of ion channels during the neuronal action potential.

However, this would surely be counted as an unwanted reification of merely mathematical variables from the interventionist-mechanist perspective inhabited by ontic mechanists. One reply might be to suggest that there is something more metaphysically fundamental or real about mechanistic components as opposed to a dynamical property of a system; another reply could involve adding new criteria to ensure that only canonical mechanistic entities can be granted this ontological status via interventions.

The former response has quite a challenge ahead of it. Mechanists hewing to this line of argument have to show how objects like neurons and brain regions are “more real” than dynamical features of a system without appealing to the outcomes of interventions. The latter response has proven the more popular one, defended in the form of the 3M requirement, which I will address later in Section 4.

None of this is to say that interventionism cannot be used to investigate what kinds of entities might exist in the world, only that interventions alone could not provide us with objective causal facts which we can then add right into the corpus of a mechanistic ontology, and subsequently interventionism provides little support for (1).

With the toolkit available to ontic mechanists unable to do much here, Craver’s (2014) declaration of independence seems to be on shakier foundations: now that the claims undergirding the objectivity of mechanism are shown to be more a set of assumptions than ontological certainties upon which an objective mode of explanation can hang, the closure of explanatory standards can be broken open.

## 4 The 3M requirement

I have just raised a few concerns with the guarantee (or lack thereof) of ontic mechanism’s undergirding assumptions. In this section, I will raise another distinct problem with the account. So far, I have dealt with that ontic account’s reliance on uncertain ontological claims for objectivity. Here I investigate how the ontic mechanist account undermines its own claim to objectivity by incorporating a normative explanatory standard—the 3M requirement.

## 4.1 The requirement

The model-to-mechanism-mapping (3M) requirement grew out of debates surrounding the explanatory power of non-mechanistic, dynamical models in the cognitive sciences. Dynamacists claimed that the descriptive and predictive successes of dynamical models elevated them to the status of genuine explanations (Stepp et al 2011, Chemero & Silberstein 2008). They have appealed to various arguments to this effect, including the equivalence of prediction with explanation, and adherence to a covering-law model of scientific explanation (Walmsley 2008). More recently Meyer (2020a, 2020b) uses the very same interventionist method core to mechanistic explanation as an argument for non-mechanistic, causal dynamical explanations in cognitive science and biology.

3M is a critical response to this fairly heterodox push for variations on mathematical, dynamical explanations (Kaplan 2011, Kaplan & Craver 2011). It stipulates that, in order to be an explanation, any model must show how its features (for instance, the variables in a dynamical model) “map onto” the components of an underlying mechanism. The assumption expressed here is that any causal relationship in a dynamical model actually maps to a causal relation present in the underlying mechanism:

(3M) A model of a target phenomenon explains that phenomenon to the extent that (a) the variables in the model correspond to identifiable components, activities, and organizational features of the target mechanism that produces, maintains, or underlies the phenomenon, and (b) the (perhaps mathematical) dependencies posited among these (perhaps mathematical) variables in the model correspond to causal relations among the components of the target mechanism.” (Kaplan 2011, pg. 347)

By introducing 3M, Kaplan & Craver (2011) create an additional explanatory standard that works to exclude potential dynamical (or otherwise non-mechanistic) explanations that meet the existing interventionist requirements to be a causal explanation. No non-mechanistic model alone, at least within certain domains of science, can provide us with genuine causal explanations along interventionist lines if 3M is to be believed. This presents a serious challenge



to the possibility of causal yet non-mechanistic explanations. If any apparent dynamical explanation necessarily maps to a mechanistic explanation, then dynamical explanations are not an explanatory alternative and are subsumed under mechanism.

However, 3M is also interesting in that Kaplan & Craver (2011) explicitly acknowledge its normative character. They claim that 3M “is justified in part because it makes sense of scientific-commonsense judgments about the norms of explanation” (Kaplan & Craver 2011, pg. 602):

The 3M constraint is the mechanist’s gauntlet: a default assumption that the phenomena of cognitive and systems neuroscience have mechanistic explanations, like so many other phenomena in the special sciences, and that cognitive and systems neuroscientists ought to (and often do) demand that explanations reveal the mechanisms underlying the phenomena they seek to explain. (Kaplan & Craver 2011, pg. 603–604)

Their case against the dynamicists, they claim, is justified because dynamicists who defy or reject 3M fail to meet the “scientific commonsense” judgement that good explanations “correctly identify features of the causal structures that produce, underlie, or maintain the explanandum phenomena.” (Kaplan & Craver 2011, pg. 607). It is this appeal to the normative, the “commonsense”, that comes back to bite Craver’s closure of explanatory standards.

