

To appear in: Glennan, S./ Illari, P. (eds.): The Routledge Handbook of Mechanisms and Mechanical Philosophy. New York: Routledge.

## 9. The Components and Boundaries of Mechanisms

*Marie I. Kaiser*

Most new mechanists agree that mechanisms are made of two kinds of components (see Chapter 1). The first kind of components are material objects that are variously called entities (Machamer, Darden and Craver (MDC) 2000), parts (Glennan 1996, 2002) or component parts (Bechtel 2006, 2008). Examples include neurotransmitters, muscle fibers and genes. The second kind of components are variously called activities (MDC 2000), interactions (Glennan 1996, 2002) or operations (Bechtel 2006, 2008). They are what entities do and what produces change (MDC 2000). Examples include binding, contracting and being transcribed. I will use the term ‘component’ to refer to these two kinds of things that make up or compose mechanisms.<sup>1</sup>

Since the origin of the new mechanical philosophy metaphysical questions about the components of mechanisms (hereinafter ‘mechanistic components’) figure prominently in the debate (e.g., Tabery 2004, Psillos 2004, Machamer 2004, Bogen 2008, Torres 2009, Illari and Williamson 2013). MDC (2000) provoked this dispute by defending a dualistic ontology of mechanistic components, according to which mechanisms are made up components belonging to two distinct ontological categories: entities (having certain properties) and activities. MDC contrast their dualism with Glennan’s (1996, 2002) allegedly monistic view that mechanisms consist of entities (having certain properties) and of interactions, which are nothing but occasions on which a property change of one entity brings about a change in properties of another entity.

The dispute between dualists and monists involves two sets of metaphysical questions: first, how can the two kinds of mechanistic components be further characterized? What are their major features and what distinguishes them from each other? How are entities/parts and activities/interactions/operations related to more well-known ontological categories, such as material objects, properties, processes, events, and dispositions? I address these questions in Section 9.1. Second, what is the relation between the two kinds of mechanistic components? Is one more fundamental than the other? Can activities, for instance, be reduced to entities and their properties? In what sense are the two kinds of mechanistic components mutually dependent? I examine these questions in Section 9.2.

The second part of this chapter (Sections 9.3 and 9.4) is concerned with the individuation of the components of mechanisms and with the boundaries of mechanisms. Mechanisms are always mechanisms *for* specific phenomena or behaviors (Glennan 1996, 2002). Accordingly, the components of a mechanism are individuated with respect to the phenomenon the mechanism is responsible for (see Chapters 1 and 8): Only those entities/parts and activities/interactions/operations that are relevant to the phenomenon count as components of the mechanism. This raises the question of how to spell out the conditions of relevance. What makes entities/parts and activities/interactions/operations relevant to a specific phenomenon? How do we draw the boundary between what belongs to a mechanism and what does not? In Section 9.3, I present different criteria for individuating the components of mechanisms – natural boundaries, robustness/stability, strength of interactions, mutual manipulability, and INUS conditions – and I discuss their merits and limitations.

The claim that the individuation of a mechanism (by individuating all of its components) depends on the characterization of the respective phenomenon invites the question of how real the boundaries of mechanisms are. Mechanistic antirealists, as I call them, argue that mechanisms do not exist as well-delineated entities in nature because it is the scientist who imposes boundaries on certain sets of components while pursuing specific explanatory interests and adopting specific perspectives (Bechtel 2015; see also Wimsatt 1972, 2007). In Section 9.4, I examine whether such an antirealistic move is inevitable and which arguments can be offered to defend the realistic view that mechanisms exist in nature and have real boundaries and a determined set of components.

## **9.1 The Ontological Nature of Mechanistic Components**

The goal of this section is to provide a general characterization of the two kinds of components that make up mechanisms. I will refer to mechanistic components as entities and as activities because entity-activity terminology has prevailed in the debate (even Glennan forthcoming speaks about activities now) and because there are good reasons for this terminological choice: First, the term ‘entity’ is more suitable than the term ‘part’ because, as I argue somewhere else (Kaiser forthcoming), being a component of a biological mechanism is not the same as being a biological part of a biological object in general. The components of a mechanism are individuated with respect to a single behavior, whereas the parts of biological objects are individuated with respect to all of their characteristic behaviors. Moreover, natural boundaries are crucial to the individuation of biological parts but they do not constrain the individuation of mechanistic components in the same way (more on this in

Section 9.3). Second, we can drop Bechtel's (2006, 2008) term 'operation' because he does not tie any metaphysical claim to this term. Third, against the term 'entity' one might object that it is commonly used as a placeholder term for *any* ontological category (including activities) and that we should be more precise and speak about material objects instead (Kaiser and Krickel 2016). But in the context of this handbook it is advisable to stick to the term 'entity' in this specific usage because it is an established term in the debate. Fourth, I prefer the term 'activity' over the term 'interaction' because interactions can plausibly be seen as special kinds of activities. Fifth, against the concept of an activity one might object that it is too restrictive and should be replaced by the broader metaphysical concept of an occurrence (which includes processes, events, and states; Kaiser and Krickel 2016). For the present purposes, however, I stick to the familiar concept of an activity while critically discussing what activities are.

I begin here by considering what entities are (as components of mechanisms). The mechanism for muscle contraction is composed of entities such as actin filaments (i.e., strands of actin molecules), calcium ions, and sarcoplasmic reticula (i.e., cell organelles involved in the storage and release of calcium ions). What the new mechanists refer to as 'entities' are concrete *material objects* that are located in space and time. Typically, entities are spatially extended, have certain shapes and sizes, and often are surrounded by characteristic spatial boundaries, such as the membrane comprising the sarcoplasmic reticulum. Entities exist in time and are wholly present at different time points. Contrary to activities, entities are not, themselves, temporally extended in the sense of having temporal parts.<sup>2</sup> A calcium ion in my muscle fiber now is wholly present; it does not consist of the calcium ion now, the calcium ion two days ago, the calcium ion in one minute, etc. Entities can engage in processes, events, or activities which are temporally extended (e.g., diffusing through a channel). But the calcium ion as such is not temporally extended. In metaphysical terms, entities are *continuants* not *occurents*, as activities are (Kaiser and Krickel 2016).

