The Case for Regularity in Mechanistic Causal Explanation

Holly Andersen
Simon Fraser University

Abstract: How regular do mechanisms need to be, in order to count as mechanisms? This paper addresses two arguments for dropping the requirement of regularity from the definition of a mechanism, one motivated by examples from the sciences and the other motivated by metaphysical considerations regarding causation. I defend a broadened regularity requirement on mechanisms that takes the form of a taxonomy of kinds of regularity that mechanisms may exhibit. This taxonomy allows precise explication of the degree and location of regular operation within a mechanism, and highlights the role that various kinds of regularity play in scientific explanation. I defend this regularity requirement in terms of regularity's role in individuating mechanisms against a background of other causal processes, and by prioritizing mechanisms' ability to serve as a model of scientific explanation, rather than as a metaphysical account of causation. It is because mechanisms are regular, in the expanded sense described here, that they are capable of supporting the kinds of generalizations that figure prominently in scientific explanations.

1. Introduction

In the last 15 years or so, a number of influential accounts of mechanisms have been offered as an alternative to law-based accounts of explanation. An important feature of mechanisms is that the entities and activities that constitute them operate in a regular fashion (Machamer, Darden, and Craver 2000; Darden 2008; Glennan 1996, 2002; Bechtel and Abrahamsen 2005; Craver 2007). Recently, however, some philosophers have weakened or removed the requirement of regularity from the characterization of a mechanism (Bogen 2005; Glennan 2010a, 2010b; Machamer 2004). I will here present a broadened and more fine-grained characterization of regularity with respect to mechanisms, and provide reasons to retain this expanded notion of regularity as a requirement on mechanisms.

There are two distinct arguments for dropping the regularity requirement on mechanisms that I will address. The first concerns the metaphysical nature of mechanisms and causation: eliminating the regularity requirement means that all causation can be construed as mechanistic in nature. The second involves examples

from the sciences that ought to qualify as mechanisms but which do not act always or for the most part, as required by at least one well-cited definition of a mechanism (Machamer, Darden, and Craver 2000). I will respond to both of these points. I argue that the characterization of a mechanism should retain a broadened requirement of regularity of operation, where such regularity is cashed out in terms of statistical predictability and is predicated of specific stages within the mechanism.

I'll show that there is a trade-off between, on the one hand, treating mechanisms as a metaphysical account of causation, and on the other hand, mechanisms' ability to serve as the basis for an account of explanation in the sciences. When we treat mechanisms as an account of explanation, we can and should bracket the metaphysical debate about the nature of causation, since mechanisms are then compatible with a range of metaphysical accounts. The taxonomy of regularity that I offer in section 2 allows for the precise description of a variety of degrees of and roles for regularity within a given mechanism, highlighting how variations in regularity figure in research. I justify retaining regularity of operation in the definition of a mechanism by showing how such regularity is required to individuate mechanisms in a non-arbitrary way, and how regularity figures crucially in the ability of mechanisms to support generalizations and thus to provide better explanations.

There are several distinct definitions of mechanisms in contemporary discussions, but all share certain features that unite them sufficiently for the purposes of this paper. For instance, "Mechanisms are entities and activities organized such that they exhibit the *explanandum phenomenon*" (Craver 2007, 6; italics in original); "A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena" (Bechtel and Abrahamsen 2005, 423). Generally, mechanisms are constituted by a coordinated sequence of causal interactions between component parts organized in such a way that the mechanism's functioning is what produces or gives rise to the phenomenon for which the mechanism is indicated as an explanation. In particular,

almost all characterizations of mechanisms involve regularity of operation as a key condition. It is useful to distinguish between mechanisms, which are actual chains of appropriately causally connected entities in the world, and mechanism models, which are descriptions or schemas of such mechanisms used to explain the phenomena for which the mechanisms are responsible (see Glennan 2005, Illari and Williamson 2010).

I will focus on the widely influential characterization of mechanisms as provided by Machamer, Darden, and Craver (MDC henceforth). "Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up conditions to finish or termination conditions" (2000, 3).[1] Entities are objects in the world, and activities are what entities *do*, by which they produce change. Activities are the causal interactions in which entities engage inside the mechanism; activities connect the entities into a coherent process. The mechanism regularly produces or gives rise to some phenomenon because of the organization of its parts: earlier stages lead reliably to final stages, so that different instances of a mechanism share patterns of activity among similar or identical entities.

Consider a simple example to illustrate this definition. The firing of a neuron can be explained by providing a mechanism involving a chain of electrochemical events that began with the stimulation of a synapse and ended with the release of specific neurochemicals. The organization in this example is spatiotemporal: presynaptic vesicles change spatial location, and must do so before subsequent activities like the release of neurotransmitter can occur. The final termination condition is that of the postsynaptic neuron firing.

Thus, the term "mechanism" may apply to either a type or a token. On one hand, the term can be used to pick out a single individual causal chain in the world. When a particular neuron fires on a given occasion, a mechanism led to that firing.

[1] Some differences between major accounts of mechanisms are not germane to the points I make here (for instance, see Tabery 2004 for an example of a reconciliation between two distinct accounts). It is worth noting that some authors, such as Bechtel, explicitly reject certain features of the MDC definition, namely start-up and termination conditions, in favor of a characterization that can handle cyclical processes as well.

On the other hand, the term is often used to indicate a type of causal chain, one that could recur on multiple instances: when a neuroscience textbook describes the mechanism for neuron firing, it does not describe a single instance, but rather a type of causal chain that presumably occurs on many occasions. In this way, mechanisms can explain both what happens on a single occasion, as well as what happens on all the occasions on which a neuron fires due to this mechanism. The way in which mechanisms comprise both types and tokens thus serves to connect the explanation of a single instance of firing with a generalization about what happens when neurons fire.

