Beate Krickel

The Mechanical World

The Metaphysical Commitments of the New Mechanistic Approach

Springer

Beate Krickel
Department of Philosophy II
Ruhr-University Bochum
Bochum, Germany

ISSN 1573-4536 ISSN 2468-399X (electronic)
Studies in Brain and Mind
ISBN 978-3-030-03628-7 ISBN 978-3-030-03629-4 (eBook)
https://doi.org/10.1007/978-3-030-03629-4

Library of Congress Control Number: 2018961738

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Chapter 6
Mechanistic Phenomena

The notion of a phenomenon plays a crucial role in the new mechanistic thinking. First, phenomena are the things that are *explained by mechanisms*, i.e., they are the things referred to in the explanandum of a mechanistic explanation.[1] Second, phenomena are what mechanisms 'are responsible for,' i.e., they are the causal or constitutive products of mechanisms. Third, identifying phenomena is crucial for *individuating mechanisms*. As I have explained in Chap. 3, mechanism types are types of causal sequences that are regular or reversely regular relative to a particular phenomenon type. Fourth, the individuation of the phenomenon determines what is a *component* of the respective mechanism. As I have explained in Chap. 5, mechanisms consist of those and only those EIOs that are causally or constitutively relevant for the phenomenon that is to be explained. But what are mechanistic phenomena?

Generally, in philosophy of science the term 'phenomenon' is used in a rather unspecific way. It is understood as an umbrella term for all features of the world that are of scientific interest (Frigg and Hartmann 2018). On the basis of this, one way to think about mechanistic phenomena might be to define them as those features of the world that scientists are interested in. But this is a non-starter, for we run into a similar dilemma as arose in the context of discussing the causal role theory of functions in Chap. 3: either we define phenomena in terms of *contemporary science*, in which case some phenomena will be left out, since contemporary science is surely incomplete; or we define phenomenon in terms of an *ideal science*, in which case we simply do not know what mechanistic phenomena are because we do not know what an ideal science will imply.

Even though a characterization of phenomena in terms of what scientists are interested in is not helpful, analyzing paradigmatic examples of mechanistic

[1] As argued before, some authors adopt a strong ontic view of explanation, according to which phenomena *are* explananda, and hence explananda exist mind-independently. As argued in the introduction, I adopt a weak ontic view, according to which explananda are descriptions that *refer* to phenomena.

B. Krickel, *The Mechanical World*, Studies in Brain and Mind 13,
https://doi.org/10.1007/978-3-030-03629-4_6

phenomena taken from actual science is crucial to ensure descriptive adequacy. Indeed, providing examples is one common starting point of the new mechanists (Bechtel and Abrahamsen 2005, 422–23; Craver 2007, 4–5). Prominent examples of mechanistic phenomena that are discussed in the mechanistic literature are

(i) spatial memory (Craver and Darden 2001; Craver 2007; Bechtel and Abrahamsen 2008; Darden 2008; Harbecke 2010; Sullivan 2010)
(ii) neurotransmitter release (Machamer et al. 2000; Bogen 2005; Craver 2007; Andersen 2012)
(iii) the action potential (Machamer et al. 2000; Bechtel and Abrahamsen 2005; Craver 2006, 2007)
(iv) protein synthesis (Machamer et al. 2000; Darden and Craver 2002; Bechtel and Abrahamsen 2005; Craver 2007; Illari and Williamson 2010; Craver and Darden 2013)
(v) muscle contraction (Garson 2013; Gebharter and Kaiser 2014; Piccinini 2015; Glennan 2016)
(vi) the heart pumping blood (Bechtel and Abrahamsen 2005; Bechtel 2006; Glennan 2010; Craver and Darden 2013)

The examples differ as to whether they are phenomena explained by *constitutive* or *etiological* mechanistic explanations (see Chap. 2; in what follows, I speak of *etiological* and *constitutive mechanistic phenomena*, respectively). For instance, example (ii) is an example of an etiological mechanistic phenomenon because it is explained by a mechanism that causally produces it (Craver 2007, 22). The others are supposed to be explananda of constitutive mechanistic explanations since they refer to phenomena that are explained by an underlying mechanism. It remains unclear whether, from a metaphysical point of view, phenomena of etiological and constitutive mechanistic explanations differ (for a discussion of this issue, see Kaiser and Krickel (2017)).

Based on the considerations regarding causation in Chap. 4, we can state what the explananda of *etiological* mechanistic explanation are. I have argued for an activity-based account of causation. According to this account, the basic units of causation are EIOs that mechanistically interact. Hence, the phenomena of etiological mechanistic explanations must be EIOs. Take example (ii), which according to Craver describes an etiological mechanistic phenomenon. Neurotransmitter release is a complex EIO in the sense described in Sect. 4.4, Chap. 4. Since the question regarding the nature of etiological mechanistic phenomena is, thus, already settled, the focus of the present chapter will be *constitutive* mechanistic phenomena.

In this chapter, I discuss and reject a view that is common in the new mechanistic literature: the view that constitutive mechanistic phenomena are *capacities*. My argument, roughly, is that this view is incompatible with the metaphysics of EA-mechanisms as described in the previous chapters. An alternative view that can be found in the new mechanistic literature, and that is prima facie compatible with the metaphysics of EA-mechanisms, is the view that constitutive mechanistic phenomena are *behaving systems*. I will present two interpretations of this claim: according to what I will call the *functionalist view*, constitutive mechanistic phenomena are behaviors of mechanisms characterized by input–output relations. According to what I will call the *behaving entity view*, constitutive mechanistic phenomena are higher-level entities that contain mechanisms that are engaged in an occurrent (as defined in Chap. 4). I will argue that the functionalist view is flawed since it conflicts with the general aims of the new mechanists, such as defending the autonomy of the special sciences (see Introduction), and defending a specific notion of levels of nature (see Chap. 5). I will show that the behaving entity view is compatible with these general goals. Hence, I will conclude, constitutive mechanistic phenomena, like etiological mechanistic phenomena, are EIOs.