Entities are *bearers of properties*, including dispositional properties (or "causal powers"; Glennan forthcoming). Properties allow entities to engage in specific activities (MDC 2000, 3; Craver and Darden 2013, 16). For example, a sarcoplasmic reticulum is surrounded by a membrane that contains different kinds of ion channels and ATP-driven calcium pumps. These properties enable the sarcoplasmic reticulum to engage in the activity of releasing calcium ions.

Psillos (2004, 312) draws our attention to the fact that entities can exist without actually engaging in activities (e.g., if entities do not manifest their dispositions). I agree that entities, in general, can exist without engaging in any activities. But the important point is that entities can only be components of mechanisms if they engage in activities. Without doing something an entity cannot be relevant to the phenomenon the mechanism is responsible for. All conditions of relevance require something to occur, something to be active (see Section 9.3). Hence, component entities are necessarily *involved in activities*, that is, they are “working entities” (Darden 2008, 961; Craver and Darden 2013, 18). But we need to be cautious at this point. Entities that are components of mechanisms need not be active at all times. For the working of some mechanisms it is important that entities engage in activities only at certain times, not at others. For example, the sarcoplasmic reticulum releases calcium ions when an action potential arrives, not during later stages of the mechanism for muscle contraction. To conclude, an entity can be a component of a mechanism only if it is involved in an activity at least once while the mechanism proceeds from its beginning to its end.

Component entities, however, can be involved in activities in different ways. They can have an *active role* and initiate or maintain the activity (i.e., be “actors”; Glennan forthcoming), such as the ATP molecule that binds to the calcium pump. On the other hand I think entities can also be *passively* involved in an activity and allow for or undergo the activity, such as the calcium pump to which the ATP molecule binds. This difference has not been recognized in the debate so far. Furthermore, there can be activities in which two or more entities with an active role are involved, such as the sliding of actin and myosin filaments past each other. According to the strict reading of the word, such cases would be examples of *interactions*. In the debate, however, the term ‘interaction’ is used in a wider sense as referring to any activity in which more than one entity is involved – let it be passively or actively (e.g., Tabery 2004, Glennan forthcoming).

Finally, it has been argued that, as components of mechanisms, entities must be “stable clusters of properties” (Craver 2007a, 131) and must “have a kind of robustness and reality apart from their place within that mechanism” (Glennan 1996, 53). For example, ATPases (i.e., transmembrane proteins that transport ions to produce ATP) can be found in a variety of different mechanisms in different organisms and they retain their properties also if studied in isolation, that is, if studied in different contexts than *in situ* (Kaiser 2015, 221-235). Even though *stability* and *robustness* may not succeed as criteria for identifying the components of mechanisms (see Section 9.3) they might still be important features that many component entities share.

### Main Features of Entities

- (1) Entities are *material objects* (i.e., continuants).
- (2) Entities are *bearers of properties*, which allow them to engage in specific activities.
- (3) As components of mechanisms, entities necessarily *engage in activities* (at least once during the mechanism).
- (4) Entities can be *actively or passively involved* in activities.
- (5) Entities are relatively *stable* and *robust*.

I move now to considering what activities are (as components of mechanisms). These are the second kind of components that make up the mechanism for muscle contraction: activities such as sliding, releasing, and binding. “Activities are the things that the entities do” (Craver/Darden 2013: 16). As such, activities are necessarily *temporally extended* and possess characteristic durations, rates, and phases. For example, the binding of ATP to the calcium pump starts with the collision of the ATP molecule with a certain region of the pump, then certain chemical bonds are formed. Because activities do not only exist in time – as entities do as well – but are also extended in time and have temporal parts (i.e., stages), they belong to the metaphysical category of occurrents, which encompasses processes, events, and states (Kaiser and Krickel 2016).

Activities are thought to be “active rather than passive” and to be the “happenings” (Machamer 2004: 29) in a mechanism. It is this active nature of activities that seems to make them indispensable if one wants to account for the alleged activeness of mechanisms (MDC 2000: 5), and it is their active nature that is said to distinguish activities from, for instance, property instantiations or property changes (Glennan forthcoming rejects this; more on this in Section 9.2). As I show elsewhere, it follows from the active nature of activities and from their ontological status as occurrents that activities must be *actualized* or manifest rather than merely potential or dispositional (Kaiser and Krickel 2016). Only manifestations of dispositions can be components of mechanisms; unmanifested dispositions cannot. It is the release of calcium ions that is a component of the mechanism for muscle contraction, not the unmanifested disposition of the sarcoplasmic reticulum to release calcium ions if a depolarization occurs.

Dualists claim that activities account not only for the active nature of mechanisms but also for their productivity (MDC 2000; Tabery 2004; Bogen 2008; Darden 2008). In their

view, the order of activities in a mechanism exhibits a “productive continuity” (MDC 2000: 3): one activity *productively causes* the next, which ensures that a mechanism runs in its typical way from beginning to end (see Chapter 10). For example, the release of calcium ions into the cytosol causes the binding of calcium ions to troponin molecules, which causes the tropomyosin complex to move off the actin binding site, which causes the myosin head to bind to the actin filament, and so on. Hence, activities are characterized as the “producers of change” (MDC 2000: 3) and as the “causal components in mechanisms” (Craver 2007a: 5).<sup>3</sup>

Another central feature of activities is that they *require entities* that engage in them. As MDC put it, “there are no activities...that are not activities *of entities*” (2000: 5). Any activity requires at least one entity that engages in it. I think we can add to this that for each activity there must be at least one entity that is *actively* involved in it. That is, activities seem to require actors. For example, a gene can be passively engaged in the activity being transcribed but there must be another entity that is actively involved in the same activity, such as the DNA polymerase.