A common feature of definitions of mechanisms is that they involve or rely on regularity. In the original MDC (2000) characterization, regularity is cashed out in terms of the mechanism functioning "always or for the most part." Once the start-up conditions occur, a chain of activities between the relevant kinds of entities reliably takes place in the same fashion, every time or almost every time that the start-up conditions occur, leading consistently to the termination conditions. This entails that if an apparent mechanism-type, such as that involving the release of neurotransmitter, fails to bring about the explanandum phenomenon always or for the most part, it is not actually a mechanism. In this definition, regularity of operation is crucial. The presence of the start-up conditions for the mechanism should be enough to ensure that, always or for the most part, the right series of causal activities will occur between the appropriately organized entities such that the phenomenon explained by the mechanism is brought about.

It is this regularity requirement that has come under fire recently (Bogen 2005; Machamer 2004). On a revised view of mechanisms without the regularity requirement, a chain of entities and activities could occur infrequently, perhaps only once, and still be called a mechanism (see especially Glennan 2010b). There are two primary arguments used as justification for dropping regularity altogether: an argument from science, involving examples that ought to count as mechanisms but which fail to work always or for the most part; and an argument from metaphysics, where mechanisms are converted away from an account of scientific explanation

and into an account of the nature of causation itself. I will now articulate and respond to these two arguments.

2. A taxonomy of regularity

The argument from science is straightforward and effective: if we think things like postsynaptic neuron firings can be explained by a mechanism (and there are many reasons we should, including but not limited to entrenched scientific usage of the term), then we must reject the "always or for the most part" element of the MDC definition. Bogen (2005) and Machamer (2004) conclude from the argument from science (among other reasons) that we should reject the regularity requirement altogether. I will agree that the argument from science is compelling enough to warrant a revision in our understanding of what mechanisms are, but argue that we should retain a broadened and more nuanced characterization of regularity as a requirement on mechanisms, in the form of a taxonomy by which to concisely and flexibly express information about regularity within mechanisms.

Bogen's example is that of the mechanism by which action potentials induce the release of neurotransmitters by vesicles in presynaptic neurons. The process of triggering vesicle release of neurotransmitter operates irregularly, in that up to 90% of the time, a given vesicle does not release neurotransmitter under these conditions. Neuroscientists are currently unable to fully account for the difference between the occasions on which the vesicles do release neurotransmitters and those on which they do not (cf. Kandel *et al*, 2000). Bogen rightly argues that we ought to call the causal process that leads to vesicle release of neurotransmitter a mechanism, even though it operates irregularly and infrequently. He wants to resist the "article of faith" that apparently unreliable mechanisms somehow instantiate hidden natural regularity. He argues that we should leave open "the possibility that some causes operate indeterministically and irregularly" in an irreducible fashion – giving up the idea that if we just knew more about the system in question, we would discover regularity in its workings.

Similar concerns have been raised with respect to natural selection as a mechanism. Skipper and Millstein (2005) demonstrate that natural selection does

not act with the kind of regularity required by the MDC definition.[2] And yet, they show there is ample reason to think that natural selection should qualify as a mechanism, including the fact that it is rampantly referred to as a mechanism and described in mechanistic terms by scientists.

Such examples undermine the "always or for the most part" regularity requirement of MDC. In order to label these examples regular mechanisms, a modified characterization of regularity is required. Barros (2008), in response to Skipper and Millstein (2005), distinguished between unbiased stochastic mechanisms, which operate with statistical frequencies at or below 50%, and deterministic mechanisms. While this is a first step in the right direction, Barros' distinction is insufficient to accommodate the rich varieties of regularity that mechanisms can display. What we need is a more precise way to specify both the frequency with which a mechanism operates, and the location within the mechanism where regularity as it is currently conceived breaks down.

We can accommodate these examples of nondeterministic mechanisms without giving up regularity, and by doing so, I'll show, we actually gain a more nuanced understanding of the role regularity plays in scientific investigation and explanations. What follows is a taxonomy of regularity, something like a map of the territory across which mechanisms can vary with respect to regularity. My claim is that any chain of entities and activities satisfying these minimal requirements should count as exhibiting sufficient regularity to qualify as a mechanism. It is certainly possible for a causal chain of entities and activities to fail to exhibit any kind of regularity in this taxonomy: in that case, it should not count as a mechanism.[3]

---

[2] Their point is not solely about the regularity requirement, however: natural selection may also violate the requirements for organization and productive continuity. While it is not the focus on this paper, I will note that the productivity requirement only poses a problem for accommodating natural selection as a mechanism if one relies on a productive, rather than counterfactual, account of causation.

[3] This implies that there can be causal chains that are not mechanisms, and causal explanations that do not cite mechanisms. By requiring mechanisms to be at least minimally regular, I thereby retain a distinction between the class of causal chains in the world, and the proper subset of that class containing causal chains that are also mechanisms. I will take this issue up again in the section on the argument from metaphysics.

There are three main parameters along which regularity in a candidate causal system can vary: the *organizational location* of regularity in terms of stages within a mechanism; the *strength of connection* between component elements in a mechanism; and the *failure pattern* by which a mechanism could fail to operate always or for the most part while still exhibiting sufficient regularity to count as a mechanism. Specifying each of these three parameters for a given mechanism allows us to conveniently express a great deal of information about that mechanism, how it may connect to other mechanisms, and where the fruitful avenues of continued research with respect to it are likely to be found. I'll first outline the taxonomy, and then demonstrate how it works with an example.

First, consider the ways in which distinct stages within a mechanism could be empirically regular in terms of their *organizational location* in the overall structure of that mechanism:

1) Regular occurrence of start-up conditions
2) Regular triggering of the mechanism once start-up conditions occur
3) Regular operation of specific activities connecting the entities within a single mechanism
4) Regular production of termination conditions once mechanism has been triggered

This parameter localizes within a mechanism where, precisely, one is attributing a specific level of regularity. Mechanisms are often organized spatiotemporally, such that 1-4 would pick out events within a temporal or spatial sequence. However, to accommodate mechanisms with organizational structures that are not clearly spatiotemporal (such as psychological mechanisms, or social mechanisms), organization location is the most precise label.