6.1 Mechanisms Do Not Explain Capacities

A common view in the new mechanistic literature is that constitutive mechanistic explanations, in contrast to etiological mechanistic explanations, explain *capacities* (Cummins 1975; Couch 2011; Piccinini and Craver 2011; Weiskopf 2011; Ylikoski 2013). This view is prima facie plausible since life scientists are in fact interested in explaining, for example, the capacity of humans to navigate through familiar environments, as in the case of spatial memory research. All paradigmatic examples listed above can be reformulated in terms of capacities. Still this view is problematic. As I will show in this section, the main reason is that explanations of capacities do not refer to mechanisms as characterized in the previous chapters (see also Kaiser and Krickel 2017).

Take the disposition of glass of being fragile (Ylikoski 2013). The explanation of why glass is fragile will mention the bonding between the molecules of the glass, which is rather weak such that it can be easily broken. This explanation mentions certain entities that have certain properties, and dispositions that manifest in a certain way if a certain stimulus were to obtain. Although this might be a valid explanation, it does not refer to a mechanism. As argued in Chap. 4, mechanisms consist of entities and *occurrents*, and not of entities and *dispositions*. The explanation, rather than being mechanistic, is a description of entities that are *disposed* in such a way that a mechanism *would* occur if a certain stimulus were present.

One might object that this is only an artifact of the formulation of the explanation. Indeed, the explanation of the fragility of glass refers to entities and their activities/occurrents, which becomes obvious when we formulate the conditional that describes the disposition: ‘Glass is fragile because when a sufficiently large force is executed on it, the bonding of the molecule breaks.’ The *breaking* of the bonding is an occurrent. Although, again, this might be a valid explanation, it does not show that capacities are explained by *constitutive* mechanistic explanations. First, since the occurrent of the glass's breaking occurs later than the execution of the force, we are dealing with an etiological mechanistic explanation, rather than a constitutive one. Second, if we argue that what happens instantaneously to the glass's breaking (the breaking of the bonding between the molecules) is what explains the disposition of breaking, we no longer explain why the glass is *disposed* to break but, rather, what happens when the glass *is breaking*. As soon as we introduce activities or other occurrents into the explanation, the phenomenon changes—we now constitutively explain a *manifestation* of the disposition rather than the disposition itself. Hence, we should not think of mechanistic phenomena as capacities.

Note that I do not want to argue that scientists are not interested in explaining capacities. Nor do I want to argue that mechanisms do not play a crucial role in explaining capacities. Rather, what the above considerations show is that the explanatory relation between mechanisms and capacities is an indirect one: mechanisms explain the *manifestations* of capacities. The explanation of the capacity can be inferred from the explanation of its manifestations. For example, when we know what constitutes the manifestation of the glass's breaking (the breaking of the bonding of the molecules), we can infer that the capacity of the glass to break is due to the capacity of the glass's molecules to break.

What, then, are constitutive mechanistic phenomena? Indeed, the considerations of the previous chapters already suggest a specific view. I have shown that, according to the new mechanists, mechanisms come in hierarchies of levels of mechanisms. On the one hand, the notion of a mechanistic level was defined as relating phenomena at higher levels and components of mechanisms at lower levels; on the other hand, the hierarchy was supposed to arise due to the fact that every component can be regarded as a phenomenon relative to which lower-level mechanistic components exist. Hence, from a metaphysical point of view, phenomena are the same kinds of things as mechanistic components: they are EIOs (Kaiser and Krickel 2017). This view seems to be in line with many implicit and explicit claims that can be found in the new mechanistic literature. Usually, the phenomenon that is to be explained is described as a "behavior of a system" (Bechtel 1994; Glennan 2005; Craver 2006, 2007; Wimsatt 2006; Hüttemann and Love 2011; Kaplan and Craver 2011; Piccinini 2015). Unfortunately, the terms 'system' and 'behavior' remain highly ambiguous. I will show that there are at least two different ways in which mechanistic authors think about what behaving systems are: I call these views the *functionalist view* of constitutive phenomena and the *behaving entity view* of constitutive phenomena. I present and discuss both views in the following sections, and I will show that only the behaving entity view gives rise to a promising account of constitutive mechanistic phenomena.

6.2 The Functionalist View of Constitutive Mechanistic Phenomena

According to the functionalist view of constitutive mechanistic phenomena, as I understand it here, the system whose behavior is to be explained is the *mechanism* itself. This idea underlies many discussions in the new mechanistic literature (Bechtel and Abrahamsen 2005; Craver 2007; Fazekas and Kertész 2011; Fagan 2012; Illari and Williamson 2012). For example, in Craver's famous diagram (Craver 2007, 7), the phenomenon is referred to as 'S's ψ-ing' and Craver states: "I often refer to the phenomenon, the property or behavior explained by the mechanism, as ψ […], and I use S […] to refer to *the mechanism as a whole*" (Craver 2007, 6–7). He expresses the same idea when specifying what he takes a constitutive mechanistic explanation to be:

> Mechanistic explanations are constitutive or componential explanations: they explain *the behavior of the mechanism as a whole* in terms of the organized activities and interactions of its components. (Craver 2007, 128; my emphasis)

A similar view can be found in Melinda Fagan's presentation of what she calls the "joint account" (2012). She summarizes the merits of her account as follows:

> Finally, [the joint account] resolves ambiguity concerning the target of explanation: the overall mechanism (M ψ-ing) rather than its downstream effects (P). The explanandum is a description of M ψ-ing […]. (Fagan 2012, 467)

Bechtel and Abrahamsen seem to endorse this view, too:

> Another point that is important to appreciate is that identifying the component parts and operations of a mechanism and their organization is only part of the overall endeavor of developing a mechanistic explanation. […] Nonetheless, it is crucial to identify them and to explore how variations affect *the behavior of the mechanism*. (Bechtel and Abrahamsen 2005, 426; my emphasis)

Hence, the idea that what is to be explained in a constitutive explanation is a behavior of the mechanism itself seems to be a common assumption among the new mechanists.