Different types of activities involve different numbers of entities. In other words, activities have “*unrestricted arity*” (Illari and Williamson 2013: 72; my emphasis). Binary activities, such as the binding of ATP to the calcium pump or the sliding of actin and myosin filaments past each other, are widespread and also activities that involve many entities, such transcription or osmosis, are common in the living world. In my view, there are also clear examples of activities that involve only one entity (so called unary activities or “un-interactive activities” Torres 2009: 243), such as the changing of the spatial conformation of a protein, the breaking of a chemical bond, and the closing of a stoma cell in a plant leaf. Other authors, however, question the existence of un-interactive activities (Tabery 2004; Fagan 2012). They claim that for an activity to be productive it must involve not less than two entities and thus be an *interaction* (or a “jointly acting complex”; Fagan 2012: 464). Tabery (2004) argues that seemingly un-interactive activities, such as shifting conformation of a protein, still involve interactions on a lower level. This argument is not convincing because the claim was that there is no other entity (besides the protein) involved in the conformation shift. This is compatible with there being interactions among the parts of the protein. Furthermore, Fagan (2012) fails to provide an argument for why all component entities must form complexes that jointly act. I conclude that there are cases of unary activities and that interactions thus constitute a mere subset of the set of all activities.

One might want to add the sixth feature that activities do not only produce changes (i.e., act as causes) but also, themselves, *involve changes* in the properties of the engaged

entities (i.e., are processes). An example that supports this claim is the release of calcium ions from the sarcoplasmic reticulum into the cytosol. This activity involves several changes, such as the calcium ions changing their location and the sarcoplasmic reticulum changing its concentration of stored calcium ions. Despite its initial plausibility, I think we should reject the claim that activities *must* involve changes because it overlooks the variety of mechanistic components. For the working of some mechanisms, it is crucial that certain properties are not changed but maintained (i.e., continuously instantiated) during a certain time span. For instance, it is essential for the mechanism of the action potential that some ion-channels remain open for a certain period of time. Moreover, the mechanism of natural selection seems to consist also of “passive properties” (Skipper and Millstein 2005: 341), such as being camouflaged or being present (see Chapter 23). Examples like these show that not all component activities involve changes, which is why this feature is not listed below.<sup>4</sup>

#### **Main Features of Activities**

- (1) Activities are *temporally extended* (i.e., occurrents).
- (2) Activities are *actualized* (rather than merely potential).
- (3) Activities *produce change* (i.e., are types of causes).
- (4) Activities require at least one *actively involved entity*.
- (5) Activities have *unrestricted arity* (i.e., involve one to many entities).

## **9.2 The Relation between Mechanistic Components**

This section examines the relation between entities and activities as components of mechanisms. Is the relation one of reduction? Can entities, for instance, be said to be more fundamental than activities? Or are both kinds of components ontologically on a par, for instance, because they are mutually dependent? These questions trace back to the already introduced dispute between dualists (MDC 2000, Machamer 2004) and monists (Glennan 1996, 2002) that has been prominent in the new mechanical philosophy. In this section, I first clarify in what respects dualists and monists have different views about the relation between mechanistic components. Then I critically discuss the arguments that have been provided in favor of dualism.

To begin with clarifying the dualism versus monism debate, MDC argue that “[m]echanisms are composed of both entities (with their properties) and activities” (2000: 3). I call the claim that mechanisms consist of components of two kinds *Duality thesis*. This is a quite weak

thesis because monists can accept it as well. Glennan argues that mechanisms are “complex systems” that consist of “parts” (1996: 52; 2002: 344) having relatively stable properties and interacting with each other in certain ways. Hence, even though Glennan (1996, 2002) denies that mechanisms are composed of entities and of activities, he also recognizes a second kind of components, namely interactions. In his recent work, Glennan (forthcoming) agrees with even MDC’s version of the *Duality thesis* because he refers to the components of mechanisms as entities (or parts) and activities and characterizes interactions as a special, important class of activities.

Contrary to dualists, however, monists reject the *Irreducibility thesis*. This is the claim that the two kinds of mechanistic components belong to two distinct ontological kinds in the sense that one cannot be reduced to the other. Glennan characterizes an interaction between parts as an “occasion on which a change in a property of one part brings about a change in a property of another part” (2002: 344). In other words, Glennan (1996, 2002) assumes that interactions are reducible to the properties of entities and rejects the *Irreducibility thesis*. He is thus only committed to the existence of entities (i.e., material objects) and their properties, including dispositional properties. By contrast, MDC (2000) postulate the existence of activities on top of the usual commitments to entities and properties. Because of this additional ontological category MDC’s approach is referred to as dualistic and contrasted with Glennan’s monistic account.<sup>5</sup>

Dualism is characterized by two further assumptions. The first is the *Interdependency thesis* which says that entities and activities necessarily exist together and determine each other. MDC state that “no activities without entities, and entities do not do anything without activities” (2000: 8). This is why the new mechanists use terms such as “acting entities” (Craver 2007a: 189) or “working entities” (Darden 2008: 961). The *Interdependency thesis* does not reinforce the dualistic character of MDC’s approach – one might even argue that it runs contrary to it – but it is central to their approach. Second, MDC (2000) defend what I call the *Parity thesis*. It is the claim that there is no priority of, for instance, entities over activities and that entities and activities are ontologically on a par (Illari and Williamson 2013: 70).<sup>6</sup> Strictly speaking, the *Parity thesis* is no independent thesis because it follows from the dualists’ assumptions of irreducibility and interdependency.

In sum, MDC’s (2000) dualism can be distinguished from Glennan’s (1996, 2002) monism by the following three claims:



### **Core Theses of Dualism**

- (1) *Irreducibility thesis*: Entities and activities belong to two distinct ontological kinds because activities cannot be reduced to the properties of entities.
- (2) *Interdependency thesis*: In mechanisms, entities and activities necessarily exist together and determine each other.
- (3) *Parity thesis*: Entities and activities are ontologically on a par.