We may find mechanisms where the start-up conditions occur very rarely, far too rarely to be called regular, but where the occurrence of those conditions is nevertheless sufficient to bring about the termination conditions for that mechanism (1). Bogen's (2005) example of conditions for human evolution fits under this heading: the failure of regularity stems from the fact that the conditions under which the mechanism could operate occur with such exceeding infrequency.

Bogen's example of vesicle release when an action potential is present illustrates (2): start-up conditions may occur much more frequently than does triggering of the mechanism. Any stage after triggering but before termination conditions qualifies under (3). The last, (4), is included in addition to (3), since we can differentiate between the regularity with which a single given stage within a mechanism operates, and the regularity with which the mechanism overall operates, i.e. how often it achieves or fails to reach termination conditions. A mechanism may contain two or more stages that do not act always or for the most part, or which do not fail independently of one another, such that the cumulative effect needs to be considered separately from the regularity of any given stage.

Mechanisms may display different kinds of regularity at different organization locations at the same time. In this regard, this parameter is not intended to convey all the information about regularity for a given mechanism. It is intended to convey precise information about a particular locus within the mechanism. For this reason, we can provide this taxonomic information about regularity for multiple different organizational locations within a single mechanism. This can ground potentially informative comparisons: one mechanism for speciation may have start-up conditions that occur extremely infrequently (1); it may have start-up conditions that obtain frequently but often fail to trigger the mechanism even when they do occur (2); that same mechanism may involve component stages that often fail to occur even once the mechanism is triggered (3). (4) accommodates cases where the failure of component elements within the mechanism may interact, where a failure at one stage triggers a back-up stage, or where two failures within a mechanism cancel one another out. The taxonomy allows us to separate information about (1), (2), (3), and (4) for a single mechanism.

This brings us to the second parameter. At each of these organizational locations within the mechanism, the term "regular" could also indicate a different *strength of connection* between stages:

a) Deterministic: cause cannot fail to act once conditions are appropriate

b) Reliable but not exceptionless: most of the time, the cause brings about the effect, but there are occasions on which it does not, and we may or may not be able to provide an explanation for the exception
c) Sporadic: the cause fails to act often enough that it cannot be considered merely an exception when this failure occurs
d) Infrequent: most of the time, the cause fails to bring about its effect, but once in a while it does

The first two correspond to MDC's "always or for the most part" regularity requirement. If one were to group (a) and (b) together, and (c) and (d) together, and consider only this parameter, one would approximate Barros' (2008) distinction. We can accommodate (c) and (d) as additional ways to instantiate regularity, when modified by the next parameter regarding failure patterns. There need not be a hard and fast line between different connection strengths; many mechanisms will fall straightforwardly into the category of sporadic or infrequent based on context-specific information about failure rates. For instance, a 70% success rate might be reliable in the context of vesicle release of neurotransmitter, while a 70% success rate for a car's braking mechanism would certainly not count as reliable.

One might be concerned that this way of categorizing the strength of connection ignores the fact that some apparently weak connections in mechanisms are cases of genuine indeterminacy, while some simply reflect our ignorance of further underlying causes. In other words, lower strengths of connection, such as (c) or (d), might be attributable to epistemic limitations, such that further information would reveal that the connection is actually much stronger. Does the failure of this taxonomy to distinguish ontological from epistemological indeterminacy turn mechanisms into merely epistemic artifacts, rather than ontological features of the world?

This is the very "article of faith" regarding regularity that Bogen claims we should resist, and I agree with him. There are several reasons why we should not require that mechanisms distinguish between the two. Our inability to distinguish ignorance-based and ontological sources of indeterminacy does not threaten the ontological status of mechanisms. We are referring to mechanisms that are in the

world, even if we misattribute certain features to them, such as a higher level of indeterminacy than they may have. Further, it is unreasonable to hold explanations to such a high standard that we only have them in cases where we are sure that we know everything there is to be known about the mechanism. The example of vesicle release of neurotransmitter illustrates this effectively. Neuroscientists are unsure as to what accounts for the particular frequency with which such release occurs. Chances are that further research will reveal underlying factors that, when they are taken into account, increase the strength of connection. However, there still may be some residual indeterminism, even if additional factors are found that account for part of the indeterminacy. We just don't know how much of the weak connection is intrinsic to the mechanism and how much is due to our ignorance, and so attempting the separate them is an impossible task.

This leads to the third parameter. There are at least two *failure patterns* by which a mechanism could fail at a given organizational location, but nevertheless exhibit some kind of regularity. If at least one of these failure patterns is discernible at an organizational location with low connection strength, I argue the mechanism is still regular:

i) Known statistical distribution of indeterminacy: the mechanism succeeds some consistent percentage of times, even though we may not be able to account for why it succeeds or fails when it does
ii) Known interfering factors: when the mechanism fails at a given organization location, we can identify factors that interfered on this occasion, whether or not we know the precise quantitative impact of such factors on the mechanism function in general

While for practical reasons it will often turn out that only one of these two can be attributed at a given time, these are not and need not be mutually exclusive categories. There will either be interfering factors that are known but which cannot be given a precise probability of occurring; or there will be probabilities that remain fairly constant, in spite of the fact that we do not know what gives rise to them. Both

of these failure patterns could be simultaneously predicated of a single organizational location with a particular strength of connection.

The first failure pattern will be examined in more detail in the next section. The second failure pattern can be illustrated by thinking about mechanisms for which we have evidence while still lacking details about the frequency with which the mechanism operates because of the widespread presence of interfering factors. The generation of exotic particles in high-speed colliders is such a case. Physicists have good reason to think that they have enough understanding of the mechanisms by which Higgs bosons are generated (if they are) that they can replicate this process, in spite of the fact that it has not yet successfully been done (at the time of writing). The interfering factors that prevent operation of the mechanism leading to production of the Higgs boson are numerous and impossible to enumerate ahead of time – witness recent trouble with the Large Hadron Collider, which included technical failures but also at least one researcher accused of terrorism. These potential interfering factors cannot be formalized into a meaningful unitary failure rate in the way that the neurotransmitter release by vesicles can, but can still be accommodated within a broad view of regularity. The problem is not with the mechanism *per se*. The interfering factors can easily be recognized as such when they occur even if we cannot assign a specific probability to their occurrence before they occur.

