Biology & Philosophy (2021) 36:53
https://doi.org/10.1007/s10539-021-09829-8

Mechanism, autonomy and biological explanation

Leonardo Bich[1] · **William Bechtel**[2]

Received: 4 December 2020 / Accepted: 10 November 2021

Abstract
The new mechanists and the autonomy approach both aim to account for how biological phenomena are explained. One identifies appeals to how components of a mechanism are organized so that their activities produce a phenomenon. The other directs attention towards the whole organism and focuses on how it achieves self-maintenance. This paper discusses challenges each confronts and how each could benefit from collaboration with the other: the new mechanistic framework can gain by taking into account what happens outside individual mechanisms, while the autonomy approach can ground itself in biological research into how the actual components constituting an autonomous system interact and contribute in different ways to realize and maintain the system. To press the case that these two traditions should be constructively integrated we describe how three recent developments in the autonomy tradition together provide a bridge between the two traditions: (1) a framework of work and constraints, (2) a conception of function grounded in the organization of an autonomous system, and (3) a focus on control.

Keywords Biological explanation · Organization · Function · Constraint · Control

Introduction

How should scientists explain biological phenomena? Many philosophers of science, who focus on practices of biologists, have embraced some form of mechanistic explanation, according to which biological phenomena are viewed as the products of mechanisms that behave as they do because of their constitution (Machamer et al. 2000; Bechtel and Abrahamsen 2005; Glennan 2017). Other philosophers of science, drawing upon a lineage of theoretical biologists and with the aim of developing a

✉ Leonardo Bich
leonardo.bich@ehu.es

[1] Department of Philosophy, IAS-Research Centre for Life, Mind and Society, University of the Basque Country (UPV/EHU), Avenida de Tolosa 70, 20018 Donostia-San Sebastian, Spain

[2] Department of Philosophy, University of California San Diego, La Jolla, CA 92093, USA

Published online: 23 November 2021

theory of living and cognitive systems, have emphasized the distinctive organization of biological organisms as autonomous systems. To maintain themselves far from equilibrium with their environments, organisms rely on a closed network of constraints that (1) are constructed by the organism and (2) direct flows of free energy through the organism to perform functional activities within it (Moreno and Mossio 2015). The first group espouses explanation in terms of mechanisms, the second in terms of closed networks of constraints that maintain themselves. Although these two philosophical traditions have largely developed independently of each other, we argue that they should be constructively integrated, as each supplies ingredients the other tradition has ignored or not accounted for in sufficient detail. We focus our discussion on the contributions of these philosophical traditions to biological explanation, and the advantages of integrating them.[1]

The new mechanists in philosophy of science have followed the lead of cell and molecular biologists over the last century in explaining biological phenomena in terms of mechanisms, where mechanisms are understood as organized parts that together perform the activities required to generate the phenomenon. The research in biology has proven extremely productive. The new mechanists have focused in part on the research strategies by which scientists develop these explanations: localizing the phenomenon in a mechanism, decomposing it into constituent parts, determining what activities or operations these parts perform, and then determining how the parts are organized to generate the phenomenon. These activities take the mechanism as the unit, with researchers directing their energies inside it. As we discuss in Sect. "Challenges confronting mechanists on which autonomy theorists can offer assistance", however, the new mechanists have confronted a number of challenges in specifying the phenomena for which mechanisms are to be sought, the boundaries of mechanisms, and how to generalize mechanistic explanations.

In Sect. "Challenges confronting autonomy theorists on which mechanists can offer assistance" we turn to the autonomy tradition. A driving question for this tradition is to explain the distinctive features of biological organisms. Maturana and Varela (1980) characterized living organisms as autopoietic—each constructs itself as a whole from materials it procures from its environment. Moreno and Mossio (2015) have recently emphasized the central importance of procuring and constraining the flow of free energy and have characterized organisms as autonomous in virtue of a mutual dependence between internally produced constraints, giving rise to 'closure of constraints'. The idea of autonomy influenced conceptual debates about life and cognition in theoretical biology, origins of life research, cognitive science and synthetic biology. Yet, as developed within the autonomy tradition, this account is highly abstract and holistic. It leaves uncertain how to ground it in biological

[1] The mechanist and the autonomy traditions are not the only perspectives on explaining biological phenomena. Nicholson and Dupré (2018), for example, have recently advocated that a process perspective is preferable to a mechanistic one. Part of their criticism of mechanists is that mechanists reify mechanisms and their parts as fixed entities. We cannot address this challenge in detail here, but note that the mechanist tradition does emphasize activities and many mechanists view mechanisms as themselves dynamical and transitory structures. More constructively, we welcome insights from process theorists that could further enrich the integration of the new mechanists and autonomy theorists that we advocate here.

research and to provide *causal explanations* of phenomena in the context of the autonomous organism.

In Sect. "Towards bridging the two philosophical traditions" we develop three recent developments in the autonomy tradition that together provide a bridge between the two philosophical traditions: (1) a framework of work and constraints, (2) a conception of function grounded in an autonomous system, and (3) a focus on control. In Sect. "Autonomy considerations provide guidance for mechanistic explanation" we develop how the notions of constraint and function and considerations of control provide insight on the issues which present challenges for the mechanistic approach, while in Sect. "Mechanistic approaches ground explanations of the autonomy of organisms" we show how recharacterizing mechanisms in terms of constraints allows the autonomy tradition to ground its accounts in terms of mechanisms. Bridging the two traditions has the potential to enhance the explanations of biological phenomena they individually provide: It allows the new mechanists to look out from mechanisms to how they are produced and controlled in the organism so that they operate as needed to maintain the organism, and the autonomy tradition to look inward to the actual mechanisms that perform the critical activities, including control activities, needed to maintain the organism.

We conclude in Sect. "Conclusion: a call for constructive engagement between mechanism and autonomy" by calling for constructive engagement between the mechanistic and autonomy traditions, focusing in particular on control as a point of engagement. The production mechanisms that transform materials to build and repair an organism are subject to multiple control mechanisms that alter constraints within them so that they serve the needs of the organism (the notions of constraints and control are explicated in Sect. "Towards bridging the two philosophical traditions"). These control mechanisms are different from those that have been the primary focus of the new mechanists in that they operate on other mechanisms in light of information they procure about the conditions in the organism and the environment. Understanding control is required to understand how mechanisms actually operate in the service of the autonomy of the organism. Thus, a focus on control provides a natural point of productive engagement between the new mechanists and autonomy theorists.

Challenges confronting mechanists on which autonomy theorists can offer assistance

The new mechanists were largely motivated to develop accounts of mechanistic explanation to address perceived inadequacies in the framework of deductive nomological (DN) explanation that had dominated discussions of explanation in the middle decades of the twentieth century. The DN framework had emphasized the importance of scientific laws. Recognizing that biologists seldom refer to laws when advancing explanations, Bechtel and Richardson (1993/2010) focused on the fact that biologists often explain a phenomenon by describing the responsible mechanism, identifying its parts and operations. They and others (Craver and Darden 2013) have characterized numerous heuristics researchers employ to discover mechanisms.

An example will provide the flavor of mechanistic explanations. After the role of DNA as the genetic material was established, researchers turned to the question of how genes give rise to proteins. Several findings pointed to the relevant components: the concentration of RNA in the newly discovered cell organelle, the endoplasmic reticulum, of structures, called ribosomes, on the rough endoplasmic reticulum, and of different forms of RNA. The different RNAs were soon linked to different functions: ribosomal RNA constituted the ribosome, messenger RNA coded for the order of amino acids in a protein, and transfer RNA ferried specific amino acids to the locus on the messenger RNA that was then in the reading frame of the ribosome. Researchers established that these different activities or operations are organized so that together they generate polypeptide chains that then folded, either on their own or with the assistance of other proteins, into the form of the mature protein. [For a detailed analysis of the steps in the discovery of these mechanisms see Darden and Craver (2002)].

As informative as the new mechanists have been about the character of mechanistic explanation and the strategies scientists use to reason from data to mechanistic explanations, there are critical questions about how to pursue scientific inquiries that lead to mechanistic explanations on which new mechanists have yet to provide definitive answers. In their practices scientists themselves embrace answers to these questions and one might hope that accounts of mechanistic explanation could illuminate these practices.

The first involves the very identification of mechanisms to investigate. The standard view among mechanists is that mechanisms are identified in terms of the phenomenon they explain. However, without an account of what to count as a phenomenon and of how to pick out the relevant phenomenon among many, appealing to phenomena does not solve the problem. Mechanists have said little about how phenomena are to be identified, largely simply embracing Bogen and Woodward's (1988) understanding of phenomena as regularities or repeatable occurrences. This is not sufficient since there are a vast number of regularities or repeatable events occurring within organisms and only some of them are picked out as worth explaining. Which one a researcher selects matters. As argued by Kauffman (1970), depending on what one takes the phenomenon to be, one will develop different decompositions of a system into mechanisms and of mechanisms into their component parts. In part, the selection a scientist makes is constrained by previous choices about what to explain: new phenomena are often identified in the context of explaining other, often-larger scale mechanisms. Craver (2007) refers to the "mosaic unity of neuroscience" to characterize how different investigations, and the phenomena each investigates, are related to one another. This, however, pushes the challenge back to how the encompassing phenomena are selected.