In his recent work, Glennan (forthcoming) seems to have partly converged to MDC's (2000) dualistic position. He speaks of entities (or parts) and activities as the components of mechanisms and he agrees with MDC in that there cannot be activities without entities, nor entities without activities (which is the *Interdependency thesis*). Moreover, Glennan rejects the reproach that his account is guilty of an entity-bias. That is, he seems to want to accept the *Parity thesis*. This is, however, the point at which the remaining difference between MDC (2000) and Glennan (forthcoming) becomes apparent. Glennan rejects the label of dualism and the claim that entities and activities are two distinct ontological categories because he thinks that the ontological distinctness of entities and activities is incompatible with the interdependency of entities and activities. In other words, he sees an incompatibility between the *Irreducibility thesis* (in particular, the distinctness claim it contains) and the *Interdependency thesis*. On my view, this casts into doubt Glennan's approval of the *Parity thesis* because if entities and activities are not ontologically distinct and activities can be reduced to property changes of entities, it seems highly questionable that entities and activities are ontologically on a par. To conclude, Glennan (forthcoming) has converged to dualism by accepting the existence of activities and the *Interdependency thesis* but he still rejects the distinctness claim that the *Irreducibility thesis* contains and thus fails to convincingly defend the *Parity thesis*.

Having examined the nature of the disagreement, I will now go on to explore challenges to dualism. Even though entity-activity talk dominates the debate, dualism and the ontological category of activities are still disputed (e.g., Woodward 2002, Psillos 2004, Machamer 2004, Tabery 2004, Bogen 2008, Torres 2009, Illari and Williamson 2013, Glennan forthcoming). In this section, I introduce and critically discuss five major arguments that have been offered in defense of dualism.

The new mechanists agree with other philosophers that philosophy of science should pay close attention to actual scientific practice. They accept “*descriptive adequacy*” (MDC 2000: 8; Kaiser 2015: 9) as an important criterion of adequacy not only for epistemic but also for metaphysical claims (Kaiser and Krickel 2016). Dualists make use of this criterion and argue that only dualism is descriptively adequate because it accounts for the fact that scientists describe mechanisms in terms of entities and activities (i.e., by using nouns, such as ‘enzyme’ and ‘repressor’, and verbs, such as ‘inhibits’) and because successful scientific practice is a practice of studying and manipulating activities (Illari and Williamson 2013: 73-5; see also Craver 2007a, 144-52). This might be convincing but it should be noted that the claim that activities play an important role in scientific practice does not represent a brute fact about science. It is already a philosophical interpretation or critical reconstruction of specific parts of scientific practice which can be contested (Kaiser 2015: Chapter 2). In my view, a monist might have equally good reasons for viewing scientific practice as a practice of studying objects, their properties and how these properties change through time. Descriptive adequacy can thus also be used as an argument *against* dualism.

A second argument stresses the “*epistemic adequacy*” (MDC 2000: 21) of dualism. Dualists argue that thinking about mechanisms in terms of entities and activities renders scientific phenomena intelligible because entity-activity language corresponds best to how we naturally describe the world (Machamer 2004: 31; Illari and Williamson 2013: 79). For example, we say that a neurotransmitter binds to a receptor (activity talk), not that a neurotransmitter changes from being unbound to being bound to a receptor (property talk), nor that a neurotransmitter manifests its bonding capacity (disposition talk). Even if this claim raises questions about the naturalness of descriptions and even if it may have exceptions I share the intuition that activity talk is more comprehensible than property or disposition talk and I agree that activity talk corresponds well to how scientists describe and explain phenomena. When compared to an entity-disposition ontology, dualism is claimed to be epistemically more adequate because we can observe a disposition of an entity only through the disposition being manifested, that is, through the entity acting (MDC 2000: 4; Machamer 2004: 30; Illari and Williamson 2013: 78). This is true but proponents of an entity-disposition ontology can accept the conceptual dependency of dispositions on their manifestations without being committed to the existence of activities in particular. In general, one might object that all of these arguments refer to mere epistemic points that have no metaphysical relevance (Psillos 2004: 313). But I agree with the dualists that metaphysicians of science should pay attention to the epistemic practices in the sciences.

MDC also appeal to common *metaphysical intuitions* to support dualism. They claim that dualism is advantageous because it captures “healthy philosophical intuitions underlying both substantivalist and process ontologies” (2000: 4). The ontological category of entities accounts for the fact that activities or processes are not free-floating but that there is always an entity involved in an activity or process. The ontological category of activities, in turn, accounts for the fact that mechanisms are dynamic and “do things” (MDC 2000, 5) and thus cannot consist of entities having certain properties only (recall Section 9.1; Kaiser and Krickel 2016). Dualism is able to account for these motivating intuitions because it conceives of entities and activities as ontologically distinct and on a par. The problem with this argument is that it remains unclear to which kind of intuitions the dualists refer. The argument would be plausible if they referred to a sort of pre-theoretical intuitions that metaphysical theories must account for (even then there might be disagreement about which intuitions are legitimate). However, MDC speak about philosophical intuitions and seem to have something more theoretical and more metaphysically laden in mind.

Another argument in favor of dualism is that only activities sufficiently explain the activeness and the *productivity* of mechanisms (MDC 2000, Tabery 2004, Bogen 2008). Only activities specify *how* in a mechanism one property change brings about another property change (i.e., how one entity interacts with another). In Tabery’s words, the ‘bringing about’ in Glennan’s interactions are “black boxes” (2004: 10) that can only be rendered intelligible by activities. For instance, the activity releasing is said to specify how the change of the sarcoplasmic reticulum’s ion channel from closed to open brings about the shift in the location of some calcium ions. Dualists claim that if you reject the *Irreducibility thesis* and argue that activities are nothing but property changes you will overlook the active and productive nature of mechanisms. This concern seems to be confirmed, for example, by Glennan’s thesis that “mechanisms are things (or objects)” (2002: 345), such as watches or cells. As I argue somewhere else, from a dualistic perspective, it becomes clear that bare objects that are neither active nor productive (i.e., that do not involve any occurrent), such as a stopped watch, cannot be mechanisms (Kaiser and Krickel 2016). But the argument that only dualism accounts for the productive nature of mechanisms encounters also serious problems. First, dualists fail to provide a clear and metaphysically satisfying analysis of what productivity is (e.g., Woodward 2002; Psillos 2004). Second, the “black box problem” seems to be not a problem for monists but rather “is one for scientists to solve in the laboratory” (Torres 2009: 239) because scientists are the ones to discover how, exactly, one particular property change brings about another.