Another example of (ii) is that of ecosystem succession. Ecologists may know the start-up conditions for ecological succession in a given ecosystem, involving disturbances to habitat or creation of new habitat. Yet not all start-up conditions lead to the termination conditions of climax communities, because there are a range of external factors governed by chance that could prevent the mechanism from working. This may include adverse weather events, human interference, geological events, etc. We may be completely unable to estimate the likelihood that one of these events will occur, such as the likelihood that a condo developer will decide to purchase and then build on some piece of habitat. Yet there is still a sense in which the mechanism operates regularly, even if more often than not the start-up conditions fail to lead to the termination conditions, and ecologists are unable to

provide a precise statistical likelihood that a given succession will occur because of such interfering factors. Interestingly, interference by known factors has been incorporated into the mechanistic explanation of succession and failure of succession (see, e.g., Sheley *et al* 2004). This leads to hybrid mechanisms used for range management purposes: a given mechanism for succession (the distribution and diversity of plant types in rangeland) can be intervened on at a certain stage by essentially grafting on a different mechanism (grazing by sheep versus cows) to bring about a desired end state (one distribution of plant types rather than another).

Mechanisms can fail to operate always or for the most part in ways that draw from each of these lists (again, there are conditions under which (i) and (ii) can co-occur). One can identify a particular organizational location in a mechanism, then identify the strength of connection between mechanism components at that location, which may then yield a known statistical pattern of success and failure, as well as known interfering factors that might lack reliable statistical distributions. Taken together, there are a lot of different ways to instantiate regularity. Bogen has pointed out a number of instances of what we should call mechanisms that fail to occur always or for the most part, but these examples do display at least a version of regularity categorized here. My claim is that so long as, and only so long as, a mechanism displays at least some minimal form of regular operation, it ought to be counted as a mechanism. Causal chains of entities and activities that do not display any of these forms of empirical regularity should not be counted as mechanisms.

3. Statistical regularity in mechanisms: an example

I'll illustrate use of this taxonomy of regularity with an example in order to show how using it can illuminate explanatory connections between different mechanisms and efficiently communicate detailed information about specific mechanisms. Consider Bogen's example of neurotransmitter release by presynaptic vesicles, which subsequently triggers depolarization in the postsynaptic neuron. There are two mechanisms here, one embedded in the other. The first is the mechanism by which an action potential triggers any given vesicle to release a quantum of neurotransmitter. The second is the broader mechanism by which a presynaptic

action potential triggers postsynaptic firing, one stage of which is the release of neurotransmitter by vesicles.

First, consider the inner mechanism. An action potential induces the vesicles to move towards and fuse with the cell surface at the active zone of the presynaptic neuron (Kandel *et al* 2000, 262). There is a relatively high failure rate – as many as 90% of vesicles fail to release their quanta of neurotransmitter when the triggering conditions are present. One can calculate the probability of finding a given number of quanta released, using at least two parameters. However, even though the overall probability can be calculated, scientists are unsure of the factors that contribute to this probability. Calcium ions seem to play some role, but it is not clear exactly how or where they do so.

> The parameters *n* and *p* are statistical terms; the physical processes represented by them are not yet known… The parameter *p* probably represents a compound probability depending on at least two processes: the probability that a vesicle has been loaded or docked onto a release site (a process referred to as vesicle mobilization) and the probability that an action potential will discharge a quantum of transmitter from a docked active zone. (Kandel *et al* 2000, 261)

For vesicle release of neurotransmitter, then, the *organizational location* is the triggering of the mechanism given start-up conditions (2); the *strength of connection* is sporadic – it is not an exception or unusual when a vesicle fails to dock and release transmitter (c). And the *failure pattern* is that of a known statistical distribution without known factors (i). Scientists are not entirely sure what factors block the mechanism running from action potential leading to release of neurotransmitter, but they have reliable means to calculate the probability of failure for a given vesicle, or a given neuron with many vesicles. In sum, for this mechanism, the taxonomy provides a label of (2-c-i).

Now consider the embedding mechanism, that of a presynaptic action potential triggering a postsynaptic depolarization. The previous mechanism now figures as a stage within this mechanism. In terms of the taxonomy, the *organizational location* in question is the release of neurotransmitters in the

presynaptic cell given the triggering of the start-up conditions, which is now an intermediary stage leading from presynaptic cell firing to postsynaptic cell firing (3). The *strength of connection* is now sufficient to count as reliable (b). This parameter is based on the vesicles' likelihood of releasing sufficient neurotransmitter to trigger postsynaptic depolarization. Yet it is stronger than the strength of connection in the vesicle mechanism just discussed. This is because even though there is a low chance that a given quantum of neurotransmitter will release, there are multiple quanta that must all fail for the mechanism to fail (Kandel *et al* 2000, 261). Finally, this is again a case where there is a *known statistical distribution* of indeterminacy and unknown interfering factors (i). We know the likelihood that a given postsynaptic neuron will respond to stimulus by neurotransmitters, and the thresholds that must be met for sufficient neurotransmitter release. However, since we are localizing our label of regularity to this particular locus within the mechanism, we end up with the same result for this parameter as with the more narrow mechanism, since that mechanism is the stage of this mechanism we are describing.[4] The strength of connection has gone up, since there are many vesicles that would need to fail together for this mechanism to fail. We can also add information about independence: using a Poisson distribution to represent the probability of a given number of vesicles releasing neurotransmitter means that the release by each vesicle is sufficiently independent from other vesicesl as to approximate such a distribution. When failure does occur, though, it is still the same process about which we lack knowledge of the physical processes that generate the probability (see Kandel quote above) We can refer to this example of regularity as type (3-b-i).