Biologists have no problem making selections of phenomena that they find worth explaining. In their practice, they select phenomena that they take to be important in the lives of organisms (accordingly, they focus on circulating blood, not the generation of heart sounds). When Galvani (1791) found that connecting the interior of a muscle to the nerve by a conductive material generated an electrical current, he assumed it to be important to the ability of organisms to contract their muscles. Much later, when Spudich observed motility when myosin was added to actin

filaments in an in vitro preparation (Spudich et al. 1985), he assumed that myosin contributed to the motility of living cells. Accounts of mechanistic explanation offer little insight into how assumptions such as these guide scientists' selection of phenomena to investigate.

Disease presents a context in which how phenomena are delineated is consequential. Mechanists have pursued two strategies to extend mechanistic accounts to disease. One views diseases as themselves phenomena (Darden 2013). Accordingly, one finds discussion of disease mechanisms—mechanisms that generate the disease (for an application of this framework to Krohn's disease, see Darden et al. 2018). The other views diseases as resulting when mechanisms that produce phenomena in healthy organisms are broken or working aberrantly (Garson 2013). This view requires not only an account of how to justify the claim that a mechanism is not working correctly but also a basis for assessing the breakdown as a disease (not every disrupted process in a living organism is regarded as constituting a disease).[2] Each approach is embraced by some medical researchers and has given rise to valuable knowledge. Adopting a pluralistic perspective on how to apply the mechanistic framework to disease may well be appropriate. Both approaches, though, require selection: which conditions to count as a disease or which mechanisms, when broken, result in disease. The mechanistic perspective itself does not answer this question.

Beyond selecting among the regularities in the world which to count as phenomena to be explained, there is a further challenge in identifying the responsible mechanism: determining which entities constitute the mechanism and which are outside it. While humans make machines with reasonably clear boundaries (there is little ambiguity about the boundaries between a car or a coffee maker and its environment), it is much more difficult to identify the boundaries of mechanisms within biological organisms. What one finds in organisms is a vast network of causally interacting entities. These interactions extend out into the environment. The challenge is to identify the entities that constitute a mechanism and distinguish them from others that are external to the mechanism but interact with it. Craver (2007) advances a constitutive relevance criterion: the mechanism for a phenomenon consists of all the causal processes that causally affect the phenomenon.[3] His goal is to provide

[2] Biological systems exhibit a high degree of robustness. Redundancy and degeneracy of functions may obviate the presence of a broken mechanism, for example when multiple copies of a gene are available so that even in case of one harmful mutation, the organism would still be able to produce functioning proteins. In other cases, the organism might employ ways to block the effects of broken mechanisms by eliminating the final product, such as misfolded proteins or malfunctioning cells. Moreover, effects of a broken mechanism may emerge as harmful only when the organism or the environment are in a particular state, for example during cell division, in presence of different food sources, predators, etc. These are all cases in which the living system needs to change its physiological regime or its behavior in the environment. Therefore, in many cases understanding disease may require looking outside the mechanism to focus on the organism as a whole.

[3] Craver as well as many other mechanists embrace Woodward's (2003) interventionist treatment of causation according to which any entity on which one can intervene to alter the effect counts as a cause. Craver fleshes this criterion for mechanism membership in terms of mutual manipulability. Some have objected that this violates Craver's own view that the constitution relation between components and mechanisms is not causal. Krickel (2018) provides a metaphysical scheme in which mutual manipula-

an objective criterion. As Kaiser (2017, p. 127) characterizes his position: "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."

It is not clear that fixing the phenomenon is sufficient to demarcate the boundaries of mechanisms in the manner that scientists commonly do. Bechtel (2015a) describes how over time research on mechanisms often leads to discovering an increasing range, both spatially and temporally, of entities and activities that make a difference to the phenomenon. In the case of circadian rhythms in organisms ranging from fungi and plants to insects and rodents, research efforts in the 1990s identified on the order of ten proteins in each that could be characterized as constituting the mechanism (for a review, see Hastings and Herzog 2004). But taking advantage of new techniques that allowed for mass screening, Zhang and Kay (2010) identified over a thousand genes that affect the period or amplitude of circadian rhythms in mice. As these are two central features of the phenomenon of circadian rhythmicity, the proteins coded for by these genes might all be considered components of the circadian clock mechanism. Bechtel also invokes research by Marom (2010) that the action potential, often appealed to as an exemplary phenomenon by Craver, lacks start and termination conditions. Marom's studies show that the effects of a single action potential on the neuron in which it occurred do not sharply terminate but drop off according to a power law, leading him to characterize the phenomenon as scale-free.

These findings, however, have not led researchers to vastly expand the spatial and temporal boundaries of the mechanisms they investigate. Bechtel's contention is that researchers determine the boundaries of mechanisms based on their epistemic objectives. Noting that this challenges Craver's strategy for rendering the boundaries of mechanisms an objective matter, Kaiser (2017) concludes that this turns "the mechanistic account into a purely epistemic account that has been detached from almost all ontological commitments." Glennan (2017) also adopts an epistemic perspective on the challenges in carving the world into mechanisms and mechanisms into their parts, but is more sanguine than Bechtel about the possibility of identifying objective boundaries once the phenomenon is chosen. He appeals, for example, to Simon's characterization of the distinctive activities that occur at the interface between mechanisms to identify boundaries. For our purposes, it is not necessary to take a stand on what considerations go into fixing the boundaries of mechanisms, but only note that it presents a challenge, of which mechanists are aware.

A further challenge confronting new mechanists stems from the fact that research on mechanisms is usually devoted to specific instances of the mechanism. Mechanistic explanations, however, are nearly always presented in general terms (exceptions, as discussed by Glennan (2017), are singular events). The scope may vary: an explanation might explain quorum sensing in a specific species (*E. coli*) or even a specific

Footnote 3 (continued)
tion can serve to identify the components of a mechanism without violating the stricture that the relation between components and mechanisms is not causal.

strain, or in gram negative bacteria, or in all prokaryotes. This raises the question of how mechanistic explanations are able to generalize over instances. Part of the answer is that the characterization of entities and activities must employ general terms. But how general? We focus in turn on two related issues: the basis for extending explanations from the instance investigated to target cases, which may involve organisms of other species, and when and how to abstract from details of individual mechanisms. Both of these are foci of current discussion by the new mechanists.

The challenge of generalization is evident in biologists' widespread practice of investigating a mechanism in one species and applying the results to a different target species. Researchers investigate sleep in mice, fruit flies, and worms even though they are really interested in sleep in humans (Joiner 2016). In recent philosophy of science there is growing interest in the epistemic issues involved in selecting appropriate model organisms and drawing conclusions from them (Ankeny and Leonelli 2011, 2020; Levy and Currie 2014; Hardesty 2018). For our purposes, what is of particular importance is that there are always important differences between the phenomenon and the mechanisms in model organisms and the target species. Continuing with the phenomenon of sleep, mice sleep during daylight and worms only during developmental stages. Yet the goal of this research is to draw conclusions about daily nocturnal sleep in humans. Philosophers such as LaFollette and Shanks (1996) and Weber (2005) have argued that researchers should not extrapolate from model organisms to target species but only use the models as sources of hypotheses that then must be investigated separately in the target organism. This makes sense in cases in which one lacks evidence for a descent relation that supports the hypothesis that the mechanisms in the two cases are conserved from a common ancestor. When there is evidence of conservation from a common ancestor, scientists treat the research on models as providing evidence for the mechanism in the target. Since evolution is a process of descent *with modification* even in the case in which the mechanisms are judged to be conserved, critical differences can arise. Steel (2007) proposes a strategy that does not require amassing evidence for the mechanism solely from the target species: if one traces those processes that are most likely to manifest differences and fails to find any, he argues, one has grounds for extending the findings from one to the other. Finding differences between the processes in the two mechanisms is often grounds for scientists to reject an application from a model organism to a target, but not all differences matter equally. This raises the question: which differences matter?

An important requirement in order to extrapolate information about mechanisms from one species to another is to abstract from the details that distinguish the species. The requirement for abstraction is broader than just cross-species research—it is required any time the same explanation is applied to multiple cases. Abstraction is commonly characterized as leaving out details, but which details can be left out? Details that distinguish the cases must be dropped. But are these the only details that should be deleted? And what is the basis for deciding which details to drop? There is general agreement among mechanists as to the importance of abstraction but not the basis for doing so. In setting out the norm for explanation, Craver and Kaplan (2020) appeal to the specification of the phenomenon as providing a criterion of relevance—one can, and should, leave out any details that are not relevant to

the phenomenon. On the other hand, all relevant details belong in an ideally complete mechanistic explanation.[4] They acknowledge that explanatory accounts scientists actually produce will fall short of this ideal, in part because not all the relevant details may be known.

Craver and Kaplan situate explanations (as opposed to explanatory models or accounts) in the world. Other philosophers follow Glennan (2017) in focusing primarily on explanatory models and address questions about what details to include on the basis of the epistemic objectives of constructing the model. Boone and Piccinini (2016), for example, describe how explanations may abstract in the service of mathematical and computational tractability or in characterizing processes at a given level of organization (leaving out details about how components in a mechanism are realized). Levy and Bechtel (2013) describe how abstracting from features of the components of the mechanism can enable researchers to show how particular patterns of organization can generate a phenomenon independently of the properties of the entities so organized. Hochstein (2016) introduces the idea that different models of a mechanism, by abstracting in different ways, can give incompatible but useful models of the same mechanism. Much of the emphasis in mechanists' discussion of abstraction focus on ways in which mechanistic accounts idealize, but except for Craver and Kaplan's injunction to leave out irrelevant details, most have not advanced norms governing practices of abstraction.