## **4.2 The problem of 3M**

Recall the reasoning behind the closure of explanatory standards: because of their objective character, mechanistic standards of explanation are immune to all sociocultural and otherwise normative considerations. These kinds of judgements only reflect the contextual values of investigators and cannot be admitted to the genuinely objective set of explanatory standards.

Yet, 3M is just this kind of thing. It is a mechanistic explanatory standard that Craver’s account adopts, advanced as part of the ontic mechanist project, and it is explicitly normative in character. Its justification is decidedly contextual—as Kaplan & Craver (2011) point out when they suggest that 3M is justified as an additional explanatory standard because it represents the judgements or expectations of a community of scientists and philosophers of science; it helps exclude potential explanations that do not suit their values or goals.

It seems difficult, then, to reconcile Craver's (2014) declaration of independence with his account's endorsement of  $3M$ . On the one hand, it is claimed that mechanistic standards of explanation are objectively derived from nature, immune to normative criteria. On the other, Kaplan & Craver include a normative criterion with the same ability to determine explanatory power (or lack thereof) as the existing explanatory standards. Even if there were strong reasons to believe that mechanistic standards of explanation were objective, the inclusion of  $3M$  would still serve to undermine confidence in that claim.

## 5 Explanatory taste

The view that objectivity is a bulwark against the intrusion of contextual values into mechanistic standards of explanation is not well-founded. The use of uncertain ontic claims to guarantee the certainty of the closure of explanatory standards, coupled with the normative character of  $3M$ , undermines the ontic mechanist guarantee that mechanistic explanations are objective explanations, outside of time and place.

How to understand these mechanistic explanatory standards, then, if not as spawned straight from objectivity? To that end, I introduce the notion of *explanatory taste*. The term refers to the taste—emphasising its contextual nature—of investigators both scientific and philosophical for certain sorts of explanation. To this end, philosophers of science gerrymander explanatory standards in order to include into explanatory canon those things they see as proper explanations, and exclude those they do not in accordance with the particular goals and assumptions of the project they are contributing to. I use the term taste not to suggest arbitrariness, but to emphasise the preference-based (rather than objective) selection of scientific aims, out of which sets of explanatory standards fall. My account here thereby contrasts strongly with Craver's view—I am arguing that explanatory standards are not fruitfully thought of as objective by themselves, but rather are primarily contrived in order to make contributions to particular scientific and related philosophical projects.

To understand what is considered explanatory versus non-explanatory—what investigators have a taste for and what they do not—it becomes necessary to look at the practical uses investigators have in mind for explanations. What contribution are the investigators attempt-

ing to make, and for what material purpose? The demarcation line for explanations is typically drawn via explanatory standards along pragmatic lines, namely the perceived end goal of a given mode of inquiry into nature. To briefly return to a historical case, Hempel's (1965) account of covering law explanations was engineered to provide the desired *understanding* of phenomena, where understanding was equated with having shown how events in nature were to be expected, where that expectation is provided by a deductive or inductive argument. Logical positivists or empiricists like Hempel had a particular explanatory taste that reflected the broader goals of their project: understanding nature in terms of a series of verifiable statements.

### 5.1 The contextual values of mechanism

The ontic mechanist accounts of Craver and Kaplan discussed so far, similarly, are concerned with explanations that serve particular purposes, though theirs differ from those motivating the covering law model. They are interested not with arguments that provide understanding (though that may be a virtue of mechanistic explanations), but with explanations that yield practical control in the biomedical sciences and better our ability to control biology and cognition:  $\mathcal{M}$  reveals explicitly the “commonsense judgements” that motivate the gerrymandering of explanatory standards to this end. By contrast dynamical explanations are judged the wrong sort of thing to be genuinely explanatory since they are not to mechanist tastes for control, and  $\mathcal{M}$  is therefore summoned to bar the gates against a dynamical encroachment on explanation.

In this section I argue, quite to the contrary of the ontic mechanist position I have discussed so far, that mechanism is very much motivated by the contextual values of investigators—their goals, interests and assumptions. In fact, the centrality of various contextual values to mechanistic explanation is a pervasive theme. Consider a few exemplary quotes from Craver (2007):

One way to justify the norms that I discuss is by assessing the extent to which those norms produce explanations that are *potentially useful for intervention and control*.