In this example, there is an instance-by-instance indeterminacy: for any single action potential on a given occasion, we can't say whether any given vesicle will release neurotransmitter, and thus whether sufficient vesicles will release so as to trigger postsynaptic depolarization. However, in spite of this, there is a meta-regularity concerning the single-case indeterminacy that justifies calling the whole

[4] While this was true when written, there are recent developments that may bear directly on this issue of identifying the factor(s) that control or influence the probability of neurotransmitter release; see (Lee *et al*, 2010).

process a regularly occurring mechanism. In any single case, we do not know if a vesicle will dock and release, or how many vesicles will do so in the larger mechanism, and thus, we do not know for a single case whether postsynaptic firing will occur. But, when we consider hundreds of thousands of neurons, we can make firm claims with a fairly high degree of precision about how many in that population will fire under such conditions. If each quantum of neurotransmitter has a .9 chance of failing to be released, we have sufficient knowledge to forecast how many neurons in a given population will fire on release of sufficient quanta of neurotransmitter (Kandel *et al* 2000, 261). We also know approximately how often neurotransmitter will release given repeated triggering of a single neuron.

This leads to an extremely interesting feature of this broadened characterization of regularity. There are good reasons to think that the sporadic operation given start-up conditions (2-c-i) of the smaller mechanism, and thus the specific regularity associated with the broader mechanism (3-b-i) together play an important role in the functioning of the brain. By only responding to some fraction of potential triggers, pre- and postsynaptic cells effectively dampen many signal propagations, thereby limiting the amount of noise in the overall system while still ensuring that a sufficient number of cells overall respond to propagate signals. The pattern by which neurons fire thus appears to play a key role in encoding information, even though the overall probability of firing remains constant.[5] Specific types of regularity naturally display the interrelationships between closely connected phenomena. A mechanism will display one particular form of regularity, rather than another, because of the role that the mechanism itself plays in some overlapping, higher-level, or lower-level mechanism. The very infrequency of operation of this mechanism given start-up conditions can be treated as another regularity, in that it calls for both explication and explanation.

We can thus explicate the existence and size of the indeterminacies of (2-c-i) and (3-b-i) in terms of the role they play in a different mechanism, that of regulating signal propagation. The statistical distribution of depolarization following

[5] See, for instance, (Mainen and Sejnowski 1995; Victor and Purpura 1996; and London *et al* 2010).

neurotransmitter release needs to fall within a rather narrow range in order to effectively serve as an intermediate stage in a larger-scale mechanism. This illustrates the way the taxonomy of regularity can represent the connections between related mechanisms. The mechanism(s) for maintaining this precise and presumably most effective rate of firing also call out for explanation. As Bogen indicates, neuroscientists do not know what factors account for why neurotransmitter release occurs or fails as it does, or if this is an irreducibly statistical phenomenon. The original regularity associated with the mechanism for triggering vesicle docking and neurotransmitter release traces out avenues of research for neuroscientists. This taxonomy thus captures the relevant regularity information of connected or embedded mechanisms as it serves the epistemic virtue of fruitfulness in science.

Thus, I agree with the argument from science to the effect that there is a strong motivation for accommodating cases like this under the heading of mechanisms. I differ from Bogen and Machamer by retaining an expanded notion of regularity. One advantage of maintaining regularity in the expanded form illustrated in this section is that it emphasizes how the failure pattern at an indeterministic functional location in one mechanism may play a role in another mechanism. Such statistical 'irregularities' thus also serve an epistemic function: statistical regularities within mechanisms that are not "always or for the most part" are phenomena that also need to be explained. They will generally occur at the rate that they do for some reason, either because of some constraints on occurrence, or because they serve a functional role in another mechanism by having the rate that they do, and so on. Patterns in failure regularities thus help outline the connections between mechanisms.

4. Using regularity: mechanism individuation and explanation

Why we should respond to the argument from science by retaining a broadened form of regularity, rather than simply rejecting regularity altogether? There are two primary reasons: regularity is required to non-arbitrarily individuate mechanisms from surrounding causal processes that do not contribute to the phenomenon in

question; and mechanisms would be unsuited to serve much of their explanatory role in the sciences without the regularity requirement. I'll discuss each of these in turn.

There are several regards in which regularity of operation is crucial for individuating the boundaries of a given mechanism from its environment or context. Individuation of a mechanism involves, among other things, differentiating the entities and activities that constitute the mechanism from those that may occur in close spatiotemporal proximity to the mechanism but without contributing to it. Regularity does just that: it provides the grounds, both epistemic and ontological, to identify causally relevant entities and activities in a mechanism against a background of spatiotemporally proximate causal interactions that do not contribute to the mechanism.

The background against which we distinguish mechanisms is replete with entities and activities that are rampantly causally interacting. When any given phenomenon of interest is produced, there is a rich nexus of such causal activity going on, most of which was not involved in bringing about the phenomenon in question.[6] To offer a mechanism as an explanation for why a certain phenomenon occur in the way and at the time that it does, a key task is essentially winnowing out, from all of the many entities engaging in activities that are candidates for causally contributing to the phenomenon, those that actually did so. Were we to provide an 'explanation' by citing all of the causal interactions that took place in the spatiotemporal vicinity in question, we would be swamped by information, most of which would be irrelevant. Some elements of the rich causal nexus contributed; some did not. Individuating the mechanism in question involves figuring out what did causally contribute. The issue of mechanism individuation is thus closely connected to the issue of causal relevance.

When a mechanism operates with at least some kind of regularity, we then have grounds by which to draw the mechanism boundary around the entities and

[6] This point is made clearly in Salmon's work (e.g. 1977) on causal processes and interactions. MDC explicitly offer an account of mechanisms that is compatible with Salmon's account of causation.

activities that were actually involved in bringing about the phenomenon in question. Multiple instances of some phenomenon allow us to judge the entities and activities that were genuinely causally involved in the mechanism that produced the phenomenon.[7] Comparing different occurrences thus provides grounds to identify the stages that genuinely contribute to the mechanism.[8] The more often a mechanism occurs, and the more regularly it occurs, the clearer it becomes which entities and activities are part of the mechanism (i.e. what the mechanism is). It also becomes clearer that there is a mechanism there at all, something that either could explain a phenomenon or which could require further explanation.