In this section we have identified a number of challenges new mechanists face in developing their accounts of explanation: challenges in selecting phenomena, determining what belongs to a mechanism, and offering appropriately general explanations. As we have noted, new mechanists have been addressing these issues. That they have not resolved them is not an indictment of mechanistic accounts. Our objective is to show, as we do in Sect. "Autonomy considerations provide guidance for mechanistic explanation", these are issues on which the autonomy tradition can provide valuable insight. But first, we consider the challenges confronting the autonomy tradition that result from its historically eschewing mechanistic explanations.

[4] A number of philosophers have challenged what they took to be an upshot of Craver's characterization of the constitution of the mechanism that explanations that include more details about the mechanism are to be preferred; in mounting her criticism Chirimuuta (2014) characterizes Craver (as well as Kaplan, 2011) as embracing the norm "the more details the better." (For other criticisms of this norm, see Levy 2013; Batterman & Rice, 2014). Recently, Craver and Kaplan (2020) have responded to these criticisms, arguing that they have been misunderstood and have advanced an alternative according to which explanations are to be preferred which provide more details relevant to a contrast between phenomena. Appealing to contrasting explananda has been invoked in discussions of causal explanation to select, among multiple factors that contribute to an event, those that make a difference between different outcomes. However, as emphasized by Glennan (2017), mechanistic explanations are directed at how-questions, not why-questions. Attempts to answers questions about how some phenomenon happened are not typically directed at contrasts but at such questions as: How is the action potential generated? How do muscles contract? How are circadian rhythms produced? Without a contrast, it is hard to see how Craver and Kaplan can delimit which among the factors that make a difference to the phenomenon are to be included in the mechanism (for discussion, see Levy and Bechtel, forthcoming).

Challenges confronting autonomy theorists on which mechanists can offer assistance

The biological autonomy approach focuses on the distinctive capability of living systems (prototypically cells and multi-cellular organisms) to produce their own components, respond actively to perturbations and maintain themselves without being completely driven by external factors. To quote Kauffman (2000), living systems are capable of "acting on their own behalf." Biological autonomy, proponents of the autonomy tradition maintain, is an organizational property realized in all living organisms. Moreover, it is a property of the whole living system. It cannot be referred to any specific component of it; rather, it rests on the peculiar and distinctive way the components—and the processes they are involved in—are related.

Inspired by the work of Claude Bernard (1865), by Cybernetics, and by Systems Theory, several theorists made pioneering theoretical contributions to the characterization of biological autonomy in the late 1960s and early 1970s: Piaget (1967), Rosen (1972), Maturana and Varela (Varela et al. 1974) and Gánti (1975). They embraced the view that organisms are far-from-equilibrium systems organized in such a way that they achieve self-production, self-maintenance, and reproduction through exchanges of matter and energy with their environment. Piaget, Rosen, and Maturana and Varela captured this with the concept of *organisational closure* according to which biological self-maintaining systems exhibit a circular topology: a network of processes of production in which each component is produced by others in the network such that the network maintains itself.

The early autonomy tradition provided a valuable characterization of the distinctive features of biological systems. The proponents emphasized the mutual dependence between an autonomous system and its components, according to which the components are produced, maintained and repaired by the system and contribute to its realization and maintenance. However, these accounts were extremely abstract. One reason for this is found in the heritage of Cybernetics and its emphasis on formal characterizations of organisms. Such formal analysis allows for multiple realizability: different types of biological, physical and mechanical components could realize the same abstract relations, such as feedback loops. A further reason is that advocates of the autonomy tradition sought to distinguish living organisms from other examples of far-from-equilibrium systems that were being advanced at the time in the physical sciences, such as dissipative structures (Glansdorff and Prigogine 1971). While accepting that living systems are maintained in far-from-equilibrium conditions, and are themselves dissipative, the autonomy framework of the 1960's and 1970's emphasized that the distinctive character of biological systems was to be found in their organization rather than in the physical properties of their components.

The autonomy tradition has focused on the fact that organisms are capable of producing their own components and maintaining themselves.[5] Organizational closure was meant as a general *explanans* from which a number of important implications about living systems could be developed, including an understanding of biological phenomenology from an internalist perspective (Maturana and Varela 1980), a neutralist account of evolution (Maturana and Varela 1987; Maturana-Romesin and Mpodozis 2000), embodied accounts of knowledge, cognition, and agency (Piaget 1967; Varela et al. 1991), and contributions to the debate on living systems and machines (Rosen 1991). However, apart from limited and still abstract attempts (Rosen 1972; Ganti 1975), the autonomous organization of the organism was not itself an *explanandum*; for that, causal relations and details about how they are realized have to be fleshed out. While this approach, especially in its early formulations, has been focusing on providing general theoretical foundations for notions such as life and cognition, it has put considerably less effort into developing the contents of the theory by providing causal explanations of actual biological and cognitive phenomena and how they specifically depend on, and contribute to autonomy.

This attitude of favoring theoretical considerations over causal explanations is not limited to the early autonomy tradition. Mossio and Pontarotti's (2019) general account of heredity through appeal to organizational closure is a recent example. The authors make explicit that their focus is on the theoretical foundations of heredity as a means of maintaining a far-from-equilibrium state in descendent organisms, not on providing an explanation of it: "Heredity, as discussed here, is a general concept designating a kind of stability. As such, it does not explain how functions are conserved across generations, which would require describing specific mechanisms and processes for each (class of) hereditary object" (p. 27). The abstract characterization of organizational closure in maintaining living systems in a far-from-equilibrium state during their lifetime and across generations is the reason why the focus on closure has found a fertile terrain in theoretical biology, Artificial Life, Astrobiology, synthetic biology and in the conceptual debates on the nature and origin of life but not, until recently, in experimental biology.

As a result of its heritage from Cybernetics, one does find references to mechanisms in the autonomy tradition. Maturana and Varela (1980) explicitly refer to living systems as autopoietic machines, and to biological explanation requiring a mechanistic approach. However, their understanding of mechanism is vastly different than that espoused by the new mechanists. As they make explicit in a paper entitled "Mechanisms and Biological Explanation," Varela and Maturana (1972) characterize both human-made machines and biological mechanisms in terms of abstract relations, not their material realization, which for them can vary without changing the identity of the machine or mechanism. Their conception of what a mechanism is

[5] We use "maintenance" as a broad term to include procuring resources, using them to construct and repair the organism, and to reproduce. In general, the autonomy tradition has placed less emphasis on reproduction than maintenance of a single organism, but reproduction simply maintains a far-from-equilibrium organization across generations (Saborido, Mossio, & Moreno, 2011).

centers on organization and activities (interactions); components are characterized functionally in terms of their activities within a given organization.

It is important to emphasize that when they apply this view to biology, Varela and Maturana conceive of the organism holistically as one mechanism, rather than a system made of distinct mechanisms. Given that the organism is characterized as a self-maintaining system whose activity and existence coincide, it follows that in a living system the phenomenon produced by the mechanism is the organism itself.[6] Moreover, in characterizing organizational closure, and the circularity it implies, Varela et al. (1974) use holistic terms, appealing to the unitary character of the organization of production processes that maintain that very network, with little or no reference to the different contribution of components and how their relationships are understood. According to autopoietic theory it is the unitary organization that the organism maintains as constant (Maturana and Varela 1980). In this view, an autopoietic system is not decomposable in the way new mechanists characterize mechanisms. This is further illustrated when Varela describes mathematically the circularity and self-reference that define an autonomous organization as an operator exhibiting a fixed point in a series of transformations, or as a self-referential equation f(f) = f (Varela 1979). Although Maturana and Varela mention a possible distinction between three types of *relations* in an autonomous organization—relations of constitution (membrane), specification (nucleic acids) and order (catalysts)—and stress the importance of a physical boundary produced from within,[7] they do not investigate how these features are combined and how they contribute to the organization of the system. Instead, they characterize an autonomous system as a unitary and non-decomposable network, whose general properties and evolution in time can be object of dynamical descriptions and computer simulations (Varela 1979). In insisting that autopoiesis depends on the whole organization, they seek to avoid a reductive focus on the intrinsic properties of specific components (especially the genome as a program). Moreover, they reject a role for the notions of function, regulation and control as describing how different types of components contribute to the physiology of the system. These notions, they claim, reflect the point of view of an external observer.

Rosen (1972) provides a somewhat more detailed mathematical characterization of organizational closure, and how different sets of functional components are

[6] Maturana and Varela (1987) go beyond merely advancing abstract descriptions of mechanisms to advance an explanatory account. On the surface, they seem to share the distinction, advanced by mechanists, between phenomenal and mechanistic models (Glennan, 2017). They argue that in order to explain, rather than describe a biological or cognitive phenomenon one needs to propose a generative mechanism for it: "a mechanism or process that if allowed to operate gives rise, as a result of its operation, to the experience or phenomena to be explained" (Maturana-Romesin and Mpodozis 2000). But there are important differences: what they propose is a generative account that appeals to identifying the mechanism responsible for the phenomenon but not to the details of the mechanism itself. Moreover, this mechanism coincides with the whole system that generates the biological or cognitive phenomenon, and their emphasis is on the distinctive overall organization of this system.