While this is not the only touchstone that one might use, it is nonetheless one, and it is objective.” (Craver 2007 pg. x, emphasis added)

...I develop a view of explanation that does justice to the exemplars of explanation

in neuroscience and to the standards by which these explanations are evaluated.”  
(Craver 2007, pg. 1–2)

If neuroscience succeeds in this...goal, it will open medical possibilities that now seem like science fiction, and it will provide human beings (for good or ill) with *new and powerful forms of control over the human condition.*” (Craver 2007, pg. 2, emphasis added)

...unlike some areas of fundamental physics, the search for neuroscientific explanations is driven by goals of treating illnesses, improving brain function, preventing cognitive decline, and developing new ways to manipulate and record from the brain in the laboratory. *Explanation is a tool for determining how to intervene into the brain and manipulate it for our various purposes.*” (Craver 2007, pg. 38, emphasis added)

To repeat a central theme: causal relevance, explanation, and control are intimately connected with one another. This is particularly true in biomedical sciences, such as neuroscience, that are driven not merely by intellectual curiosity about the structure of the world, but more fundamentally by the desire (and the funding) to cure diseases, to better the human condition, and to make marketable products.”  
(Craver 2007, pg. 93)

There are clear themes running through Craver’s rationale for an account of mechanistic explanation: concerns about human control over biomedical phenomena; pragmatic concerns about efficacy in medicine; the onerous matter of obtaining funding; justification of certain hypotheses for pragmatic ends; a confidence in neuroscience and what neuroscientists consider to be canonical explanations.

Most telling about the explanatory taste of mechanism is the way Craver speaks about explanation and the pragmatic goal of *control*. Mechanistic explanation is, as Craver says above, not driven merely by “intellectual curiosity” but by a need to control phenomena “for our various purposes”; “[s]uch explanations are useful precisely because they identify loci in a mechanism that can be commandeered for the purposes of control.” (Craver 2007 pg. 200). This

notion of control, of explanation as an inspection that yields various levers for the fine-grained manipulation of nature, is a mechanist contextual value which finds itself reflected in mechanistic standards of explanation. Control is valued—it is to the explanatory tastes of ontic mechanists—precisely because it is perceived to serve these pragmatic ends.

This specific focus on control evident in Craver’s approach reflects a more general interest in control across mechanism, both historical and contemporary. The original mechanism of early modernity was in part valued by its champions because “...it put control of phenomena, at least in principle, in human hands rather than in control of independent vital spirits, and it made understanding accessible to all rather than only to those who had special sympathetic abilities” (Longino 1990 pg. 97). More recently interventionism (Woodward 2003)—which forms the core of Craver’s (2007) mechanist account and has been widely adopted amongst other neo-mechanists (i.e. Bechtel 2017; Glennan 2017)—has at the centre of its conception of explanation a notion of control: “explanatory relationships are relationships that are potentially exploitable for purposes of manipulation and control.” (Woodward 2003 pg. v). So while the importance of control sits front and centre in Craver’s account, it is also a motif in Mechanism more broadly that runs deep and broad. Hence while my focus is on Craver’s position, this discussion of the contextual values motivating Craver’s account of explanation is also relevant to other neo-mechanistic accounts.

How precisely does control motivate the specific explanatory standards of mechanism, then? I argue that the explanatory standard of *completeness* results from this taste for control—not from an objective grounding in a mechanistic reality. Completeness is born of the contextual values of mechanist-oriented investigators: specifically, the goal of control over a phenomenon. This connection between the two notions has not gone unacknowledged:

The pragmatic import of developing norms of mechanistic completeness links to the fact that mechanistic explanations often provide the rationale for developing technologies for gaining control over phenomena, such as experimental techniques and medical treatments. (Baetu 2015, pg. 779–780)

Craver’s position wants explanations that afford control over certain phenomena in certain ways. To that end, mechanists of this strain have developed standards of explanation, specifically

completeness, designed to equate control over nature with explanation.

## 5.2 What becomes of objectivity?

I have given reasons to reject the notion that mechanistic explanations are objective explanations, at least in the manner prescribed by ontic mechanists. In doing so I have deferred to a standard view of objectivity, where the objectivity of science means value–freedom, impartiality and the absence of perspective. Objectivity in this sense “is bound up with questions about the truth and referential character of scientific theories, that is, with scientific realism” (Longino 1990, pg. 62). Objectivity is achieved when science accurately reveals the real world, when “it is a correct view of the objects to be found in the world and of their relations with each other.” (Longino 1990, pg. 62); its criteria for theory selection and success are “nonarbitrary and non-subjective” (Longino 1990, pg. 62). This is the notion of objectivity that seems to line up closest with Craver’s own view.