In addition to the claim about the *mechanism's* behavior as the explanandum of constitutive explanations, the behavior that is to be explained is characterized in terms of a complex input–output relation or a causal role (Craver 2007, 214; Bechtel 2008, 201–202; Fazekas and Kertész 2011; Baetu 2012; Kuorikoski 2012, 146; Soom 2012; Casini and Baumgartner 2017). These inputs and outputs are connected by the mechanism. The combination of the claims (i) that the explanandum is the behavior of *a mechanism*, and (ii) that the behavior can be characterized in terms of *inputs* and *outputs* of the mechanism, at least implicitly underlies many discussions. For example, when describing different kinds of interlevel experiments, Craver (2007, 146) argues that "[i]n each case, the goal is to show that X's ϕ-ing is causally between the inputs and outputs that constitute S's ψ-ing." Similarly, Kuorikoski describes constitutive mechanistic explanations as explaining system-level properties in the following way:

> That the system-level property f_{system} is realized by the causal structure means that the causal dependencies of the structure (f_1 … f_4) provide a more fine-grained picture of *how the causal inputs to* f_{system} *lead to its causal outputs*. (Kuorikoski 2012, 375; my emphasis)

Bechtel seems to endorse the functionalist view of mechanistic constitution when he describes the mechanism of fermentation:

> For example, we conceptualize the fermentation mechanism in yeast as taking in sugar and out-putting alcohol. Typically, the reactions are diagrammed linearly: sugar is shown at one end and arrows (reactions) lead the eye through a sequence of intermediate products to alcohol at the other end […] Additional chemical substances, such as inorganic phosphate

> (Pi), oxidized and reduced nicotinamide adenine dinucleotide (NAD+, NADH), and adenosine diphosphate and triphosphate (ADP, ATP) enter and leave the main linear pathway in what are typically appended as "side reactions." The focus is on the main pathway: fermentation as a way to turn grapes and grains into alcohol that we can enjoy drinking. (Bechtel 2008, 202)

These quotations show that many authors assume that the description of the phenomenon specifies the inputs and outputs of a particular mechanism. This is what I call the *functionalist view* of constitutive mechanistic phenomena.

(*Functionalist View of Constitutive Mechanistic Phenomena*) Constitutive mechanistic phenomena are characterized in terms of the inputs into, and the outputs out of the mechanism that constitutes the phenomenon at hand.

There are two possible interpretations of the metaphysics of the functionalist view of mechanistic phenomena in line with what is often called *realizer* and *role functionalism* (McLaughlin 2007; Levin 2016). Realizer functionalism implies that the realizer and the realizee are in fact identical. The characterization of the phenomenon in terms of an input–output relation is a means to identify the realizer, in the present case the mechanism, that connects the inputs with the outputs. For example, protein synthesis might be characterized as whatever process starts with mRNA molecules leaving the cell nucleus and ends with there being new proteins; and then the mechanism that realizes this input–output relation is found, which is then identified with protein synthesis. Hence, the phenomenon just is the mechanism under a functional description. The phenomenon turns out to be identical with the mechanism (Fazekas and Kertész (2011) and Soom (2012) argue that the new mechanists are committed to this identity claim).

(*Realizer Functionalist View of Constitutive Mechanistic Phenomena*) Constitutive mechanistic phenomena just are the mechanisms that constitute them under a functional description.

The realizer functionalist view of constitutive mechanistic phenomena nicely captures a certain reasoning strategy in the life sciences: scientists use 'black boxes' or 'filler terms', which, after careful investigation, are filled with assumptions about entities and occurrents and their organization (i.e., with details about the mechanism) (Craver 2007; Piccinini and Craver 2011). According to Craver, scientists often characterize phenomena in terms of causal roles if the phenomenon is "some-process-we-know-not-what" (Craver 2007, 114). The research goal, then, is to specify the process (i.e., the mechanism) that actually is the phenomenon that has thus far been characterized in terms of an input–output relation only.

A second way to interpret the functionalist view of constitutive mechanistic phenomena is *role* functionalism. According to role functionalists, the phenomenon is *the causal role*, rather than the realizer of that role. Hence, it is a relational higher-order property (McLaughlin 2007). According to this reading, constitutive mechanistic phenomena are not identical with the mechanism. Rather, they are relational properties *realized* by mechanisms. For example, protein synthesis would simply be the property of using mRNA (etc.) to produce new proteins. The protein synthesis mechanism constitutes protein synthesis in the sense that it fills this causal role, i.e., it instantiates the property of having the causal role of producing new proteins from mRNA.

In the context of the new mechanistic approach, role functionalism is a non-starter. The reason is that it is incompatible with the general metaphysical convictions of the new mechanists, according to which, first, only entities and occurrents exist, and second, phenomena are supposed to have spatiotemporal parts. Metaphysically speaking, properties are nothing but entities and occurrents in some sense. Hence, the role functionalist interpretation either collapses into the realizer functionalist interpretation or is metaphysically dubious (e.g., how can phenomena have spatiotemporal parts if they are abstract relational properties?). In a nutshell: the realizer functionalist view is the only plausible candidate to make sense of the functionalist view of constitutive mechanistic phenomena. Thus, if one wants to be functionalist with regard to phenomena, one has to assume that phenomena turn out to be identical with their mechanisms (under a functional description).

> (*Role Functionalist View of Constitutive Mechanistic Phenomena*) Constitutive mechanistic phenomena are causal roles (realized by the mechanism).

But there are good reasons not to be a functionalist with regard to constitutive mechanistic phenomena. Although metaphysically sound, the realizer functionalist view is incompatible with the broader goals of the new mechanists. First, one central motivation for many new mechanists was to argue for the autonomy of the special sciences (Bechtel 2007; Craver 2007, Chap. 7). Bechtel defends the autonomy of the special sciences by arguing that knowledge about lower levels is insufficient for inferring knowledge about the behavior of higher-level phenomena. He highlights that the higher levels provide information that the corresponding lower levels do not contain, namely *organizational* and *contextual* information (Bechtel 2007, 182–83). Adopting a realizer functionalist view of constitution seems to defeat this goal. Since this view implies that the phenomenon just is the mechanism, one can no longer uphold the view that there can be information about the phenomenon that is not implied in the information about the mechanism (for a similar line of argument, see Fazekas and Kertész 2011, 380–81). Since all knowledge about the phenomenon just is knowledge about its causal role and the realizer of that causal role (that, therefore, has the *same* causal role) there cannot be any knowledge about the phenomenon that is not already implied in the knowledge about the mechanism.