[7] It is interesting that the most explicit biological aspect discussed by Maturana and Varela is also the one that generated most of the experimental work derived from the theory of autopoiesis, specifically in the areas of origins of life and synthetic biology focused on prebiotic compartments (Luisi 2006).

required in order to achieve self-production and self-maintenance, but his account remains abstract and he does not connect mathematical functions to biological components. Rosen explicitly rejects the idea of appealing to mechanisms in explaining biological phenomena. He seeks to emphasize the distinctive character of biological systems and sees invocation of mechanisms as obscuring the difference in organization between biological organisms and artifacts and ordinary physical objects.[8] Nevertheless, there is a sense in which Rosen's models are compatible with mechanistic accounts since his aim was to account for biological phenomena in terms of components and their operating on processes.[9] Gánti (1975) is more concrete in advancing a hypothetical model, the Chemoton, of a minimal living system that is based on chemistry. In particular, he proposes a coupling of three chemical cycles—a metabolic cycle, a template cycle, and a system boundary. Part of his account is a proposal for how Chemotons can reproduce, maintaining the minimal living system. Still, this model is very abstract and far from real chemistry. Nonetheless, Gánti (2003) stresses the importance of accounting for how such organization, in which components continuously undergo transformations, can be realized by actual chemical components. Despite supplying more causal detail than Varela and Maturana on the distinctive biological features of autonomous organizations, both Rosen and Gánti avoid addressing real biological cases in discussing autonomous systems. As a result, they, as well as many others in the autonomy tradition, are not able to demonstrate the applicability of the framework to actual biological organisms.

Towards bridging the two philosophical traditions

As characterized so far, the mechanist and autonomy traditions direct their attention differently—mechanists inside mechanisms, proponents of autonomy towards the whole organism. This results in the new mechanists downplaying what happens outside individual mechanisms and the autonomy tradition downplaying how the actual components constituting an autonomous system causally interact and contribute in different ways to realize and maintain the system. As a result, both approaches face several challenges in developing explanations of biological phenomena. However, there have been developments in the autonomy tradition that point to a means of bridging between the two philosophical traditions. In this section we focus on three: (1) the introduction of the framework of work and constraints, allowing for the recharacterization of closure in terms of constraints, (2) the formulation of an account of function and (3) a focus on control processes. As we will see

[8] It is important to point out that the specific objects of Rosen's criticisms are Cartesian and Newtonian mechanisms and the application of machine models to biology (based on the separation between hardware and software). According to Rosen, both are unable to capture the distinctive capability of biological systems to realize self-producing and self-maintaining closed causal loops (Rosen, 1991).

[9] This sense is close to the contemporary meaning of a mechanism as a causal mechanism, which provides "a step-by-step explanation of the mode of operation of a causal process that gives rise to a phenomenon of interest," (as discussed by Nicholson, 2012, p. 153).

in subsequent sections, these provide the means for each philosophical tradition to draw upon the other.

First, Pattee (1972) invoked the notion of constraint, drawn from classical mechanics, in explaining the activities of organisms. A constraint is a structure that affects a process yet is locally unaffected by the process itself (i.e., it remains the same at the end of the process or returns to the same state; see Montevil and Mossio 2015). A paradigmatic example in biology is an enzyme which catalyzes a chemical reaction without being consumed. The term *constraint* implies limiting, but, as argued by Hooker (2013), constraints also create possibilities—a skeleton restricts the movements an organism can make, but also enables it to move in ways it couldn't otherwise. One thing that is constrained in organisms is the diffusion of Gibb's free energy, resulting, as characterized by Kauffman (2000), in the performance of work. Gibb's free energy (hereafter, free energy) is a measure, developed in thermodynamics, of available energy due to disequilibrium within a system that can be used to produce mechanical work—produce physical changes in components of the system.[10]

Although constraints are not changed during the interval during which they enable work to be done, they are themselves constructed over a longer time frame. As Kauffman also develops, it takes work, and hence free energy, to construct a constraint. The interplay of constraints and work gives rise to what Kauffman characterizes as work-constraint cycles: constrains are needed to perform work, but work is required to establish constraints (Kauffman 2019). This applies as well to constraints that enable organisms to perform work: they must be constructed. This idea provides a means to reformulate Rosen's (1972) formal model of closure in more concrete terms as the closure of constraints (Moreno and Mossio 2015; Montevil and Mossio 2015): all the constraints operative in an organism must be constructed from other constraints. Some of these might be inherited from those constructed in a parent organism, but most are constructed by the organism itself during its lifespan. An organism is autonomous because it creates the set of constraints responsible for its own constitutive activities that maintain its existence. Accordingly, a cycle involving constraints is distinguished from other instances of cyclic organization such as the water cycle in which material is recycled. In the closure of constraints, it is the organization of constraints that recycles itself through its generative activities.

Reframing the notion of autonomy in terms of constraints allows the autonomy tradition to incorporate consideration of the material realization of autonomous systems (Ruiz-Mirazo and Moreno 2004). It also provides a direct bridge to the new mechanists: mechanisms and their parts are constraints that direct flows of free energy into the performance of the activity associated with the mechanism (Winning and Bechtel 2018). Autonomy itself can be characterized in terms of mechanisms: through its mechanisms, the organism constructs and maintains other mechanisms, including those of its descendants. By linking constraints to mechanisms, proponents

[10] An information theoretic notion of free energy has been invoked in neuroscience and cognitive science by Friston (2009) as a unifying principle. While drawing upon the same roots in thermodynamics, there are important differences between his use of the concept and ours.

of autonomy can now fill in the missing biological causal grounding of that tradition by identifying the mechanisms responsible for autonomy. We develop this in Sect. "Mechanistic approaches ground explanations of the autonomy of organisms".

Second, theorists in the autonomy tradition have recently developed a new account of biological function. Mechanists often speak of mechanisms and their parts as performing functions, but the accounts of function traditionally on offer are not sufficiently focused to characterize functions of biological mechanisms. Accounts that identify the function of a component with its causal role in a larger system (Cummins 1975) do not provide a normative basis for distinguishing which among many causal effects to count as the function of a component. The etiological approach, which appeals to evolutionary history to identify effects that led the organism to be selected as functions (Millikan 1989), faces a number of problems, including that previous adaptations may no longer be functional for the organism. Mossio et al. (2009) draw upon the autonomy tradition to offer a more fine-grained characterization of the function of a constraint in terms of its role within a closed system of constraints within an existing organism.[11] Drawing upon the above linking of mechanisms with constraints, one can characterize the functions of mechanisms in terms of their contributions to the maintenance of particular biological organisms.[12] This allows one to explain the existence and operation of a mechanism by referring to its function within the organism.[13] We consider how this provides insight addressing some of the challenges mechanists confront in the next section.

Third, the autonomy tradition has expanded its focus on what organismic activities support autonomy. Early autonomy theorists focused on how the parts of a system are produced, transformed and repaired from within the system. However, to remain viable, especially in changing environments, autonomous systems must also

[11] This account of functions was inspired by the work of McLaughlin (2001) and Christensen and Bickhard (2002), among others, which grounds functional attributions in the contributions of biological traits to the maintenance of the system that harbors them. Within this framework functional roles are usually attributed to physiological traits involved in the maintenance of the organism. However, a functional grounding has been provided for cross-generational functions such as reproduction and for altruistic behaviors, which do not directly contribute to the maintenance of the organism. In this view, reproductive traits are functional because they are produced by the autonomous biological organization of the parents at some point in their life cycle, and they contribute to re-establishing that very organization in the offspring (Saborido et al., 2011; Mossio and Pontarotti, 2019). For a criticism and response see, respectively, Artiga and Martinez (2016) and Mossio and Saborido (2016). For a functional characterization of altruistic behaviors, see Frick et al. (2019).

[12] In embracing Mossio et al.'s account of function as appropriate in understanding functions in biological organisms, we are not arguing that it provides a universal account for all function attributions. Wimsatt (1972) advanced a general theoretical account that covers a broad range of ways functions are attributed to entities. A key variable in Wimsatt's account is purpose. Maintaining themselves and replicating are critical purposes for living organisms as far-from-equilibrium systems. If they fail to satisfy these purposes, organisms cease to exist. Wimsatt's framework, though, makes clear that functions can be attributed in other contexts for systems with different purposes.

[13] Scientists may identify functions in other ways, for example by adopting an etiological evolutionary account (Millikan 1989; Neander 1991). Yet, in that case one should expect that when scientists identify a function, they do so on the basis of evolutionary evidence. However, this is not what typically happens in scientific practice. When biologists characterize a function, they may do it on the basis of things happening at the present time within an organism, and only afterwards attempt to reconstruct its evolutionary past (see Frick et al. 2019 for a discussion of this issue in studies on biological communication).

respond adaptively, changing what activities they perform in ways appropriate to the circumstances they face. Only through such adaptive response can organisms counteract potentially destabilizing interactions with the environment and amongst internal components. Di Paolo (2005) emphasized the importance of developing the autonomy framework in this direction. To do so, some autonomy theorists have introduced a different type of constraint—regulatory constraints—into the account of closure (Bich et al. 2016). Among the new mechanists, Winning and Bechtel (2018) have characterized a type of mechanism that they term a *control mechanism* in similar terms. Regulatory constraints or control mechanisms are sensitive to environmental or internal conditions requiring adaptation and act upon other constraints or mechanisms to selectively shift between different available regimes of self-maintenance in such a way as to contribute to the viability of the system.[14] A classic example of this type of shift is the one occurring in bacteria between distinct metabolic regimes based on the availability of different sources of sugar (lactose and glucose). We also explore how the distinction between control and controlled mechanisms addresses some of the challenges facing the new mechanists in the next section.