It is crucial for individuation in this way that mechanisms have at least *some* minimal form of regularity. One-offs, causal chains of entities and activities that by definition occur once, cannot be individuated in this way. In those cases, there is no comparison class by which to gauge which entities and activities were causally relevant to producing the phenomenon in question, versus the causal interactions which occurred in the vicinity but did not contribute to producing the phenomenon in question. In particular, if one takes mechanisms to have something like start-up conditions, one-off causal chains will have no non-arbitrary cut-off point where the mechanism starts – there will only be a long chain of activities and entities.

Consider an example of individuation of the start-up conditions for a mechanism, the mechanism explaining a postsynaptic cell firing. There are ongoing activities performed by a variety of entities within the presynaptic neuron, including, for instance, some that involve the vesicles that subsequently release neurotransmitter into the synaptic cleft. Of all the activities that have some causal impact on the vesicles, however, most will not lead to neurotransmitter release (using the rich causal nexus idea of what counts as a causal interaction with a vesicle, there will be many causal interactions in which the vesicle is involved but

[7] One might be tempted to say that constant conjunction alone is insufficient to justify causal claims. I am certainly not suggesting that scientists somehow either do or should rely on constant conjunction alone to find causal relations - there is no reason to pretend the rest of scientific practice doesn't exist while discussing the role of regularity in it.
[8] It's worth clarifying that the first step in individuating a mechanism is often the identification of a phenomenon of interest for which a mechanism is sought as an explanation.

which do not trigger the release). When an action potential arrives at the nerve terminal, it opens calcium ion channels, which causes the vesicles to fuse with the presynaptic cell membrane. This part of the process is now part of the mechanism – the way in which the opening of calcium ion channels causes vesicles to move towards and fuse with the membrane is a crucial stage of entities engaging in activities required for operation of this mechanism. The start-up conditions involve specific kinds of causal interactions for the vesicle, including the opening of the calcium ion channels. Start-up conditions do not involve the other causal interactions in which the vesicle engaged prior to, or even during, this stage.

The justification for drawing the line, spatiotemporally, at this particular influence on vesicles is not rooted in any single instance of a neuron firing. If only one such event ever occurred, we would lack the means to determine which of the many activities engaged in by the relevant entities were the ones that constituted the start-up condition for the mechanism in question. There are two closely related points here. Epistemically, we justify our choice of some causal interaction(s) as the start-up conditions for this mechanism because we know, based on multiple such instances, that these are the conditions that are both necessary for the mechanism to be triggered and which lead, in some kind of regular fashion, to the termination conditions. Ontologically, however, it is already true that in this single instance, some causal interactions but not others led to the termination conditions. Thus, the epistemological access that regularity of operation can provide is to the genuine causal structure of the individual instance. Regularity thus provides epistemic access to the causal structure of mechanisms. We should not, however, assume that *only* regularity does so, or that regularity alone is sufficient to ascertain the causal structure of mechanisms. Regularity, in the context of, for instance, Bayes' nets methods, may only provide us with a class of possible causal structures. But regularity is not the only epistemic tool we have; we also have interventions, for instance, which supplement regularity information for epistemic purposes.

But my main point here is ontological in character. Even if we knew the entire causal structure at hand – even if we had constructed the mechanism ourselves – we would still require a non-arbitrary way to identify the boundary

between mechanism and context.[9] Without at least some regularity, the line between mechanism and environment either disappears or becomes gerrymandered. For a one-off causal chain, we need a reason to call a particular set of entities or activities the start-up condition for the mechanism, rather than the stage before that, or the stage after it. We need a reason to highlight a particular set of activities and entities with finite boundaries, and regularity is a primary way to do so.

The issue with individuation is not merely that of ascertaining which entities and activities were causally involved, but also to ascertain the range of variability in those entities and the ways in which they perform activities across which the mechanism is still robust enough to produce the phenomenon. Even if we somehow knew which entities and activities were involved in a singular causal chain, we would lack the information about range of variability: how robust to perturbation are the entities and activities in this one-off 'mechanism'? How differently could the same entities have interacted, or which entity substitutions could have occurred, without substantial change to the phenomenon thereby produced? Some changes will be sufficient to destroy the mechanism, but many will not. In order to garner this information, we need at least some kind of regularity.

This segues into my second point, namely, the key role that regularity plays with respect to mechanisms as an account of scientific explanation. One of the consequences of dropping the regularity requirement from the definition of a mechanism is that of collapsing the distinction between mechanisms and causation in general. Any chain of causes could be called a mechanism; we could add another causal interaction to a given chain, or take some away, and it would still be a mechanism. Moreover, calling it a mechanism would not give us any predictive or control capacities with respect to *other* causal chains – such a mechanism would

[9] An interesting discussion on the distinction between mechanism and environment can be found in Delehanty (2005). She argues that, for the case of token-token reduction of higher-level to lower-level descriptions, a lower level mechanism that accounts for the type of phenomena to be explained can be extended in the token instance at hand to include what is ordinarily part of the context as an additional part of the mechanism. I note that there must already be previously delineated boundaries to the type-mechanism in order to apply this strategy of extension.

only be capable of explaining a single occurrence. Explanations of unique historical events display these features, but very few scientific explanations do.

Retaining regularity in the definition of a mechanism involves retaining the ability to differentiate a stronger and *more explanatorily useful* subset of causal explanations from other kinds of causal explanations (see also Darden 2008). The regularly recurring chains of entities and activities that constitute a mechanism make them more useful in explanations other chains of causal-activity-connected entities because regularity grounds generalization from one instance to further instances. The more regular a mechanism is, the more potential instances to which the same explanation applies.