A final argument in favor of dualism pertains to *parsimony*. Illari and Williamson state that an entity-activity ontology is more parsimonious than an entity-disposition ontology (2013: 79-81). At first sight, this claim might be surprising because dualism accepts the existence of activities on top of the usual ontological commitments and thus entails a rather unparsimonious move. However, Illari and Williamson do not compare dualism with monism, but rather an entity-activity ontology with an entity-disposition ontology. They argue that, due to the restricted arity of dispositions (they attach to only one entity), an entity-disposition ontology results in a proliferation of indefinitely many dispositions, which is why an entity-activity ontology is more parsimonious. This argument raises the question which kind of parsimony is most important. On which level of graininess should we evaluate parsimony? If the goal of metaphysics is to reveal the fundamental structure of reality I doubt that it is worse to have fewer general ontological categories (entity, property) and more numerous ontological sub-categories (different entities, many different properties).

### 9.3 Individuating Mechanistic Components

The new mechanists agree that mechanisms consist of all and only those components that are relevant to the specific phenomenon that the mechanism is responsible for. The term “responsible for” (Bechtel and Abrahamsen 2005: 422; Illari and Williamson 2012: 123) can be specified in two ways: etiological mechanisms *cause* their phenomena, whereas constitutive mechanisms *constitute* their phenomena (Kaiser and Krickel 2016). Figure 1 illustrates the triangular relationship between mechanistic components, a mechanism, and a phenomenon.

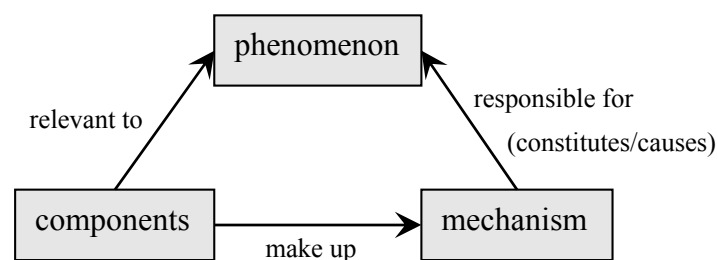


Fig. 1: Relationships between components, mechanisms, and phenomena

Figure 1 raises the question, for instance, how we should spell out the relevance relation between components and phenomena. This question has an epistemic and a metaphysical side. Epistemically, we are asking for a criterion of explanatory relevance: What are the explanatorily relevant factors that must be represented in a mechanistic explanation of

a given phenomenon? How do we distinguish adequate mechanistic explanations from mere descriptions?<sup>7</sup> Metaphysically, the question of relevance is a question about where the boundary of a particular mechanism runs: Which entities and activities belong to the mechanism for a specific phenomenon? Which criteria tell apart genuine components of a mechanism from mere pieces, correlates, redundant parts, or background conditions?

The two types of mechanisms allow us to distinguish two kinds of relevance relations. The components of etiological mechanisms are causally relevant to their phenomena, whereas the components of constitutive mechanisms are said to be constitutively relevant to their phenomena. Since the notion of causal relevance has been subject to extensive philosophical discussion the new mechanists focus on the notion of “constitutive relevance” (Craver 2007a: 139-60; 2007b; see Chapter 14).<sup>8</sup> I conclude this section by reviewing different possible criteria for individuating the components of constitutive mechanisms.

One might first suggest identifying the boundaries of mechanisms with what I call “natural boundaries” (Kaiser 2015: 176) of biological objects (e.g., Darden seems to suggest this; 2008: 960). Paradigmatic examples of natural boundaries are the cell membrane, the exoskeleton of insects, the skin, the blood-brain barrier, and the chain of mountains that borders a particular ecosystem. I agree that the existence of natural boundaries is essential to the working of many mechanisms. For example, the mechanism for muscle contraction requires that the muscle fiber is surrounded by a membrane that functions as a selective barrier and receives extracellular signals. In most cases, however, these natural boundaries will not be the boundaries of the mechanism but of the corresponding biological object (e.g., of the muscle fiber).

According to my account, the boundary of a mechanism differs in two respects from the boundary of the corresponding object (what Craver refers to as ‘S’ in his characterization of the phenomenon ‘S’s  $\psi$ -ing’; Kaiser and Krickel 2016). First, mechanisms frequently transgress natural boundaries (Craver 2007b: 9). For instance, the mechanism for muscle contraction is composed also of entities, such as neurotransmitters, that are located outside the membrane of the muscle fiber. Also the gecko adhesion mechanism includes entities that belong to the gecko’s environment (Kaplan 2012: 552). Hence, natural boundaries seem to be crucial to identifying the parts of biological objects in general (e.g., muscle fibers or geckos; Kaiser forthcoming) but not to individuating the components of mechanisms (e.g., the mechanism for muscle contraction or the gecko adhesion mechanism).

Second, biological objects, such as cells or geckos “do many things at once” (Glennan 1996: 52) and all of these characteristic behaviors are relevant to individuating the parts of the biological object (Kaiser forthcoming). By contrast, mechanisms are responsible only for a single phenomenon (e.g., muscle contraction or gecko adhesion) and mechanistic components must be relevant to this phenomenon only. Hence, a mechanism typically includes many fewer entities than are parts of the corresponding system. To conclude, natural boundaries fail to provide us with adequate criteria for demarcating mechanisms because the condition of being located inside a natural boundary is neither necessary nor sufficient for an entity to be a component of a mechanism.