Autonomy considerations provide guidance for mechanistic explanation

As we have seen, to explain a phenomenon the new mechanists identify it with a mechanism and then look inside it to explain how it produces the phenomenon in terms of its organized parts and operations. Besides identifying inputs or start-up conditions and how they arise, the only motivation for looking outside the mechanism is if researchers determine that the boundaries of the mechanism were set too narrowly—that entities outside those boundaries causally affect the phenomenon. In contrast, the autonomy tradition advocates looking to the organism and how individual mechanisms fit into the closure associated with the organism—how the mechanism was built and maintained by other components of the organism and how it in turn contributes to building and maintaining other components. New mechanists can analyze how the phenomenon associated with one mechanism may figure in a larger-scale mechanism. However, by identifying larger and more comprehensive mechanisms that carry out higher-level activities, one will not arrive at an account of the organism as an autonomous system. Rather, one will only arrive at individual higher-level phenomena exhibited by autonomous systems, such as moving or feeding. From this perspective, autonomous systems themselves can be construed, to use Glennan's terminology (2017), as 'mechanistic systems', that is, as systems of

[14] The ability of control mechanisms (or regulatory constraints) to be sensitive to conditions that affect the maintenance of the autonomous system depends on making measurements. These alter constraints in the control mechanism and so determine what work a control mechanism performs on the constraints of the controlled mechanism.

mechanisms[15] that are not necessarily mechanisms themselves. The two approaches do not coincide but are not incompatible and can engage fruitfully with each other. In the previous section we showed how the autonomy tradition has recently offered an account of biological function and has come to emphasize regulation or control. In this section we explore how these foci of the autonomy tradition provide resources to help address the challenges confronting new mechanists that we identified in Sect. "Challenges confronting mechanists on which autonomy theorists can offer assistance".

First, the autonomy tradition's account of function provides a perspective from which to select which among activities occurring in a living organism are phenomena to be explained. Specifically, it is important to explain those activities that count as functions in virtue of their contribution to the maintenance, repair, and reproduction of the current organism. As argued by Garson (2013) among others (see also Glennan 2017), adopting a functional approach is not only compatible with mechanistic accounts but also widely embraced and recognized as useful. By incorporating a focus on organisms as maintaining autonomy, the mechanist framework acquires a basis for picking out phenomena that is in line with scientific practice. Activities that do not affect the autonomy of the organism are not regarded as important as those that contribute to it. Such considerations seem to figure in scientists' judgments about what to investigate.

As articulated by Glennan (2017), functional considerations provide a normative perspective on mechanisms.[16] Norms enable both characterizing proper functioning but also what counts as impaired functioning. The autonomy tradition provides a perspective for articulating a mechanistic perspective on disease—what diseases do is impair the ability of the organism to maintain autonomy. Some diseases are brought about by external agents such as bacteria that are acting to promote their autonomy (whether viruses count as autonomous systems is more controversial). But for those organisms afflicted with diseases, it is the challenges to their ability to maintain itself that are most relevant (Saborido and Moreno 2015). In some cases, disease involves the inability of a production mechanism to perform its function (e.g., for the kidney to filter and remove waste from the blood). But in many others, it involves failure of control mechanisms to regulate these mechanisms in the service of the organism. Cancers (Bechtel 2018) and psychopathologies (Bechtel 2020) provide illuminating examples.[17] While it is challenging to characterize what

[15] A mechanistic system is a system in which mechanisms act. It can also participate in larger mechanisms: an example is the contribution of yeast to the phenomenon of fermentation in a bioreactor.

[16] Recently, working within the mechanist tradition, Winning (2020) has developed a related but distinctively different account of function in terms of the perspective of a biological organism on its proper functioning. It too advances a normative perspective, one grounded in the activities through which an organism enforces its own standards.

[17] A consideration of control mechanisms provides a useful complement to Darden's (2013) account of mechanisms that produce disease symptoms. In many cases, the mechanism that results in disease symptoms comprises components that figure in healthy functioning but are operating in an altered manner. Often is it altered control processes that result in the components operating differently. Darden's account identifies the start or setup conditions and the productive continuity between components in generating the disease state, while the control perspective can provide insight into how these components came to operate differently than they do in a healthy organism.

counts as disease using just the resources of the new mechanists, by drawing upon the autonomy perspective's construal of function, the mechanist tradition acquires a perspective for assessing which mechanistic processes, or disruptions in them, constitute disease.

Second, the autonomy tradition can provide insight into how to fix the boundaries of mechanisms. As we discussed in Sect. "Challenges confronting mechanists on which autonomy theorists can offer assistance", theorists such as Glennan and Bechtel appeal to epistemic objectives of scientists. To others, such as Craver and Kaiser, this results in an objectionably subjective account of mechanisms. Just as considerations of contributions to autonomy can be used to delineate phenomena, they can be used to settle where the boundaries of a mechanism should be drawn. When Zhang and Kay (2010) identified over a thousand additional proteins that affected the period and amplitude of circadian oscillations, in presenting their analysis, they nonetheless distinguished core clock genes from the extended set of proteins that, when altered, alter circadian rhythms. This is not an arbitrary choice but principled. What supports this distinction? In its account of adaptive responses, the autonomy tradition provides insight. It makes a distinction between constraints that direct flows of energy within a system and control processes or mechanisms that operate on these constraints to alter their function. Although circadian biologists generally do not make special note of the free energy deployed in maintaining circadian rhythms, nonetheless the proteins constituting the core clock constitute an organized systems in which free energy is constrained to produce an oscillation in the expression of proteins, including those that regulate other activities. The other proteins that can affect circadian rhythms all do so by modulating the activity of these proteins. They are performing a control activity by altering the flexible constraints in the mechanism generating circadian oscillations. Thus, one can distinguish components that are core to a given mechanism and those that act on it to alter its constraints and hence its operation.

The distinction between control activities and the activities within the controlled mechanism applies far beyond this one case. To distinguish control processes from those within controlled mechanisms, one needs to focus on the work that a mechanism is doing. Some mechanisms in living organisms perform work that directly serves to maintain the organism—procuring and digesting food, synthesizing body parts, locomoting, etc. These are the ones we have called production mechanisms. But most components of organisms perform work on the constraints within other mechanisms, often other control mechanisms. Indeed, the circadian clock is itself a control mechanism, regulating other mechanisms. The other mechanisms affecting its flexible constraints are regulating it. Tracking the work that is performed, and when that work is being performed on other mechanisms, provides a principled way of drawing boundaries around mechanisms.

Third, the autonomy perspective provides resources to address challenges in addressing the generalizability of mechanistic explanations. One challenge is to explain when investigations of mechanisms in model organisms, often of different species, can be extrapolated to a target organism. Evidence that the model and the target are related phylogenetically in such a manner that they are conserved versions

of the same ancestorial mechanism is viewed as supporting extrapolation from the model to the target. However, since evolution also results in modifications, there will be differences. The question is when differences undercut the inference. The autonomy tradition offers an important consideration: are the mechanism in the model and the target serving the maintenance of the organism in the same way? If they are not, the differences may well be consequential. To the degree that they contribute to the maintenance of the organism in the same manner, the case for extrapolating results from one to the other is stronger. In the case of sleep mechanisms, the evidence is strong that sleep in fruit flies serves the same function—flies and humans exhibit similar responses when deprived of daily sleep, including suffering memory impairments and eventually dying (Shaw et al. 2002). The fact that worms only enter a sleep-like state during development makes the challenge of demonstrating common function more difficult. In this case, however, researchers have turned the reasoning around—since many of the components involved in sleep in worms are homologous to those in animals, researchers have turned to them as a potential source for identifying the contribution of sleep to autonomy and answering the question of why animals sleep (Keene and Duboue 2018).

Functional considerations also provide guidance as to how much abstraction makes sense in characterizing a mechanism. If abstracting from details of a given mechanism still enables researchers to explain how it contributes to the function it performs, the account will be regarded as informative. Moreover, insofar as it shows how the phenomenon is produced in different cases as a result of them all satisfying the abstractly characterized operations, it will reveal that the differing details do not make a difference to the phenomenon (as described). Thus, when Alon (2007) demonstrated that motifs—subnetworks of two, three, or four units—under broad assumptions about the character of the entities realizing the units, would exhibit comparable behaviors on which the organism relied, he showed that it was the pattern of connection, not the specific constitution of the units that accounted for the function (Levy and Bechtel 2013). If, however, the function is lost as one leaves out detail, then the account will be dismissed as too abstract.