Second, the identity claim conflicts with the notion of a mechanistic level as discussed in Chap. 5. If phenomena just are mechanisms, there are no levels of nature at all since there are no distinguishable relata. Nor does it help to say that a mechanistic hierarchy relates explananda and explanantia, i.e., representations of mechanisms. Mechanistic levels are supposed to be *levels of nature* (Craver 2007, 177ff.) that relate things *in the world*, rather than epistemic constructs, descriptions, models, or the like.

Third, Craver and Bechtel (2007) and Bechtel (2016) attempt to provide an account of top-down causation in terms of mechanistically mediated effects that is supposed to make sense of downward-causation talk in the sciences without being committed to a mysterious metaphysical picture. Mechanistically mediated effects are supposed to account for talk about downward causation without rendering it mysterious by interpreting the downward relation in terms of a horizontal, intra-level causal relation and a vertical, inter-level constitution relation. For example, my playing tennis does not cause my cells to start using more glucose (Craver and Bechtel 2007, 559), but rather my playing tennis is *constituted* by my muscles moving, which *causes* them to metabolize the available ATP to ADP which in the end *causes* the glycolysis. The effect is not caused by the tennis playing but mediated via the mechanism that constitutes the tennis playing.

If the realizer functionalist picture were correct, the tennis playing would be identical to the activity of the muscles that constitute it. Fazekas and Kertész (2011, 366–67) show that this is in conflict with the account of mechanistically mediated effects, as it makes this account redundant. If phenomena are identical with mechanisms, downward causation just is horizontal causation. Hence, no need for mechanistically mediated effects. Indeed, Bechtel (2016) admits that his objectors are right in this respect.[2]

> This exegesis of Craver's diagram suggests that the critics who viewed Craver and my account as rendering higher levels epiphenomenal were right. It suggests a highly reductionistic picture of levels according to which causal relations that were supposed to be between entities at higher levels of organization dissolve into causal interactions at the lowest level considered. (Bechtel 2016)

Fourth, a further central motivation for highlighting the relevance of constitutive explanations is to stress the importance of *structural decomposition* of a system into relevant and irrelevant parts (Bechtel and Abrahamsen 2005; Craver 2007, 109). Structurally decomposing a system means to find the structural, i.e., spatiotemporal

[2] Indeed, at this point it is unclear whether Bechtel takes his opponents to argue that his view of levels implies a reductionist view with regard to levels (i.e., an identity between the levels) or an epiphenomenalism with regard to higher-level phenomena. Epiphenomenalism implies a non-reductionist claim with regard to higher-level phenomena (i.e., they are not identical with lower-level phenomena) but implies that, due to this irreducibility, the higher-level phenomena are causally inert. A reductionist view with regard to higher levels implies that higher-level phenomena are causally efficacious but only due to their being identical with lower-level phenomena.

parts (the entities and activities) of a system that are relevant to the system's behavior that one wants to explain. The realizer functionalist interpretation of constitutive mechanistic phenomena cannot make sense of the idea that structural decomposition consists in decomposing the systems whose behaviors we want to explain into relevant and irrelevant parts, given that they assume that the systems just are the mechanisms. This consequence of the realizer functionalist view can be show with help of the following deductive argument:

1. Each phenomenon just is the mechanism that constitutes it. [assumption: realizer functionalist view]
2. Mechanisms are composed of those and only those entities and occurrents that are relevant to the phenomenon. [assumption from Chap. 5]
3. Phenomena are composed of those and only those entities and occurrents that are relevant to it. [from 1 and 2]
4. *Structural decomposition* of a system means to distinguish between relevant and irrelevant parts of the system relative to a particular behavior of that system. [assumption]
5. Structurally decomposing phenomena into relevant and irrelevant parts is redundant as there are no irrelevant parts of phenomena. [from 3 and 4]

It follows that the distinction between relevant and irrelevant parts of the behaving systems that are the phenomena is empty, as there are no irrelevant parts if the realizer functionalist view is presupposed.

Fifth, given that the realizer functionalist view identifies the phenomenon with the mechanism, this view vitiates the first criterion of Craver's account of constitutive relevance ('X's ϕ-ing is a part of S's ψ-ing,' see Chap. 5, Sect. 5.2). If the phenomenon (S's ψ-ing) *just is* the mechanism, this condition amounts to the requirement that X's ϕ-ing must be a spatiotemporal part of the mechanism. But this is exactly what we want to get from an account of constitutive relevance. The realizer functionalist view amounts to saying that an X's ϕ-ing is a component of a mechanism iff it is a component of that mechanism. A consequence is that we end up in an epistemic circle, since we have to know the components of a mechanism in order to be able to determine its components. Bechtel (2016) does not seem to be aware of this circle when he argues that

> [i]n the life sciences, investigators developing explanations often (1) *begin by identifying the mechanism* responsible for a specific phenomenon to be explained, (2) proceed to *decompose the mechanism* into its parts and the operations they perform […]. (Bechtel 2016, my emphasis)

It is at best unclear how mechanistic explanations are supposed to succeed if it is presupposed that one knows the mechanism (i.e., the entities, occurrents, and their organization) in order to identify the phenomenon (or, in Bechtel's case, in order to know where to look for parts).

Finally, the functionalist view of constitutive mechanistic phenomena is guilty of committing the mereological fallacy and the reification fallacy as introduced in Chap. 4, Sect. 4.1. First, by holding that mechanistic phenomena are behaviors of mechanisms, one ascribes predicates to parts that can only be ascribed to the whole. It is not the moving mechanism that moves—it is, say, the car that moves; it is not the spatial memory mechanism that navigates the Morris water maze—it is the mouse; it is not the contracting mechanism that contracts—it is the muscle that does so. Second, if one takes phenomena to be behaviors of mechanisms, one treats a system that consists of various acting and interacting entities as one unified entity. It is a category mistake to say that, for example, the mechanism for muscle contraction is an entity such that it can be engaged in contracting behavior. This would be gerrymandering. The mechanism for muscle contraction is *composed* of various entities and activities (interacting actin and myosin filaments) that are *responsible* for the contracting of the muscle, but they do not together form an entity (additional to the muscle) that contracts (Kaiser and Krickel 2017).