Generalizations based on mechanisms can vary along numerous parameters: in terms of the degree of precision with which they are formulated, the scope of phenomena to which they apply, their accuracy within that range, and their stability under perturbations, to name a few.[10] Mere causal explanations, those that do not involve mechanisms because the phenomena that they explain lack any variety of regularity, do not vary along these parameters because they only apply once. There is no range or scope of applicability, no range of perturbations to the system under which the explanation still holds, not even multiple occurrences to which it can be applied. There is only a single event being explained. There are occasions, both within science and especially outside of it, when particular explanations of token events are what we need. But these are not the rule, and they are insufficient motivation to assimilate mechanisms to mere one-offs. Mechanistic explanatory practices in the sciences hinge on considering single instances as instances of a type, and on providing explanations based on mechanisms that constitute a type of causal chain, not merely a single instance of one. What happens when a single neuron fires once is not the target of investigation or of explanation. What happens when neurons fire, and why and how it happens, in general: this is what mechanisms can explain, but only if they retain regularity.

[10] For an illuminating and more complete look at the relationship between generalizations and causation in the sciences, see Mitchell (2008).

In sum, what the argument from science shows is that we need a more fine-grained understanding of the sorts of regularities that causal mechanisms can exhibit at various organizational locations. There is no need to eliminate regularity based on these examples. Even accounts of mechanisms that include regularity will benefit from the clarity and conciseness that this taxonomy can provide.

5. Metaphysics and mechanisms

I will now consider the argument from metaphysics for dropping the regularity requirement on mechanisms. The taxonomy of regularity, and the discussion of role of regularity in individuation and explanation, are primarily responses to the argument from science. I will argue for a deflationary response to the argument from metaphysics. No single notion of mechanisms will serve both explanatory and metaphysical needs.

There are two versions of the argument from metaphysics against regularity that can be found in contemporary literature. The first is the line taken by Bogen (2005, 2004) and Machamer (2004), concerning the metaphysical status of counterfactuals. Bogen (2005, 2004) emphasizes that while counterfactuals, such as those found in Woodward (2003), may serve as the means for keeping track of what causal claims are true or false, they cannot provide the grounds on which they are true or false. Causation in Woodward's account involves "systematic patterns of counterfactual dependence" (2003, 191); Bogen interprets this as a kind of regularity account of causation, where the relevant regularity includes both actual and counterfactual occurrences.

Rather than take a stand on this point, I want to offer a skeptical view of the importance of counterfactuals versus production for understanding mechanisms themselves. The notion of regularity that the above taxonomy maps out is actual and not counterfactual, namely, multiple occurrences in the actual world. As such, the notion of regularity defended in the earlier section is not undermined by a metaphysical argument against counterfactuals: it is agnostic with respect to this debate. One can therefore require that mechanisms display at least some form of regularity without thereby requiring that they display counterfactual regularity. The

causation within mechanisms could be construed in counterfactual terms, or it could be construed in productive terms. Since the expanded regularity requirement does not hinge on counterfactuals, this version of the argument from metaphysics gets no traction on it.

The second version of the argument from metaphysics takes up just this point: namely, what is the metaphysical nature of causation? One might think it is counterfactual, or productive; Glennan (see especially 2010a, 2010b) thinks it is mechanistic. The metaphysical argument against regularity is this: if we eliminate the regularity requirement, then we can use mechanisms as an account of all causation, not just for those instances of causation privileged by some kind of regularity. His account has a number of metaphysical advantages, such as its ability to undermine the apparent epiphenomenality of higher level causes that results from the Causal Exclusion argument.

However, there is a trade-off in terms of explanatory power when treating mechanisms as a general account of causation. The trade-off is between metaphysical viability as an account of causation versus usefulness or adequacy as an account of scientific explanation: using mechanisms for metaphysics precludes also using the same notion of mechanisms for explanation, and vice versa. Furthermore, the parameter that varies between mechanisms-as-metaphysics and mechanisms-as-scientific-explanation is precisely that of regularity. If we eliminate the regularity requirement in order to use mechanisms as an account of causation, we thereby lose what made mechanisms so useful as an account of scientific explanation. If we keep the regularity requirement so as to use mechanisms as an account of explanation in the sciences, we lose the ability to use it to cover all instances of causation. We cannot use a single account to both do the explanatory work with respect to the sciences that the new mechanism approach is supposed to do, while also working as an account of causation.[11]

[11] One could advocate a mechanistic account of causation while also maintaining that mechanisms, as defined by authors like Machamer, Darden, Craver, Bechtel, Glennan, etc., offer an account of explanation. But to do so, two different accounts are needed: these require two distinct notions of mechanisms, which is why a single account of mechanisms cannot accomplish both tasks. Both the new mechanism approach and the mechanism-as-causation approach have

Consider this tension in more detail. As we saw in sections 2 through 4, a broadened notion of regularity serves a number of important purposes. It provides non-arbitrary means to individuate the boundaries of mechanisms. It provides the grounds for generalizations, making mechanisms genuinely useful for explanation in the way that one-off causal chains are not. It is capable of conveying an enormous amount of information about a mechanism and how it functions, how it fits into its environmental context, how it connects to other mechanisms, where we need to investigate further, and more. If we give up the regularity requirement, we give up the ability to use mechanisms as an account of *scientific* explanation as an enterprise with distinctive features when compared with, for instance, legal, social, or historical explanation, any of which may rely to a large extent on singular causal instances for explanation.

Why not just keep regularity, then, and still use mechanisms as a metaphysical account of causation? The second half of the trade-off is that keeping regularity as a requirement on mechanisms involves a commitment to a distinction between mechanisms and causation in general. Mechanisms (with the regularity requirement) are a proper subset of all causal chains: mechanisms do not exhaust causation if we retain regularity. This means that there exist causal chains that only occur once and thus are not mechanisms. It would be a poor metaphysical account of causation that labeled one-off causal chains either nonexistent or inexplicable.