Before turning to the contributions appeals to mechanisms can make to the autonomy tradition, we briefly consider how mechanists' engagement with two biological phenomena, homeostasis and self-organization, might benefit from engagement with the autonomy tradition. In the case of homeostasis, Bechtel and Abrahamsen (2009) articulate Krebs' reasoning about how cycles such as the ones he discovered allow chemical reactions to regenerate their components, while Craver and Darden (2013) introduce negative feedback as a strategy for explaining homeostatic phenomena by maintaining an equilibrium state. Homeostasis is often viewed as a good in itself, but the autonomy tradition motivates inquiring how, in different cases, it serves or does not serve the maintenance of the organism. Deviations from a usual range of values are not necessarily bad, and the maintenance of homeostasis is not necessarily good: under threat conditions restoration of normal values may be disastrous. As Bich et al. (2020) argue, whether maintaining homeostasis is good or bad depends on the general state of the organisms and its adaptive needs. Likewise, self-organization can be advantageous in organizing living systems and Glennan (2017) has shown how it can be understood in terms of affinitive relations between

parts of mechanisms. Discussions of self-organization have figured prominently on discussion of the origin of life (Moreno and Ruiz-Mirazo 2009), in morphogenesis during developmental processes (Newman and Forgacs 2005), and more recently in cell biology (Karsenti 2008), offers microtubule patterning as an example). But, as Mossio and Bich (2017) emphasize, self-organization is not sufficient for autonomy—that requires self-organizing subsystems to contribute to the maintenance of an organism. This often requires that self-organizing capacities by brought under various control processes. Homeostasis and self-organization each are important for understanding how biological organisms are autonomous systems. Yet, they alone do not make a system autonomous. In the case of both homeostasis and self-organization, mechanists can benefit by following the autonomy tradition and focusing on when they contribute to the autonomy of the organism.

Mechanistic approaches ground explanations of the autonomy of organisms

Turning to the autonomy tradition, the ambivalent attitude towards appeal to mechanisms and the emphasis on abstraction, which can be traced back to its early roots but which is still present in more recent work, runs the risk of hindering the ability to flesh out the actual causal relationships underlying autonomy. While not denying the paramount importance of foundational theoretical work, it is also crucial to integrate that work with an understanding of the underlying biological processes. Developing a correct theory of living organisms, such as the autonomy approach aims to do, should be grounded in an understanding of how autonomy is actually realized by biological systems.

The other limitation identified in Sect. "Challenges confronting autonomy theorists on which mechanists can offer assistance" is that the autonomy tradition often seems to embrace a holism that countenances undifferentiated organization. It is not surprising then, that most recent scientific work on autonomy has been carried out by employing tools such as network modeling, which have been applied to investigate the dynamical properties (e.g., stability and robustness) of abstract theoretical models such as Rosen's (see, for example, Piedrafita et al. 2010).[18] As recently pointed out by Moreno and Suárez (2020), among others, network modelling is a holistic tool used to study and predict global dynamical properties of large sets of interacting entities.[19] While useful in characterizing and studying global activities of biological systems, network models, unless supplemented, do not provide a *causal explanation* of how the individual parts operate and interact within autonomous

[18] Properties less related to basic self-maintenance such as reproduction have been given less attention, although with some exceptions; see, for example, Mavelli and Ruiz-Mirazo's (2007) proposal advanced in the context of discussions on the origins of life.

[19] For an analysis of the relationship between networks and mechanisms, and how to use network description to identify mechanisms see Bechtel (2015a, 2015b).

systems to produce these global activities. Not only that: for network models to provide a more reliable and biologically relevant understanding of global properties, they must be constructed on information about the parts and their operations.[20]

The abstractness and holism of the autonomy approach can be properly grounded by developing causal mechanistic explanations of biological functions that identify the constraints generated within the autonomous system. As we have noted, the autonomy framework provides a way to identify functions in the context of the organization of the system. The multiplicity of functions points to a multiplicity of constraints (mechanisms) that contribute differentially to the maintenance of that organization (Mossio et al. 2009). By examining these, the autonomy tradition can ground its abstract accounts in concrete causal details about how autonomous systems in fact maintain themselves.[21]

One further holistic aspect which is still manifest in some recent accounts of autonomy is the identification of one function with one constraint subject to closure, as proposed by Mossio et al. (2009) and Montevil and Mossio (2015). Sometimes, depending on the scale of observation, it can be the case that a single constraint can be associated with a biological function without losing explanatory power. For example, for some explanatory purposes, the heart can be considered as an individual constraint which functionally contributes to the circulation of blood. Yet pursuing this strategy risks overlooking complexity underlying the realization of a biological function and how this complexity matters for the overall functioning of the system. Different parts of the heart (the muscles, valves, etc.) contribute differently to the function of the heart. What renders it plausible to treat it as one constraint instead of a mechanism employing several interacting constraints is that it appears as a single organ. Yet analyzing the organ as a mechanism constituted by several constraints would be especially informative when the organ is malfunctioning or to understand how it is controlled.

The strategy of identifying one function with one constraint is less plausible for functions that are realized through different steps by different types of constraints. Consider vascularization. The whole vascular system can be considered as one macroscopic constraint acting on the flow of blood. However, if the objective is understanding the rate at which blood is flowing to a certain part of the body, it will be beneficial to describe the system as a mechanism and to identify the relevant constraints that constitute it, such as the heart, arteries and veins, muscles compressing veins, etc. A similar consideration applies at smaller scales. The activity of complex molecular machinery such as ATPases and how they perform their function in the cell, can be explained if they are described in mechanistic terms as ordered structures of interacting constraints (Militello and Moreno 2018). Mechanistic details are useful for understanding allosteric enzymes in which it is informative to distinguish as different constraints the allosteric and catalytic sites, as well as additional phosphorylation sites. Controlled catalysis results from the interaction of these components.

[20] Bechtel and Abrahamsen (2010) draw a contrast between modeling of circadian rhythms, which is typically grounded on information about the component parts and operations of circadian clocks, and modeling of cognitive capacities in cognitive science, which is typically not so grounded.

[21] Initial work in this direction has been done in the case of glycaemia regulation (Bich, Mossio, & Soto, 2020) and organogenesis (Montévil et al. 2016).

As the case of allosteric enzymes makes clear, the risks implied by a holistic view of the relationship between constraints and functions, and the advantages of adopting mechanistic explanations in the autonomy approach are especially evident in the case of control. To achieve control, controllers must base their actions on states of sensors that measure the state of the system or the environment. In biological systems these two capabilities are often embedded in distinct and even noncontiguous structures, and the connection between them is realized by further constraints acting as signal transducers. Consider the control of the direction of movement in bacteria chemotaxis. Chemotaxis is a form of behavioral adaptive response to changing environmental conditions. Bacteria such as *E. coli* alter the direction in which they swim on the basis of the gradient in the concentration of metabolites in the medium they inhabit (Eisenbach 2004). They do this by changing the mode of rotation of their flagella while exploring their environment. The control of movement relies on at least three subsystems: (a) a complex receptor subsystem of transmembrane methyl-accepting chemotaxis proteins (MCPs) activated by specific chemoattractants or chemorepellants; (b) proteins which transduce the state of the receptors to the flagella through the cytoplasm by means of non-covalent post-translational modifications (phosphorylations); and (c) the flagella themselves, molecular complexes which include control sites where the modulation of the mode of rotation takes place.

Considering the whole control subsystem (receptors, transduction proteins and control sites on the flagella) as an individual functional constraint fails to explain why and how a bacterium moves adaptively in the environment to maintain itself. The complexity of the sensor complex in integrating different types of measurements, for example, can be better understood by adopting a mechanistic explanation within the autonomy framework. In *E. coli* the receptor complex is made of five types of MCPs which are sensitive to different ligands, plus Che proteins that, by methylating and demethylating the MCPs, allow the latter to adapt in response to the previous concentration so that they measure whether concentrations are increasing and decreasing. The differing number and position of these molecules also accounts for differences in responses between individual bacteria belonging to the same species. The adaptive behavior of the bacterium, therefore, can be better explained as a functional phenomenon within the autonomy approach if control is explained mechanistically as the result of the organization of a set of constraints.

In addition, the holistic approach may overlook another central aspect of control: the *loci* where control is exerted. Control can be exerted on different structures that constrain different steps of a process. The presence of specific *loci* where control is exerted plays an important role in the context of the organism. It allows it to integrate and coordinate different functions. Consider a functional pathway realized by more than one constraint. If one or more constraints responsible for some given steps are flexible, their operations can, in principle, be modulated by control constraints in the system, which in turn operate on the basis of measurements of the state of the system and the environment. When more than one constraint is flexible, and each is controlled by distinct control constraints sensitive to different variables, the mechanism can work as an integration hub and operate on the basis of different sources of

information (which in turn can be controlled by several other constraints). Consider the phenomenon of insulin release in mammals, which is critical in controlling the metabolism and concentration of glucose, with glucose in turn contributing to the maintenance of the constraints involved.[22] Pancreatic beta-cells continuously produce and store insulin in vesicles and secrete it into the bloodstream. The secretion of insulin occurs when the insulin-containing vesicles already present in the cells fuse with the cell membrane and release the hormone. In the presence of a high concentration of glucose, ATP production in beta-cells is increased and is measured by sensitive constraints such as K^+-channels that invert the polarity of the cell membrane and trigger a cascade of changes of activation states of several constraints in the cells, involving changes in potentials, that lead to the fusion of insulin containing vesicles with the membrane and the massive release of the hormone. This control mechanism, however, is not only sensitive to glucose concentrations (indirectly through ATP levels); it is responsive to outputs from other control mechanisms in the system which are sensitive to many other, non-metabolic factors. As a result, the nervous system, the intestines, and adipose and muscle tissues also participate in the modulation of insulin (Roder et al. 2016). Addressing glycaemia regulatory function in terms of control mechanisms enables investigators to understand the integration of all these contributions to the maintenance of the system.