A second idea for how to identify the boundaries of mechanisms is using strength of interactions. The idea that the intensity of interactions determines part-whole relations can be traced back to Simon (1962). This claim is based on the observation that, at least in “nearly decomposable” (1962: 474) systems, interactions among the parts of an object are generally stronger and more frequent than the interactions between an object’s parts and its environment. For example, the forces holding together a molecule are much weaker than those holding together atoms within the molecule. Other authors have picked up Simon’s idea and applied it to the boundaries of mechanisms (e.g., Haugeland 1998; Wimsatt 1974, 2007).

This account, however, faces various problems. In general, it is quite difficult to assess and compare the strength of interactions. We seem to need to specify a threshold that will depend on pragmatic considerations of researchers. In addition, the strengths-of-interaction-criterion seems unable to exclude background conditions and sterile effects as components of mechanisms (Craver 2007a: 143f). Finally, when looking at actual biological practice we recognize that the strength of interactions may play a crucial role in individuating biological objects such as populations or ecosystems (e.g., Hunemann 2014; Kaiser forthcoming) but not in individuating the components of biological mechanisms.

As a third possibility, we might turn to an account of constitutive relevance. Craver (2007a: 139-60; 2007b) has offered the most elaborated and widely discussed account of constitutive relevance. He identifies two conditions under which an acting entity ( $X$ ’s  $\varphi$ -ing) is constitutively relevant to a phenomenon ( $S$ ’s  $\psi$ -ing) and thus is a component of the mechanism for this phenomenon: the parthood condition and the mutual manipulability condition.

The parthood condition requires that “ $X$  is a part of  $S$ ” (Craver 2007a, 153) or that  $X$  is “contained within  $S$ ” (Craver and Tabery 2015). Hence, Craver seems to identify parthood with spatial inclusion. According to my view, the assumption that any component entity  $X$  must be spatially included in the corresponding object or system  $S$  encounters two major objections. First, it overlooks that temporal relations between component activities and the phenomenon of interest are important as well (Kaiser and Krickel 2016). Second, it conflicts with the fact that many mechanisms transgress natural boundaries (a fact that Craver, himself, recognizes; 2007a, 141) and include also entities and activities that are spatially located outside the corresponding object or system  $S$  (recall the discussion of natural boundaries in this section).

The second condition of Craver’s account of constitutive relevance relies on Woodward’s interventionist theory of causation (2003) and requires mechanistic components and phenomena to be mutually manipulable. This means that there must be an ideal intervention on the putative component  $X$ ’s  $\varphi$ -ing that changes the phenomenon  $S$ ’s  $\psi$ -ing and there must be an ideal intervention on  $S$ ’s  $\psi$ -ing that changes  $X$ ’s  $\varphi$ -ing. This part of Craver’s account of constitutive relevance has recently attracted much philosophical attention (e.g., Harbecke 2010; Couch 2011; Leuridan 2012; Baumgartner and Gebharder 2016; Harinen forthcoming). It has been objected that Craver’s account is merely epistemic and does not provide a satisfactory metaphysical analysis of what constitutive relevance or constitution really is (Harbecke 2010, Couch 2011). It seems to me that this criticism is too premature. Even though Craver’s analysis of the explanatory and investigative practices of neuroscience (2007) is not intended to be a metaphysical analysis it can easily be used to draw metaphysical conclusions about the relation of constitution (Kaiser and Krickel 2016). Another criticism is that Craver’s notion of constitutive relevance turns out to be a subtype of causal relevance and thus requires rethinking either Woodward’s notion of an ideal intervention or the mutual manipulability account altogether (Leuridan 2012; Baumgartner and Gebharder 2016; Harinen forthcoming; see also Woodward 2014).

Finally, I will consider what are known as ‘INUS conditions’. Regularity theories of constitution (Harbecke 2010; Couch 2011) develop criteria for individuating mechanistic components by making use of Mackie’s idea of a cause being an INUS condition (i.e., an Insufficient but Non-redundant part of an Unnecessary but Sufficient condition for the effect; Mackie 1965). The general idea is that a component is an insufficient but necessary part of an unnecessary but sufficient mechanism for a phenomenon. This idea nicely captures the fact

that a single component is insufficient for a phenomenon. A component must always be a non-redundant member of a “team” (Gillett 2013: 311) of components, which, together, are sufficient for the phenomenon. The criterion also accounts for the possibility that different particular mechanisms might be sufficient for a certain type of phenomenon.

It is problematic that regularity theories of constitution either require mechanistic phenomena to be capacities, which is an inadequate view of what constitutive mechanistic phenomena are (Kaiser and Krickel 2016), or they conceive of constitution as a relation between types only, which presupposes a realism about biological kinds that is heavily contested. Moreover, the claim that a mechanistic component must be a necessary member of a team of components raises the well-known problem of how to exclude mere correlates from being components (see also Craver and Tabery 2015).

#### **9.4 Are Boundaries of Mechanisms Real?**

A scientific realist might begin with the intuition that mechanisms exist out there in nature independently of the scientists discovering and representing these mechanisms. Where one mechanism ends and another starts seems to be an objective feature of reality, which depends neither on our ability to individuate mechanistic components nor on our having an account of constitutive relevance. Mechanisms seem to be real things with real boundaries.

Although many mechanists share these realist intuitions, a realist interpretation of the boundaries of mechanisms encounters serious challenges. A commonly held assumption is that how one identifies the components of mechanisms depends upon the phenomenon one seeks to understand (recall Section 9.3). Characterizing the phenomenon differently will yield different sets of relevant components and thus result in drawing the boundary of a mechanism differently. Since the choice of the phenomenon is relative to scientists’ explanatory interests some authors have argued that the identification and decomposition of mechanisms is an inherently “perspectival” matter (Darden 2008: 960; see also Kauffman 1971; Wimsatt 1972, 2007). From here it seems to be only a small step to the antirealistic claim that mechanisms do not exist as well-delineated things in the world but that scientists impose boundaries on mechanisms relative to their explanatory purposes (Bechtel 2015). The aim of this section is to discuss whether such an antirealistic move is inevitable and which arguments a realist might invoke to defend the independent existence of boundaries of mechanisms.