There is however another notion of objectivity that is detached from these commitments. On this alternative view of objectivity, it is accepted that science cannot be purged of the prior commitments or perspectives of investigators—science is not value–free, impartial and absent perspective, and so “it would be a mistake to identify the objectivity of scientific methods with their empirical features alone” (Longino 1990 pg. 75). For example, in the case of ontic mechanism, it is a mistake to think interventions provide objectivity. Between data and hypothesis—intervention and mechanism—there remains an interpretive gap where values determine the sorts of hypotheses that are selected, and the standards by which they are selected.

Instead objectivity can be thought of as an achievement founded on intersubjective criticism. When the commitments of a community of investigators are subjected to public critique from different critical perspectives, where those perspectives represent communities holding to different contextual values, it makes those commitments visible and hence evaluable:

When...background assumptions are shared by all members of a community, they acquire an invisibility that renders them unavailable for criticism. They do not become visible until individuals who do not share the community’s assumptions can provide alternative explanations of the phenomena without those assump-

tions, as, for example, Einstein could provide an alternative explanation of the Michelson–Morley interferometer experiment. Until such alternatives are available, community assumptions are transparent to their adherents. In addition, the substantive principles determining standards of rationality within a research program or tradition are for the most part immune to criticism by means of those standards. (Longino 1990, pg. 80)

A critical alternative brings to light the contextual values of the community. It does not purge science of those values but makes them explicit and open to discussion and revision. Objectivity is a condition that arises when these biases have been evaluated, criticised, and controlled for. Hence objectivity is achieved not in spite of the inherently value-laden character of scientific inquiry—manifest in those features that Craver (2014) suggests might be “fascinating to anthropologists, psychologists, and sociologists”, but irrelevant to objective scientific explanation—but precisely because of it.

Mechanist criticism of dynamical explanation (Kaplan 2015) has illuminated the otherwise invisible commitments of dynamicists—for instance their explanatory taste for prediction—and spurred reappraisal, further proof that “[c]riticism is thereby transformative” (Longino 1990, pg. 73). Abandoning Craver’s closure of explanatory standards is necessary to permit such criticism of the values of mechanism to be aired out. The alternative account of dynamical explanation I have discussed also serves as a criticism of many conceptual commitments of mechanists, showing how the standards used to separate genuinely explanatory mechanistic models from merely descriptive dynamical models fail to do so. While the ontic mechanist could respond that no matter the criticism, ontic mechanistic explanatory standards are objectively correct, this would be a misstep.

There is another reason beyond, to again borrow Craver’s phrase, intellectual curiosity, as to why should we be particularly concerned about this process of criticism. Longino’s (1990) discussion is motivated by the very real problem of biases in science that entrench or reinforce social inequalities, for instance conceptual problems in research on the biology or cognitive basis of sex and gender differences, as well as criticism thereof by feminist philosophers of science. Making the assumptions of researchers explicit through external, public criticism, generated by

a community who do not share those assumptions, preserves the successes of science while also subjecting its assumptions to illuminating criticism. These criticisms can and have driven science to produce less biased and more conceptually sound research (Longino 1990). For instance, feminist critique of sex determination literature in cell biology shows how previously unanalysed background assumptions about “active maleness” versus “passive femininity” pervading said literature curtail a more comprehensive understanding of the phenomenon (Beldecos et al 1988). My point is not to accuse contemporary Mechanism of holding to pernicious biases, but rather to caution against eliminating the possibility of fair and open criticism which acts as prophylaxis against them.

## **6 Conclusion**

In this paper I have criticised the claim that mechanistic standards of explanation are closed to criticism because they are objectively derived from nature. To this end I targeted the underdeveloped justification for this objectivity which fails to guarantee objectivity as required. I also showed how the closure of explanatory standards is undermined when ontic mechanists introduce the normative explanatory standard of  $\mathfrak{M}$ . I have argued that mechanistic standards of explanation are influenced heavily by explanatory taste; the contextual values and goals of investigators. A prime example of explanatory taste at work is the mechanist taste for control, in pursuit of which they gerrymander mechanistic standards of explanation, introducing the standard notion of completeness. I also suggest that, per Longino (1990), criticism between communities holding to different contextual values is a prerequisite for objectivity, and hence mechanism would be better served by acknowledging the contextual nature of its explanatory commitments.

## **Compliance with Ethical Standards**

The author has no relevant conflicts of interest to disclose.