Decomposition of systems into their components is the hallmark of mechanistic approaches. Adopting it within the autonomy tradition can ground explanations of autonomy. Moreover, it can do so without sacrificing the distinctive feature of the autonomy tradition—a focus on contributions to the maintenance of the organism. Functions and the components to which functions are ascribed are identified *top-down*, starting with their contribution to the maintenance of the system rather than from their structural features (Bich 2012). Once a function is identified as the phenomenon to be explained mechanistically, the constraints involved in its realization and the processes they take part in need to be characterized. First, production constraints and processes need to be identified. Second, an account of how they are produced and maintained within the system needs to be provided. Third, the possible control constraints acting upon them need to be identified. This implies also providing an account of how the controllers detect conditions in the organism or its environment that require adaptation and operate appropriately on constraints that determine activities in the organism.

Conclusion: a call for constructive engagement between mechanism and autonomy

The new mechanists and the autonomy tradition advance alternative perspectives for explaining the capacities of living organisms. Each approach on its own has produced important insights. We have argued that each can also benefit from

[22] See Bich et al. (2020) for more details of the explanation of glycaemia regulation from the point of view of autonomy.

engagement with the other. Mechanists can draw upon insights of the autonomy tradition in addressing questions about how to characterize phenomena, identity components of mechanisms, and generalize mechanistic accounts. Theorists in the autonomy tradition can ground their accounts of how autonomy is achieved in an understanding of the mechanisms employed. In the last two sections we have presented examples of ways in which the concepts of constraint, function, and control provide foundations for successfully bridging the two philosophical traditions. We view these examples as just a start of what could be a prolonged and productive engagement between mechanism and autonomy.

We conclude by emphasizing control as a promising locus of engagement. Recognizing the richness of control processes in biology makes clear that mechanisms do not just explain phenomena but explain how they are regulated to serve the needs of organisms and contribute to their autonomy. We have singled out a specific kind of mechanism—a control mechanism that operates on constraints in other mechanisms on the basis of measuring states of the system or the environment. Many of the mechanisms examined by the new mechanists are in fact control mechanisms (brain mechanisms are a paradigmatic example). However, their role in controlling other mechanisms has not assumed any special significance within the new mechanistic framework. They are treated like other mechanisms—causal systems that generate phenomena such as producing action potentials, spatial navigation (Craver 2007), or visual perception (Bechtel 2008). These activities are not ends in their own right, but serve to regulate what we earlier characterized as production mechanisms. Recognizing that role allows us to understand why they, together with processes such as cell signaling (Marks et al. 2017), are characterized as processing information—through making measurements they acquire information about the organism or its environment and use this information to operate on other mechanisms. The importance of this distinction can be captured from the autonomy perspective. Production mechanisms do not just exist to generate the phenomenon for which they are responsible. Rather, they operate to maintain the autonomy of the organism. This requires that they generate the phenomenon when and in the manner that is needed by the organism. For this, control is paramount. Control mechanisms perform the measurements (or draw upon the measurements) of variables that determine whether the activity of the mechanism is needed.

Recognizing the role of control mechanisms provides a new perspective on another feature of accounts of mechanism. Production mechanisms are often presented as operating in a regular manner—any time the start-up conditions are realized, they produce the phenomenon for which they are responsible (Machamer et al. 2000). Such regularity, however, is often only achieved in experimental arrangements in which mechanisms are studied—it is the specific arrangement of conditions in the experiment that render the operation of a mechanism regular. This enables researchers to develop accounts of how they operate. But in living organisms, regular behavior might be the exception. As adaptive systems, the phenomenon required—and hence the work required from a biological mechanism—depends on the state of the system and its environment. Producing the phenomenon is a means to achieve self-maintenance. Accordingly, an organism can vary its behavior, and therefore the operation of the mechanisms responsible for it, when the specific

internal and external conditions require an adaptive response. Inability to adapt—for example, a heart that cannot change the frequency of contractions in relation to the needs for oxygen in the muscles of the organism—is pathological.

Control mechanisms are ubiquitous in biology—both in single- and multi-cellular organisms, they vastly outnumber production mechanisms. Moreover, many of the changes over the course of evolution have involved the incorporation of new control mechanisms that enable organisms to better deploy production mechanisms that perform the work needed to maintain, repair, and reproduce themselves (Bechtel and Bich 2021). In addition, control is central to understanding how different biological functions are coordinated to make an organism viable. Insofar as new mechanists have focused on inquiries directed within individual mechanisms, which are often studied in conditions in which their activity is regularized, they have failed to recognize the importance of these distinctive features of control. Such a focus on individual mechanisms has also characterized much of traditional biology. Only recently have topics such as cell signaling become prominent topics of research. Control is a challenge to the autonomy framework as well, for understanding it requires examining the details through which control is realized—decomposing and analyzing the operation of control mechanisms, especially how they procure and use information. Understanding biological control will benefit from the collaboration of the two approaches. If philosophers of science interested in mechanistic explanation want to understand how mechanisms are controlled to serve organisms, they can lean on the autonomy tradition. At the same time, the autonomy framework can benefit from a mechanistic approach to account for how these functions are realized and how they are integrated into the organization of the system.

Acknowledgements The authors thank Alvaro Moreno for the careful reading and useful comments on a previous version of this paper.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. The authors acknowledge funding from the Basque Government (Project: IT1228-19 for LB), Ministerio de Ciencia, Innovación y Universidades, Spain (research project PID2019-104576 GB-I00 for LB and WB, and 'Ramon y Cajal' Programme RYC-2016–19798 for LB).

References

Alon U (2007) Network motifs: theory and experimental approaches. Nat Rev Genet 8(6):450–461

Ankeny RA, Leonelli S (2011) What's so special about model organisms? Stud History Philos Sci Part A 42(2):313–323

Ankeny RA, Leonelli S (2020) Model organisms. Cambridge University Press, Cambridge

Artiga M, Martinez M (2016) The organizational account of function is an etiological account of function. Acta Biotheor 64(2):105–117

Batterman RW, Rice CC (2014) Minimal model explanations. Philos Sci 81(3):349–376
Bechtel W (2008) Mental mechanisms. Philosophical perspectives on cognitive neuroscience. Routledge, London
Bechtel W (2015a) Can mechanistic explanation be reconciled with scale-free constitution and dynamics? Stud History Philos Sci Part C Stud History Philos Biol Biomed Sci 53:84–93
Bechtel W (2015b) Generalizing mechanistic explanations: Using graph-theoretic representations. Explanation in biology: an enquiry into the diversity of explanatory patterns in the life sciences. Springer, Dordrecht, pp 199–225
Bechtel W (2018) The importance of constraints and control in biological mechanisms: insights from cancer research. Philos Sci 85(4):573–593
Bechtel W (2020) Rethinking psychiatric disorders in terms of htererarchical networks of control mechanisms. In: Kendler KS, Parnas J, Zachar P (eds) Levels of analysis in psychopathology. Cambridge University Press, Cambridge, pp 24–46
Bechtel W, Abrahamsen A (2005) Explanation: a mechanist alternative. Stud Hist Philos Biol Biomed Sci 36(2):421–441
Bechtel W, Abrahamsen A (2010) Dynamic mechanistic explanation: computational modeling of circadian rhythms as an exemplar for cognitive science. Stud History Philos Sci Part A 41(3):321–333
Bechtel W, Bich L (2021) Grounding cognition: heterarchical control mechanisms in biology. Philos Trans R Soc London B Biol Sci 376(1820):20190751
Bechtel W, Abrahamsen A (2009) Complex biological mechanisms: Cyclic, oscillatory, and autonomous. In: C. A. Hooker (Ed), Philosophy of complex systems. Handbook of the philosophy of science, vol 10, pp 257–285. New York: Elsevier
Bechtel W, Richardson RC (1993/2010) Discovering complexity: Decomposition and localization as strategies in scientific research. Cambridge, MIT Press. 1993 edition published by Princeton University Press
Bich L (2012) Complex emergence and the living organization: an epistemological framework for biology. Synthese 185(2):215–232
Bich L, Mossio M, Ruiz-Mirazo K, Moreno A (2016) Biological regulation: controlling the system from within. Biol Philos 31(2):237–265
Bich L, Mossio M, Soto AM (2020) Glycemia regulation: from feedback loops to organizational closure. Front Physiol, 11(69)
Bogen J, Woodward J (1988) Saving the phenomena. Philos Rev 97:303–352
Boone W, Piccinini G (2016) Mechanistic abstraction. Philos Sci 83(5):686–697
Chirimuuta M (2014) Minimal models and canonical neural computations: the distinctness of computational explanation in neuroscience. Synthese 191(2):127–153
Christensen WD, Bickhard MH (2002) The process dynamics of normative function. Monist 85:3–28
Craver CF (2007) Explaining the brain: mechanisms and the mosaic unity of neuroscience. Oxford University Press, New York
Craver CF, Darden L (2013) In search of mechanisms: discoveries across the life sciences. University of Chicago Press, Chicago
Craver CF, Kaplan DM (2020) Are more details better? On the morms of completeness for mechanistic explanations. Br J Philos Sci 71(1):287–319
Cummins R (1975) Functional analysis. J Philos 72:741–765
Darden L (2013) Mechanisms versus causes in biology and medicine. In: Chao H-K, Chen S-T, Millstein RL (eds) Mechanism and causality in biology and economics. Springer, Dordrecht, pp 19–34
Darden L, Craver C (2002) Strategies in the interfield discovery of the mechanism of protein synthesis. Stud Hist Philos Biol Biomed Sci 33:1–28
Darden L, Kundu K, Pal LR, Moult J (2018) Harnessing formal concepts of biological mechanism to analyze human disease. PLOS Comput Biol 14(12):e1006540
Di Paolo EA (2005) Autopoiesis, adaptivity, teleology, agency. Phenomenol Cogn Sci 4(4):429–452
Eisenbach M (2004) Chemotaxis. Imperial College Press, London
Frick R, Bich L, Moreno A (2019) An organisational approach to biological communication. Acta Biotheor 67(2):103–128
Friston K (2009) The free-energy principle: a rough guide to the brain? Trends Cogn Sci 13(7):293–301
Galvani L (1791) De viribus electricitatis in motu musculari commentarius. Bologna: Ex typographia Instituti Scientiarum
Gánti T (1975) Organization of chemical reactions into dividing and metabolizing units: the chemotons. BioSystems 7:15–21
Gánti T (2003) The principles of life. New York, Oxford