In a nutshell: the functionalist view fails as an account of constitutive mechanistic phenomena. Fortunately, there is an alternative interpretation of the idea that phenomena are 'behaving systems.' I will present this view in the next section.

6.3 The Behaving Entity View of Constitutive Mechanistic Phenomena

Implicit in the new mechanistic thinking is a second view on constitutive mechanistic phenomena that is, as I will show, compatible with the overall goals of the new mechanistic approach. I call this view the *behaving entity view*.

The crucial difference between the functionalist view and the behaving entity view is that, according to the latter, the system whose behavior is to be explained is not the mechanism but a *larger entity that contains the mechanism*. This interpretation is suggested by Craver's discussion of spatial memory (see Chap. 5, Sect. 5.3). He identifies spatial memory as a multi-layered phenomenon, where higher-level behaving entities contain lower-level behaving entities: a mouse navigating the Morris water maze at the highest level contains the hippocampus generating spatial maps, which contains neurons inducing long-term potentiation, which again contain NMDA-receptor activating at the lowest level. At each level there is an entity (the mouse, the hippocampus, a neuron, an NMDA-receptor) that contains the lower-level mechanism. Similarly, Glennan (2002, 1996) seems to think about phenomena in this way. He takes mechanisms to be located in larger entities or systems such as watches, cells, organisms, and toilets (see Chap. 2, Sect. 2.2).[3]

Similarly, Gillett takes mechanistic constitution to hold between larger entities that contain mechanisms and their parts (Gillett 2013, 327–328). He quotes the following passage from Craver 2007:

[3] Note that, as shown in Chap. 2, Sect. 2.2, Glennan calls the systems/objects 'mechanisms', and not what is going on inside of them.

> Not all parts are components […]. The hubcaps, mudflaps, and the windshield are all parts of the automobile, but they are not part of the mechanism that makes it run. They are not *relevant* parts of that mechanism. (Craver 2007, 140)

Gillett comments on this passage:

> Notice that here we have a higher level individual, the car, whose properties allow it to move around. But the automobile has many individuals that are parts of it. Craver is interested in why certain parts of this higher-level individual are counted as elements of the process of moving itself around and others are not? (Gillett 2013, 326)

Here Gillett (and Craver) seems to assume that constitutive mechanistic explanations explain behaviors of larger entities (individuals), like cars, that contain various mechanisms that are responsible for different behaviors the larger entity can be engaged in (e.g., moving around). The crucial task is to identify those parts of the larger entity that are relevant for the particular behavior that is to be explained and that are, thus, part of the constitutive mechanistic explanation.

A second assumption that characterizes the behaving entity view of constitutive mechanistic phenomena is that the behaviors that are to be explained are activities, or rather occurrents (as characterized in Chap. 4, Sect. 4.2), that the larger entity is engaged in. These occurrents might be characterized in terms of inputs and outputs. But these input–output descriptions do not exhaustingly characterize the relevant occurrent. As argued in the previous chapter, although they might be picked out in terms of inputs and outputs, from a metaphysical perspective they do not reduce to them (Illari and Williamson 2011, 2013; Machamer et al. 2000, 5). The combination of the claims that (i) what is to be explained are behaviors of *larger entities* containing mechanisms and (ii) that behaviors are (irreducible) *occurrents* or *activities*, is often summarized by the claim that the explanantia of constitutive mechanistic explanations are "acting entities" (Craver 2007, 189). In Chap. 4 I argued that we should think of entities and activities in terms of *entity-involving occurrents* (EIOs).[4]

(*The Behaving Entity View of Constitutive Mechanistic Phenomena*) Constitutive mechanistic phenomena are EIOs that contain the mechanism that constitute them (such that all components of the mechanism are spatial EIO-parts of the phenomenon).

The behaving entity view (or *EIO-view*) of constitutive mechanistic phenomena is more promising than the functionalist view with regard to the general goals of the new mechanistic approach. The reason is that it is not committed to the view that the phenomenon *just is* the mechanism. First, the behaving entity that contains the

[4] Note that constitutive mechanistic phenomena can be *simple* (like a muscle contracting) or *complex* (protein-protein binding or osmosis) EIOs. In line with the considerations made in Chap. 4, complex phenomenon-EIOs contain mechanisms in the sense that all entity-components of the mechanism are parts of one of the entities that participates in the phenomenon EIO. All occurrent-components of the mechanism occur during the occurrence of the complex phenomenon-EIO.

mechanism is usually larger than the mechanism that is responsible for the behavior. For example, the moving car has parts that are not parts of the driving mechanism; the contracting muscle has parts that are not relevant to its contracting, and so on.

Second, higher-level behaving entities are what Gillett calls *qualitatively different* from lower-level entities (Gillett 2002, 2010, 2013; Gillett and Aizawa 2016).

> [W]e must carefully mark that the various entities bearing 'making-up' relations in these cases, and the many like it, are usually qualitatively distinct—that is, the relata of these relations usually differ in their features. […] And a survey of any number of examples of mechanistic explanation in the sciences, or the entities found at the distinct 'levels' related by such explanations, establishes that the relevant relata are usually of *qualitatively different* kinds. (Gillett 2010, 172)

The qualitative distinctness shows that the relata cannot be identical (Gillett 2010, 174). Gillett highlights the difference in *powers*. Take the example of muscle contraction: the contracting muscle has powers (i.e., moving a limb) that no actin filament has. Similarly, the contracting muscle has powers that even the whole contracting mechanism does not have (i.e., the power to swell, the power to displace a certain amount of water). Gillett primarily focuses on the relation between higher-level entities and the *components* of the mechanisms that are responsible for their behaviors. Here, we are concerned with the relation between the behaving entity and the mechanism as a whole. In our case, the qualitative distinctness seems to go even further: the higher-level behaving entity and the mechanism belong to two different metaphysical categories. While the contracting muscle is an entity that shows a certain behavior, the contracting mechanism is a continuous causal sequence consisting of various entities and behaviors. Hence, it would involve a category mistake to say that the phenomenon (the higher-level behaving entity) is identical with the mechanism.