Let us begin with the claim that characterizing a phenomenon guides and constrains the individuation of the components of a mechanism. According to the new mechanists,



characterizing the phenomenon confines the space of possible mechanisms in so far as it uses the language of a given field and implicitly calls up a set of explanatory concepts and a “store of accepted entities, activities, and organizational structures that people in the field are licensed to use” (Craver and Darden 2013: 67). The relation between mechanisms and phenomena is thought to be reciprocal: not only does the nature of the phenomenon constrain what is to be included in the mechanism, findings about the mechanism can also force one to “reconstitute the phenomenon” (Bechtel and Richardson 2010: 173).

But does this imply that the boundaries of mechanisms are not real in the sense of being no objective feature of reality? I think it does not. A realist can agree that for different token phenomena there exist different token mechanisms with different boundaries and that the choice of scientists of which phenomena to study may change over time and is dependent on pragmatic factors such as the scientist’s perspective and explanatory interests. A realist would, however, emphasize that as soon as the phenomenon of interest is sufficiently specified and fixed the boundaries of the mechanism for this phenomenon are fixed, too, and do not depend on pragmatic factors. This way a realist can agree that the choice of the phenomenon is relative to the scientist’s interests and perspective while holding on to the idea that for each token phenomenon there exists a token mechanism out there in the world that has a well-delineated, real boundary. Such a realistic position is taken up, for instance, by Glennan (forthcoming), and it seems to be the position that Craver is also committed to (not least because of his ontic conception of explanation; 2014).

Other authors reject the realistic assumption that boundaries of mechanisms exist in nature independently of the fact that scientists take specific perspectives and pursue certain explanatory goals when decomposing mechanisms. Wimsatt, for instance, argues that in the case of “interactionally complex” systems, different “theoretical perspectives” give rise to different non-isomorphic decompositions of a system into parts with “non-coincident boundaries” (Wimsatt 1974: 69-71). This sounds quite antirealistic, as if Wimsatt claimed that parts and boundaries would exist only relative to theoretical perspectives. But at other places in his work, Wimsatt explicitly defends realism – though he is keen to add that his realism is a realism for a “messy world” and is “piecemeal and usually satisfied with a local rather than a global order” (2007: 5f.; see also 1974: 672).

We can find antirealism about the boundaries of mechanisms more explicitly in the recent work by Bechtel. He draws on research in Neurobiology to show that “mechanisms as bounded entities don’t exist” (2015: 84). Marom’s (2010) argument that the time-course of many phenomena in the life sciences is scale-free, in Bechtel’s view, does not challenge the

mechanistic account. It merely shows that the common view of mechanisms as temporally and spatially well-delineated entities is literally false – even though it is an idealization that facilitates developing scientific explanations. Bechtel argues that it is “the scientists who impose boundaries around entities and activities in nature and impose a time scale on which their functioning is characterized” (2015: 85). A realist may object that if boundaries of mechanisms do not exist in the actual world, mechanisms do not exist either, because it is indeterminate which entities and activities belong to which mechanism. Bechtel’s antirealism seems to result in the implausible view that the world consists of loose entities and activities only. Alternatively, Bechtel’s antirealism can be read as turning the mechanistic account into a purely epistemic account that has been detached from almost all ontological commitments.

## References

- Baumgartner, M., Gebharder, A. (2016): “Constitutive Relevance, Mutual Manipulability, and Fat-Handedness”, *British Journal for the Philosophy of Science*.
- Bechtel, W. (2008): *Mental Mechanisms: Philosophical Perspectives on Cognitive Neurosciences*, New York, Routledge.
- (2015): “Can mechanistic explanation be reconciled with scale-free constitution and dynamics?”, *Studies in History and Philosophy of Biological and Biomedical Sciences* 53, 84-93.
- Bechtel, W., Abrahamsen, A. (2005): “Explanation: a mechanist alternative”, *Studies in History and Philosophy of Biological and Biomedical Sciences* 36, 421-441.
- Bechtel, W., Richardson, R.C. (2010): *Discovering Complexity. Decomposition and Localization as Strategies in Scientific Research*, Cambridge: MIT Press.
- Bogen, J. (2008): “Causally productive activities”, *Studies in History and Philosophy of Science* 39, 112-23.
- Couch, M.B. (2011): “Mechanisms and constitutive relevance”, *Synthese* 183, 375-88.
- Craver, C. (2007a): *Explaining the Brain*. Oxford: Clarendon Press.
- (2007b): “Constitutive explanatory relevance”, *Journal of Philosophical Research* 32, 3-20.
- (2014): “The Ontic Conception of Scientific Explanation”, in M.I. Kaiser, O. Scholz, D. Plenge, A. Hüttemann (eds.), *Explanation in the Special Sciences – The Case of Biology and History*, Dordrecht: Springer.
- Craver, C., Darden, L. (2013): *In Search of Mechanisms*. Chicago: University of Chicago Press.