There is a sound historical tradition to Glennan's use of mechanisms to physically connect causes and effects, and I do not hereby simply dismiss that view. But I do want to emphasize that it comes at a cost. "Old school" mechanisms serve a solid metaphysical purpose, but the main impetus of the "new school" mechanisms was to capture the explanatory practices of sciences such as biology, which do not rely on covering laws to explain phenomena. Insofar as the primary focus of the new mechanism accounts is scientific explanation, the regularity requirement should be retained, contra the argument from metaphysics.

---

historical claim to the term, but in my view, they need to be terminologically distinguished to indicate that it is not the same notion of mechanism underwriting both.

6. Conclusion

Bogen's argument from science identifies scientific examples that ought to count as mechanisms but which fail to meet a more narrow regularity requirement that mechanisms must act always or for the most part. I have argued for retaining and broadening the regularity requirement on mechanisms to include the varieties of regularity detailed in the taxonomy introduced in section 2. Regularity on my taxonomy is something that is predicated of particular organization locations within a mechanism, with varying degrees of strength, and as potentially instantiating several different informative failure patterns. Retaining this broadened regularity requirement has several advantages. It provides nonarbitrary means of distinguishing mechanisms from other chains of causes. It also provides the basis for generalizations from one instance to another, which is crucial for the role mechanisms play in scientific explanation. The details of how regularity is instantiated, per the taxonomy divisions (organizational location, failure pattern, etc.), carry important information about the scope of generalizations that can be made based on a single instance of mechanism operation.

In response to the use of mechanisms as a metaphysical account of causation, I have shown how this sacrifices mechanisms' usefulness in scientific explanation. If we use mechanisms for a metaphysical account of causation, we must give up regularity and thereby explanatory capabilities. If we retain regularity and explanatory capabilities, we cannot use mechanisms as an account of causation. I have thus argued that we should retain a distinction between causation in general, and the stronger version of recurrent causal chains that constitute mechanisms. This is why we should continue to require regularity as part of the definition of a mechanism.

References

Barros, D. Benjamin (2008). "Natural Selection as a Mechanism." *Philosophy of Science* 75(3): 306-322.

Bechtel, William and Adele Abrahamsen (2005). "Explanation: A Mechanistic Alternative."*Studies in History and Philosophy of Biological and Biomedical Science* 36: 421-441.

Bogen, Jim (2004). "Analysing Causality: The Opposite of Counterfactual is Factual." *International Studies in the Philosophy of Science* 18(1): 3-26.

Bogen, Jim (2005). "Regularities and Causality; Generalizations and Causal Explanations", *Studies in History and Philosophy of Biological and Biomedical Science* 36: 397-420.

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

Darden, Lindley (2008). "Thinking Again about Biological Mechanisms." *Philosophy of Science* 75(5): 958-969.

Delehanty, Megan (2005). "Emergent properties and the context objection to reduction." *Biology and Philosophy* 20:715–734.

Glennan, Stuart S. (1996). "Mechanisms and The Nature of Causation." *Erkenntnis* 44: 49-71.

Glennan, Stuart S. (2002). "Rethinking Mechanistic Explanation." *Philosophy of Science* 69(S): S342-S353.

Glennan, Stuart S. (2005). "Modeling Mechanisms." *Studies in the History and Philosophy of Science Part C* 36(2): 443-464.

Glennan, Stuart S. (2010a). "Mechanisms, Causes, and the Layered Model of the World." *Philosophy and Phenomenological Research* 81(2) 362-381.

Glennan, Stuart S. (2010b). "Ephemeral Mechanisms and Historical Explanation." *Erkenntnis* 72(2): 251-266.

Illari, Phyllis McKay, and Jon Williamson (2010). "Mechanisms are real and local." in *Causality in the Sciences*, eds. Illari, Russo, and Williamson. Oxford University Press.

Kandel, Eric; Schwartz, James; and Jessel, Thomas (2000). *Principles of Neural*

Science: Fourth Edition. McGraw-Hill Companies.

Lee, Han-Ki; Yang, Yoosoo; Su, Zenglui; Hyeon, Changbong; Lee, Tae-Sun; Lee, Hong-Won; Kweon, Dae-Hyuk; Shin, Yeon-Kyun; Yoon, Tae-Young (2010). "Dynamic Ca2+ Dependent Stimulation of Vesicle Fusion by Membrane-Anchored Synaptotagmin 1." *Science* 328: 760-763.

London, Michael; Roth, Arnd; Beeren, Lisa; Hausser, Michael; and Latham, Peter (2010). "Sensitivity to perturbations *in vivo* implies high noise and suggests rate coding in cortex." *Nature* 466: 123-127.

Machamer, Peter; Darden, Lindley; and Craver, Carl F. (2000). "Thinking About Mechanisms." *Philosophy of Science* 67(1): 1-25.

Machamer, Peter (2004). "Activities and Causation: The Metaphysics and Epistemology of Mechanisms." *International Studies in the Philosophy of Science* 18(1): 27–39.

Mainen, ZF, and Sejnowski, TJ. (1995). "Reliability of spike timing in neocortical neurons." *Science* 268(5216): 1503-1506.

Mitchell, Sandra D. (2008). "Exporting Causal Knowledge in Evolutionary and Developmental Biology." *Philosophy of Science* 75(5): 697-706.

Salmon, Wes (1977). "An At-At Theory of Causal Influence." *Philosophy of Science* 44(2): 215-224.

Sheley, Roger; Jacob, James, and Svejcar, Tony (2005). "Integrating disturbance and colonization during rehabilitation of invasive weed-dominated grasslands" *Weed Science*. 53(3): 307-314.

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

Tabery, James (2004). "Synthesizing Activities and Interactions in the Concept of a Mechanism." *Philosophy of Science* 71(1): 1-15.

Victor, JD, and Purpura, KP (1996). "Nature and precision of temporal coding in visual cortex: a metric-space analysis." *Journal of neurophysiology* 76(2): 1310-1326.

Woodward, James (2003). *Making Things Happen*. New York: Oxford University Press.