## References

- Baetu, T. M. (2015). The completeness of mechanistic explanations. *Philosophy of Science*, 82(5), 775–786. <https://doi.org/10.1086/683279>
- Bechtel, W. (2017). Explicating top-down causation using networks and dynamics. *Philosophy of Science*, 84(2), 253–274. <https://doi.org/10.1086/690718>
- Bechtel, W. & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 421–441. <https://doi.org/10.1016/j.shpsc.2005.03.010>
- Beldecos, A., Bailey, S., Gilbert, S., Hicks, K., Kenschaft, L. & Niemczyk, N. (1988). The importance of feminist critique for contemporary cell biology. *Hypatia*, 3(1), 61–76.
- Chemero, A. & Silberstein, M. (2008). After the philosophy of mind: Replacing scholasticism with science\*. *Philosophy of Science*, 75(1), 1–27. <https://doi.org/10.1086/587820>
- Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience* [OCLC: ocm77573093]. Oxford : New York : Oxford University Press, Clarendon Press.
- Craver, C. F. (2014). The ontic account of scientific explanation (M. I. Kaiser, O. R. Scholz, D. Plenge & A. Hüttemann, Eds.). In M. I. Kaiser, O. R. Scholz, D. Plenge & A. Hüttemann (Eds.), *Explanation in the special sciences: The case of biology and history*. Springer Verlag.
- Craver, C. F. & Kaplan, D. M. (2011). Towards a mechanistic philosophy of neuroscience (S. French & J. Saatsi, Eds.). In S. French & J. Saatsi (Eds.), *Continuum companion to the philosophy of science*.
- Craver, C. F. & Kaplan, D. M. (2018). Are more details better? on the norms of completeness for mechanistic explanations. *The British Journal for the Philosophy of Science*. <https://doi.org/10.1093/bjps/axy015>
- Glennan, S. (2017). *The new mechanical philosophy*. Oxford University Press.
- Haken, H., Kelso, J. A. S. & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics*, 51, 347–356.

- Hempel, C. G. (1965). Aspects of scientific explanation. *Journal of Symbolic Logic*, 37(4), 747–749.
- Hempel, C. & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15(2), 135–175.
- Illari, P. (2013). Mechanistic explanation: Integrating the ontic and epistemic. *Erkenntnis*, 78, 237–255. <https://doi.org/10.1007/s10670-013-9511-y>
- Kaiser, M. I. (2013). *Explanation in the special sciences: The case of biology and history*. New York, Springer.
- Kaplan, D. M. (2011). Explanation and description in computational neuroscience. *Synthese*, 183(3), 339–373. <https://doi.org/10.1007/s11229-011-9970-0>
- Kaplan, D. M. (2015). Moving parts: The natural alliance between dynamical and mechanistic modeling approaches. *Biology & Philosophy*, 30(6), 757–786. <https://doi.org/10.1007/s10539-015-9499-6>
- Kaplan, D. M. & Craver, C. F. (2011). The explanatory force of dynamical and mathematical models in neuroscience: A mechanistic perspective. *Philosophy of Science*, 78(4), 601–627.
- Kitcher, P. (1989). Explanatory unification and the causal structure of the world (P. Kitcher & W. Salmon, Eds.). In P. Kitcher & W. Salmon (Eds.), *Scientific explanation*. Minneapolis: University of Minnesota Press.
- Longino, H. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry* (Vol. 25). Princeton University Press.
- Machamer, P. (2004). Activities and causation: The metaphysics and epistemology of mechanisms. *International Studies in the Philosophy of Science*, 18(1), 27–39. <https://doi.org/10.1080/02698590412331289242>
- Meyer, R. (2020). The non-mechanistic option: Defending dynamical explanations. *The British Journal for the Philosophy of Science*. <https://doi.org/10.1093/bjps/axy034>
- Quine, W. V. (1951). Main trends in recent philosophy: Two dogmas of empiricism. *The Philosophical Review*, 60(1), 20. <https://doi.org/10.2307/2181906>
- Quine, W. V. (1975). On empirically equivalent systems of the world. *Erkenntnis*, 9, 313–328.

Salmon, W. (1984). *Scientific explanation and the causal structure of the world*. Princeton University Press.

Stepp, N., Chemero, A. & Turvey, M. T. (2011). Philosophy for the rest of cognitive science. *Topics in Cognitive Science*, 3(2), 425–437. <https://doi.org/10.1111/j.1756-8765.2011.01143>.

x

Walmsley, J. (2008). Explanation in dynamical cognitive science. *Minds and Machines*, 18(3), 331–348. <https://doi.org/10.1007/s11023-008-9103-9>

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford University Press.

Woodward, J. (2013). Mechanistic explanation: Its scope and limits. *Aristotelian Society Supplementary Volume*, 87(1), 39–65. <https://doi.org/10.1111/j.1467-8349.2013.00219.x>