Since the behaving entity view implies the non-identity between higher-level phenomena and lower-level mechanisms, we can do justice to the general goals of the new mechanists that I already discussed in the last section. First, one can reformulate Bechtel's arguments for the autonomy of higher-level sciences (Bechtel 2007). As mentioned before, Bechtel argues that higher levels provide information that the corresponding lower levels do not contain, namely organizational and contextual information (Bechtel 2007, 182–83). Clearly, the level of the muscle's contracting provides information that the level of the interacting actin and myosin filaments does not provide. For example, the higher level contains information about when the muscle is contracted and to which degree (which is a kind of temporal organization), and the exact size and shape of the muscle (which is a kind of spatial organization). This kind of information we do not get by merely looking at actin and myosin filaments, and their interactions. Furthermore, we get contextual information about how the muscle's contraction influences other muscles, where it is located relative to other body parts, and how the contraction of the muscle changes depending on the activity of motor neurons. Again, we do not get this kind of information from merely looking at the interacting actin and myosin filaments.

Second, the behaving entity view can make sense of the notion of mechanistic levels as *levels of nature*. Contracting muscles, as well as interacting actin and myosin filaments, are distinct things that exist in the world. Hence, the behaving entity view can make sense of the idea that there are two *distinguishable*, mind-independent relata of mechanistic levels.

Third, the notion of a mechanistically mediated effect can straightforwardly be applied if the behaving entity view is presupposed. Causation, according to the account of mechanistically mediated effects, can only be intra-level between, for example, actin and myosin filaments, but the actin or myosin filaments do not cause the behavior of the muscle, nor does the muscle cause any behaviors of the filaments.[5] Rather, the interaction between myosin and actin filaments constitutes state s_1 of the behaving muscle at t_1, then ATP causes a change in the interaction at t_2, where at t_2 the changed myosin and actin interaction constitutes the changed state s_2 of the behaving muscle.

Fourth, the behaving entity view can make sense of structural decomposition, as we can now make sense of the idea that phenomena have relevant and irrelevant parts. Behaving entities can have irrelevant parts. Clearly, not every part of a muscle is relevant to its contracting behavior. Hence, the behaving entity view can make sense of the claim that "[n]ot all parts are components" (Craver 2007, 140).

Fifth, we can make sense of the first condition of Craver's mutual manipulability criterion without depriving it of its content: X's ϕ-ing is a component of a mechanism only if it is a spatial EIO-part of the behaving larger entity. Furthermore, we avoid the epistemic circle since we do not have to know the mechanism before we can identify its components. Rather, the contracting muscle can be identified independently of the myosin and actin filaments.

Finally, if we adopt the behaving entity view we can avoid committing the mereological and the reification fallacies. Predicates like 'driving,' 'navigating,' and 'seeing' are not ascribed to mechanisms but to entities such as cars (that contain driving mechanisms), mice (that contain spatial memory mechanisms), and organisms (that contain vision mechanisms). The reification fallacy is avoided since it is descriptively adequate to the sciences and everyday talk to treat cars, mice, and organisms as real unified entities.

One might object against the behaving entity view that it cannot account for all sorts of constitutive mechanistic phenomena. First, one might doubt whether the behaving entity view captures the phenomenon of spatial memory. How is spatial memory a behaving entity? 'Spatial memory' is an umbrella term for various different kinds of phenomena. Indeed, this poses a challenge for spatial memory research since in order to perform experiments on spatial memory one first has to determine what is to count as an instance of that phenomenon (this is part of what is called *operationalization*[6]). First, the description of the phenomenon has to be

[5] Note that I accept this claim here only for the sake of argument. In Chap. 7 I show that there can be interlevel causation in mechanisms. See also Krickel 2017.

[6] The term was introduced by the physicist Percy Bridgman (1882–1961), claiming that "in general, we mean by a concept nothing more than a set of operations; the concept is synonymous with the corresponding sets of operations" (Bridgman 1927, 5). Here, I use the notion of operationalization in its methodological reading, according to which operationalizations are definitions of

disambiguated (Feest 2010). Spatial memory is often defined as the ability of mammals to navigate through familiar environments. Thus, in the first step, the rather diffuse phenomenon description 'spatial memory' is clarified as referring to 'the mammalian ability to navigate through familiar environments.' The validity of a concrete specification of a concept like 'spatial memory' is usually restricted to particular disciplines, or even studies, and is often only temporary. As Feest (2005) argues, these concept specifications serve the purpose "to get empirical investigations off the ground" (2005, 134). After disambiguating the diffuse phenomenon description, the second step consists in specifying the experimental setup. It has to be specified which *particular entities* are to be investigated that count as instances of the phenomenon referred to by the description developed in the disambiguation step. Spatial memory, for example, in the end, is operationalized in terms of behaviors of *single individuals*, such as mice or human subjects that navigate through mazes (or the like). Hence, spatial memory can be analyzed in terms of entity-involving occurrents.

How does the behaving entity view handle the *generalizations* that are often the explananda of mechanistic explanations? Take for example the generalization 'When an action potential reaches the axon terminal, the release probability is p.' In line with the view on regularity as presented in Chap. 3, these generalizations depend on their being concrete tokens, which are behaving entities, that resemble each other in certain ways. For example, the generalization concerning the release probability of axon terminals describes the fact that there are various axon terminals (tokens) that release neurotransmitters when an action potential occurs. The probability mentioned in the generalization describes the fact that among all axon terminals (tokens), only some release neurotransmitters when an action potential occurs. If one wants to explain *why* the probability has the specific value p, one has to compare those axon terminals (tokens) that release neurotransmitters with those that do not. Hence, we can make good sense of the idea that generalizations describe resemblances between behaving entities.