- Craver, C., Tabery, J. (2015): "Mechanisms in Science", in Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition, forthcoming URL = <http://plato.stanford.edu/archives/spr2016/entries/science-mechanisms/>).
- Darden, L. (2006): *Reasoning in Biological Discoveries: Essays on Mechanisms, Interfield Relations, and Anomaly Resolution*, Cambridge: Cambridge University Press.
- (2008): "Thinking Again about Biological Mechanisms", *Philosophy of Science* 75, 958-69.
- Fagan, M.B. (2012): "The Joint Account of Mechanistic Explanation", *Philosophy of Science* 79, 448-72.
- Glennan, S. (1996): "Mechanisms and the nature of causation", *Erkenntnis* 44, 49-71.
- (2002): "Rethinking Mechanistic Explanation", *Philosophy of Science* 69, 342-53.
- (2010): "Mechanisms, Causes, and the Layered Model of the World", *Philosophy and Phenomenological Research* 81, 362-381.
- (forthcoming): *The New Mechanical Philosophy*.
- Gillett, C. (2013): "Constitution, and Multiple Constitution, in the Sciences: Using the Neuron to Construct a Starting Framework", *Minds and Machines* 23, 301-37.
- Harbecke, J. (2010): "Mechanistic Constitution in Neurobiological Explanations", *International Studies in the Philosophy of Science* 24, 267-85.
- Harinen, T. (forthcoming): "Mutual manipulability and causal inbetweenness", *Synthese*.
- Haugeland, J. (1998): *Having Thought*, Cambridge, MA: Harvard University Press.
- Huneman, P. (2014): "Individuality as a Theoretical Scheme. II. About the Weak Individuality of Organisms and Ecosystems", *Biological Theory* 9, 374-381.
- Illari, P., Williamson, J. (2012): "What is a mechanism? Thinking about mechanisms *across* the sciences", *European Journal for Philosophy of Science* 2, 119-35.
- (2013): "In Defense of Activities", *Journal for General Philosophy of Science* 44, 69-83.
- Kaiser, M. I. (2015): *Reductive Explanation in the Biological Sciences*, Cham: Springer.
- (forthcoming): "Individuating Part-Whole Relations in the Biological World", in O. Bueno, R.-L. Chen, M. B. Fagan (eds.), *Individuation across Experimental and Theoretical Sciences*, Oxford: Oxford University Press.
- Kaiser, M. I., Krickel, B. (2016): "The Metaphysics of Constitutive Mechanistic Phenomena", *The British Journal for the Philosophy of Science*, DOI: 10.1093/bjps/axv058.
- Kaplan, D.M. (2012): "How to demarcate the boundaries of cognition", *Biology and Philosophy* 27, 545-70.

- Kauffman, S.A. (1971): “Articulation of Parts Explanation in Biology and the Rational Search for Them”, *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1970, 257-72.
- Leuridan, B. (2012): “Three Problems for the Mutual Manipulability Account of Constitutive Relevance in Mechanisms”, *British Journal for the Philosophy of Science* 63, 399-427.
- Lewis, D.K. (1986): *On the Plurality of Worlds*, Oxford: Blackwell.
- Machamer, P. (2004): “Activities and Causation: The Metaphysics and Epistemology of Mechanisms“, *International Studies in the Philosophy of Science* 18, 27-39.
- Machamer, P., Darden, L., Craver, C. F. (2000): “Thinking About Mechanisms“, *Philosophy of Science* 67, 1-25.
- Mackie, J.L. (1965): “Causes and Conditions”, *American Philosophical Quarterly* 24, 245-64.
- Marom, S. (2010): “Neural timescales or lack of thereof”, *Progress in Neurobiology* 90, 16-28.
- Psillos, S. (2004): “A Glimpse of the *Secret Connexion*: Harmonizing Mechanisms with Counterfactuals”, *Perspectives on Science* 12, 288-319.
- Simon, H.A. (1962): “The Architecture of Complexity”, *Proceedings of the American Philosophical Society* 106, 467-82.
- Skipper, R.A., Millstein, R.L. (2005): “Thinking about evolutionary mechanisms: natural selection”, *Studies in History and Philosophy of Biological and Biomedical Sciences* 36, 327-47.
- Tabery, J.G. (2004): “Synthesizing Activities and Interactions in the Concept of a Mechanism”, *Philosophy of Science* 71, 1-15.
- Torres, P.J. (2009): “A Modified Conception of Mechanisms”, *Erkenntnis* 71, 233-51.
- Wimsatt, W.C. (1972): “Complexity and Organization”, *Proceedings of the Philosophy of Science Association* 1972, 67–86.
- (1974): “Reductive Explanation: A Functional Account”, *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1974, 671-710.
- (1994): “The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets”, *Canadian Journal of Philosophy* 24, 207-274.
- (2007): *Re-Engineering Philosophy for Limited Beings. Piecewise Approximations to Reality*. Cambridge: Harvard University Press.
- Woodward, J. (2002): “What Is a Mechanism? A Counterfactual Account”, *Philosophy of Science* 69, 366-77.

- (2003): *Making Things Happen: A Theory of Causal Explanation*, New York: Oxford University Press.
- (2014): “A Functional Account of Causation; or, A Defense of the Legitimacy of Causal Thinking by Reference to the Only Standard That Matters—Usefulness (as Opposed to Metaphysics or Agreement with Intuitive Judgment)”, *Philosophy of Science*, 691–713.

---

<sup>1</sup> I do not use the term ‘constituent’ because some mechanisms are said to constitute their phenomena (Kaiser and Krickel 2016) and the relation between mechanisms and phenomena should be kept apart from that between components and mechanisms.

<sup>2</sup> This claim presupposes endurantism (Lewis 1986).

<sup>3</sup> In the debate, there is no agreement on what exactly a cause is. Some hold that causes are difference-makers (e.g., Craver 2007a), in the sense that if they had not occurred, the purported effect would not have occurred. Others claim that causation is production (e.g., Bogen 2008, Glennan 2010), where causes are actively producing their effects.

<sup>4</sup> Alternatively, one might stick to the claim that activities must involve changes and argue that mechanisms consist of more than just entities and activities – namely property instantiations or maintenances (i.e., states).

<sup>5</sup> Strictly speaking, these numbers are false because material objects and properties are already two ontological categories.

<sup>6</sup> Only Machamer (2004) deviates from the *Parity thesis* insofar as he sometimes suggests the priority of activities.

<sup>7</sup> According to the ontic conception of explanation, the question of explanatory relevance is not an epistemic but also a metaphysical question because explanations are regarded as things in the world (Craver 2014).

<sup>8</sup> Whether Craver’s account of constitutive relevance succeeds in keeping apart constitutive and causal relevance or whether it collapses into an account of causal relevance is, however, contested (e.g., Leuridan 2012, Baumgartner and Gebharder 2016, Harinen forthcoming).