6.4 Summary

In this chapter I developed an approach to mechanistic phenomena. I argued that the nature of etiological mechanistic phenomena straightforwardly follows from the account of causation defended in the previous chapter: they are entity-involving occurrents. I discussed two views of constitutive mechanistic phenomena that can be found in the literature. According to one view, mechanistic phenomena are capacities. This view turned out to be incompatible with the metaphysics of

---

concepts that are "either temporary assumptions about typical empirical indicators of a given subject matter, which allow[s] researchers to get empirical investigations 'off the ground', or they [are] presentations of the outcomes of experiments, which [are] assumed to individuate a given phenomenon particularly well" (Feest 2005, 134).

EA-mechanisms as analyzed in the previous chapters. Capacities can be mechanistically explained only *via* their manifestations that are caused or constituted by mechanisms.

According to a second view, constitutive mechanistic phenomena are behaving systems. I argued that there are two interpretations of this view. Many authors seem (at least implicitly) to hold that the phenomena explained in constitutive mechanistic explanations are behaviors of mechanisms that are characterized in terms of input–output relations. I call this the *functionalist view* of constitutive mechanistic phenomena. Although this view (at least its realizer functionalist interpretation) is compatible with the metaphysics of mechanisms and their components, it is incompatible with the general goals of the new mechanistic approach. Since, according to this view, constitutive mechanistic phenomena are identical with their underlying mechanisms, this view cannot make sense of the autonomy of the special sciences, mechanistic levels as levels of nature, mechanistically mediated effects, and the notion of structural decomposition. The alternative interpretation, what I called the *behaving entity view*, does justice to these goals because it can make sense of the idea that phenomena and mechanisms are distinct. According to this view, constitutive mechanistic phenomena are EIOs that contain mechanisms. As a consequence, constitutive mechanistic phenomena turn out to be the same kinds of things as etiological mechanistic phenomena.

Now that we have a clear understanding of what constitutive mechanistic phenomena are, we can start thinking about what mechanistic *constitution* is. This will be the topic of the next chapter.

References

Andersen, H. K. (2012). The case for regularity in mechanistic causal explanation. *Synthese, 189*, 415–432. https://doi.org/10.1007/s11229-011-9965-x.

Baetu, T. M. (2012). Filling in the mechanistic details: Two-variable experiments as tests for constitutive relevance. *European Journal for Philosophy of Science, 2*, 337–353. https://doi.org/10.1007/s13194-011-0045-3.

Bechtel, W. (1994). Biological and social constraints on cognitive processes: The need for dynamical interactions between levels of inquiry. *Canadian Journal of Philosophy, 24*, 133–164.

Bechtel, W. (2006). *Discovering cell mechanisms*. Cambridge: Cambrdige University Press.

Bechtel, W. (2007). Reducing psychology while maintaining its autonomy via mechanistic explanations. In M. Schouten & H. Looren de Jong (Eds.), *The matter of the mind: Philosophical essays on psychology, neuroscience and reduction* (pp. 172–198). Oxford: Basil Blackwell. https://doi.org/10.1017/CBO9781107415324.004.

Bechtel, W. (2008). *Mental mechanisms. Philosophical perspectives on cognitive neuroscience*. New York/London: Routledge.

Bechtel, W. (2016). Explicating top-down causation using networks and dynamics. *Philosophy of Science*. 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*, 421–441. https://doi.org/10.1016/j.shpsc.2005.03.010.

Bechtel, W., & Abrahamsen, A. (2008). From reduction back to higher levels. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the cognitive science society* (pp. 559–564). Austin: Cognitive Science Society.

Bogen, J. (2005). Regularities and causality; generalizations and causal explanations. *Studies in History and Philosophy of Science Part C :Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 397–420. https://doi.org/10.1016/j.shpsc.2005.03.009.

Bridgman, P. (1927). *The logic of modern physics*. New York: Arno Press.

Casini, L., & Baumgartner, M. (2017). A Bayesian theory of constitution. *manuscript*.

Couch, M. B. (2011). Mechanisms and constitutive relevance. *Synthese, 183*, 375–388. https://doi.org/10.1007/s11229-011-9882-z.

Craver, C. F. (2006). When mechanistic models explain. *Synthese, 153*, 355–376. https://doi.org/10.1007/s11229-006-9097-x.

Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. New York: Oxford University Press.

Craver, C. F., & Bechtel, W. (2007). Top-down causation without top-down causes. *Biology and Philosophy, 22*, 547–563. https://doi.org/10.1007/s10539-006-9028-8.

Craver, C. F., & Darden, L. (2001). Discovering mechanisms in neurobiology: The case of spatial memory. In P. Machamer, R. Grush, & P. McLaughlin (Eds.), *Theory and method in neuroscience* (pp. 112–137). Pittsburgh: University of Pitt Press.

Craver, C. F., & Darden, L. (2013). *Search of mechanisms. Discoveries across the life sciences*. Chicago/London: University of Chicago Press.

Cummins, R. (1975). Functional analysis. *The Journal of Philosophy, 72*, 741–765.

Darden, L. (2008). Thinking again about biological mechanisms. *Philosophy of Science, 75*, 958–969. https://doi.org/10.1086/594538.

Darden, L., & Craver, C. F. (2002). Strategies in the interfield discovery of the mechanism of protein synthesis. *Studies in History and Philosophy of Science Part C :Studies in History and Philosophy of Biological and Biomedical Sciences, 33*, 1–28. https://doi.org/10.1016/S1369-8486(01)00021-8.

Fagan, M. B. (2012). The joint account of mechanistic explanation. *Philosophy of Science, 79*, 448–472. https://doi.org/10.1086/668006.

Fazekas, P., & Kertész, G. (2011). Causation at different levels: Tracking the commitments of mechanistic explanations. *Biology and Philosophy, 26*, 365–383. https://doi.org/10.1007/s10539-011-9247-5.

Feest, U. (2005). Operationism in psychology: What the debate is about, what the debate should be about. *Journal of the History of the Behavioral Sciences, 41*, 131–149. https://doi.org/10.1002/jhbs.20079.

Feest, U. (2010). Concepts as tools in the experimental generation of knowledge in cognitive neuropsychology. *Spontaneous Generations: A Journal for the History and Philosophy of Science, 4*, 173–190. https://doi.org/10.4245/sponge.v4i1.11938.

Frigg, R., & Hartmann, S. (2018). Models in science. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*, Summer 201. Metaphysics Research Lab, Stanford University.

Garson, J. (2013). The functional sense of mechanism. *Philosophy of Science, 80*, 317–333. https://doi.org/10.1086/671173.

Gebharter, A., & Kaiser, M. I. (2014). Causal graphs and biological mechanisms. In M. I. Kaiser, O. R. Scholz, D. Plenge, & A. Hüttemann (Eds.), *Explanation in the special sciences: The case of biology and history* (pp. 55–85). Dordrecht: Springer. https://doi.org/10.1007/978-94-007-7563-3_3.

Gillett, C. (2002). The dimensions of realization: A critique of the Standard view. *Analysis, 62*, 316–323.

Gillett, C. (2010). Moving beyond the subset model of realization: The problem of qualitative distinctness in the metaphysics of science. *Synthese, 177*, 165–192. https://doi.org/10.1007/s11229-010-9840-1.

Gillett, C. (2013). Constitution, and multiple constitution, in the sciences: Using the neuron to construct a starting framework. *Minds and Machines, 23*, 309–337. https://doi.org/10.1007/s11023-013-9311-9.

Gillett, C., & Aizawa, K. (2016). *Scientific composition and metaphysical ground*. London: Palgrave Macmillan. https://doi.org/10.1057/978-1-137-56216-6.

Glennan, S. (1996). Mechanisms and the nature of causation. *Erkenntnis, 44*, 49–71. https://doi.org/10.1007/BF00172853.

Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science, 69*, S342–S353. https://doi.org/10.1086/341857.

Glennan, S. (2005). Modeling mechanisms. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 443–464. https://doi.org/10.1016/j.shpsc.2005.03.011.

Glennan, S. (2010). Ephemeral mechanisms and historical explanation. *Erkenntnis, 72*, 251–266. https://doi.org/10.1007/s10670-009-9203-9.

Glennan, S. (2016). Mechanisms and mechanical philosophy. In P. Humphreys (Ed.), *The Oxford handbook of philosophy of science* (Vol. 1). New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199368815.013.39.

Harbecke, J. (2010). Mechanistic constitution in neurobiological explanations. *International Studies in the Philosophy of Science, 24*, 267–285. https://doi.org/10.1080/02698595.2010.522409.

Hüttemann, A., & Love, A. C. (2011). Aspects of reductive explanation in biological science: Intrinsicality, fundamentality, and temporality. *British Journal for the Philosophy of Science, 62*, 519–549. https://doi.org/10.1093/bjps/axr006.

Illari, P. M. K., & Williamson, J. (2010). Function and organization: Comparing the mechanisms of protein synthesis and natural selection. *Studies in History and Philosophy of Science Part C :Studies in History and Philosophy of Biological and Biomedical Sciences, 41*, 279–291. https://doi.org/10.1016/j.shpsc.2010.07.001.

Illari, P. M. K., & Williamson, J. (2011). Mechanisms are real and local. In *Causality in the sciences* (pp. 818–844). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199574131.003.0038.

Illari, P. M. K., & Williamson, J. (2012). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science, 2*, 119–135. https://doi.org/10.1007/s13194-011-0038-2.

Illari, P. M. K., & Williamson, J. (2013). In defence of activities. *Journal for General Philosophy of Science, 44*, 69–83. https://doi.org/10.1007/s10838-013-9217-5.

Kaiser, M. I., & Krickel, B. (2017). The metaphysics of constitutive mechanistic phenomena. *The British Journal for the Philosophy of Science, 68*, 745–779. https://doi.org/10.1093/bjps/axv058.

Kaplan, D. M., & Craver, C. F. (2011). The explanatory force of dynamical and mathematical models in neuroscience: A mechanistic perspective. *Philosophy of Science, 78*, 601–627. https://doi.org/10.1086/661755.

Krickel, B. (2017). Making sense of interlevel causation in mechanisms from a metaphysical perspective. *Journal for General Philosophy of Science, 48*, 453–468. https://doi.org/10.1007/s10838-017-9373-0.

Kuorikoski, J. (2012). Mechanisms, modularity and constitutive explanation. *Erkenntnis, 77*, 361–380. https://doi.org/10.1007/s10670-012-9389-0.

Levin, J. (2016). Functionalism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy.* Metaphysics Research Lab, Stanford University.

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1–25.

McLaughlin, B. P. (2007). Mental causation and shoemaker-realization. *Erkenntnis, 67*, 149–172. https://doi.org/10.1007/s10670-007-9069-7.

Piccinini, G. (2015). *Physical computation: A mechanistic account*. Oxford: Oxford University Press.

Piccinini, G., & Craver, C. F. (2011). Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese*. https://doi.org/10.1007/s11229-011-9898-4.

Soom, P. (2012). Mechanisms, determination and the metaphysics of neuroscience. *Studies in History and Philosophy of Science Part C :Studies in History and Philosophy of Biological and Biomedical Sciences, 43*, 655–664. https://doi.org/10.1016/j.shpsc.2012.06.001.

Sullivan, J. A. (2010). Reconsidering "spatial memory" and the Morris water maze. *Synthese, 177*, 261–283. https://doi.org/10.1007/s11229-010-9849-5.

Weiskopf, D. A. (2011). Models and mechanisms in psychological explanation. *Synthese, 183*, 313–338. https://doi.org/10.1007/s11229-011-9958-9.

Wimsatt, W. C. (2006). Aggregate, composed, and evolved systems: Reductionistic heuristics as means to more holistic theories. *Biology and Philosophy, 21*, 667–702. https://doi.org/10.1007/s10539-006-9059-1.

Ylikoski, P. (2013). Causal and constitutive explanation compared. *Erkenntnis, 78*, 277–297. https://doi.org/10.1007/s10670-013-9513-9.