Garson J (2013) The functional sense of mechanism. Philos Sci 80(3):317–333
Glansdorff P, Prigogine I (1971) Thermodynamic theory of structure, stability and fluctuations. Wiley, London
Glennan S (2017) The new mechanical philosophy. Oxford University Press, Oxford
Hardesty RA (2018) Much ado about mice: standard-setting in model organism research. Stud History Philos Sci Part C Stud in History Philos Biol Biomed Sci 68–69:15–24
Hastings MH, Herzog ED (2004) Clock genes, oscillators, and cellular networks in the suprachiasmatic nuclei. J Biol Rhythms 19(5):400–413
Hochstein E (2016) One mechanism, many models: a distributed theory of mechanistic explanation. Synthese 193(5):1387–1407
Hooker CA (2013) On the import of constraints in complex dynamical systems. Found Sci 18(4):757–780
Joiner WJ (2016) Unraveling the evolutionary determinants of skoizleep. Curr Biol 26(20):R1073–R1087
Kaiser MI (2017) The components and boundaries of mechanisms. In: Glennan S, Illari P (eds) The routledge handbook of mechanisms and mechanical philosophy. Routledge, London, pp 116–130
Kaplan DM (2011) Explanation and description in computational neuroscience. Synthese 183(3):339–373
Karsenti E (2008) Self-organization in cell biology: a brief history. Nat Rev Mol Cell Biol 9(3):255–262
Kauffman SA (1970) Articulation of parts explanation in biology and the rational search for them. In: Bluck RC, Cohen RS (eds) PSA 1970. Reidel, Dordrecht, pp 257–272
Kauffman SA (2000) Investigations. Oxford University Press, Oxford
Kauffman SA (2019) A world beyond physics: the emergence and evolution of life. Oxford, Oxford
Keene AC, Duboue ER (2018) The origins and evolution of sleep. J Exp Biol, 221(11): jeb159533
Krickel B (2018) Saving the mutual manipulability account of constitutive relevance. Stud History Philos Sci Part A 68:58–67
LaFollette H, Shanks N (1996) Brute science: Dilemmas of animal experimentation. Routledge, London
Levy A (2013) What was Hodgkin and Huxley's Achievement? Br J Philos Sci 65(3):469–492
Levy A, Bechtel W (2013) Abstraction and the organization of mechanisms. Philos Sci 80(2):241–261
Levy A, Currie A (2014) Model organisms are not (theoretical) models. Br J Philos Sci 66(2):327–348
Levy A, Bechtel W (forthcoming). Doing without (explanatory) completeness
Luisi PL (2006) The emergence of life: from chemical origins to synthetic biology. Cambridge University Press, Cambridge
Machamer P, Darden L, Craver CF (2000) Thinking about mechanisms. Philos Sci 67:1–25
Marks F, Klingmüller U, Müller-Decker K (2017) Cellular signal processing: An introduction to the molecular mechanisms of signal transduction (Second edition ed)
Marom S (2010) Neural timescales or lack thereof. Prog Neurobiol 90(1):16–28
Maturana H, Varela FJ (eds) (1980) Autopoiesis: the organization of the living. In: Autopoiesis and cognition: the realization of the living. Reidel, Dordrecht, pp 73–138
Maturana H, Varela FJ (1987) The tree of knowledge: the biological roots of human understanding. Shambhala, Boston
Maturana H, Mpodozis J (2000)The origin of species by means of natural drift. Rev Chil Hist Nat 73(2):261–310
Mavelli F, Ruiz-Mirazo K (2007) Stochastic simulations of minimal self-reproducing cellular systems. Philos Trans R Soc B Biol Sci 362(1486):1789–1802
McLaughlin P (2001) What functions explain: functional explanation and self-reproducing systems. Cambridge University Press, Cambridge
Militello G, Moreno Á (2018) Structural and organisational conditions for being a machine. Biol Philos 33(5–6). https://doi.org/10.1007/s10539-018-9645-z
Millikan RG (1989) In defense of proper functions. Philosophy of Science 56:288–302
Montévil M, Speroni L, Sonnenschein C, Soto AM (2016) Modeling mammary organogenesis from biological first principles: Cells and their physical constraints. Prog Biophys Mol Biol 122(1):58–69
Montevil M, Mossio M (2015) Biological organisation as closure of constraints. J Theor Biol 372:179–191
Moreno A, Mossio M (2015) Biological autonomy: A philosophical and theoretical inquiry. Springer, Dordrecht
Moreno A, Ruiz-Mirazo K (2009) The problem of the emergence of functional diversity in prebiotic evolution. Biol Philos 24(5):585–605
Moreno A, Suárez J (2020) Plurality of explanatory strategies in Biology: mechanisms and networks. In: Gonzalez WJ (ed) Methodological prospects for scientific research: from pragmatism to pluralism. Springer International Publishing, Cham, pp 141–165

Mossio M, Bich L (2017) What makes biological organisation teleological? Synthese 194:1089–1114
Mossio M, Saborido C (2016) Functions, organization and etiology: a reply to Artiga and Martinez. Acta Biotheor 64(3):263–275
Mossio M, Saborido C, Moreno A (2009) An organizational account of biological functions. Br J Philos Sci 60(4):813–841
Mossio M, Pontarotti G (2019) Conserving functions across generations: Heredity in light of biological organization. British J Philos Sci
Neander KL (1991) The teleological notion of 'function'. The Australasian J Philos, 69(454–468)
Newman SA, Forgacs G (2005) Complexity and self-organization in biological development and evolution. In: Bonchev D, Rouvray DH (eds) Complexity in chemistry, biology, and ecology. Springer, Boston, pp 49–95
Nicholson DJ (2012) The concept of mechanism in biology. Stud Hist Philos Biol Biomed Sci 43(1):152–163
Nicholson DJ, Dupré J (eds) (2018) Everything flows. Oxford University Press, NewYork
Pattee HH (1972) Laws and constaints, symbols and languages. In: C. H. Waddington (Ed), Towards a theoretical biology. Chicago: Adine-Atherton
Piaget J (1967) Biologie et connaissance. Gallimard, Paris
Piedrafita G, Montero F, Moran F, Cardenas ML, Cornish-Bowden A (2010) A simple self-maintaining metabolic system: Robustness, autocatalysis, bistability. Plos Comput Biol, 6(8)
Roder PV, Wu B, Liu Y, Han W (2016) Pancreatic regulation of glucose homeostasis. Exp Molecul Med, 48, e219
Rosen R (1972) Some relational cell models: The metabolism-repair systems. In: Rosen R (ed) Foundations of mathematical biology, vol II. Academice, New York, pp 217–253
Rosen R (1991) Life itself: A comprehensive inquiry into the nature, origin, and fabrication of life. Columbia University Press, New York
Ruiz-Mirazo K, Moreno A (2004) Basic autonomy as a fundamental step in the synthesis of life. Artif Life 10:235–259
Saborido C, Moreno A (2015) Biological pathology from an organizational perspective. Theor Med Bioeth 36(1):83–95
Saborido C, Mossio M, Moreno A (2011) Biological organization and cross-generation functions. Br J Philos Sci 62(3):583–606
Shaw PJ, Tononi G, Greenspan RJ, Robinson DF (2002) Stress response genes protect against lethal effects of sleep deprivation in Drosophila. Nature 417(6886):287–291
Spudich JA, Kron SJ, Sheetz MP (1985) Movement of myosin-coated beads on oriented filaments reconstituted from purified actin. Nature 315(6020):584–586
Steel D (2007) Extrapolation, capacities, and mechanisms. In: D. Steel (Ed.), Across the boundaries: extrapolation in biology and social science: Oxford
Varela FJ (1979) Principles of biological autonomy. North Holland, New York
Varela FJ, Maturana H (1972) Mechanism and biological explanation. Philos Sci 39:378–382
Varela FJ, Maturana H, Uribe R (1974) Autopoiesis: the organization of living systems, its characterization and a model. BioSystems 5:187–196
Varela FJ, Thompson E, Rosch E (1991) The embodied mind. MIT Press, Cambridge
Weber M (2005) Philosophy of experimental biology. Cambridge University Press, Cambridge
Wimsatt WC (1972) Teleology and the logical structure of function statements. Stud History Philos Sci 3:1–80
Winning J (2020) Internal perspectivalism: the solution to generality problems about proper function and natural norms. Biol Philos 35(3):33
Winning J, Bechtel W (2018) Rethinking causality in neural mechanisms: constraints and control. Mind Mach 28(2):287–310
Woodward J (2003) Making things happen: a theory of causal explanation. Oxford University Press, Oxford
Zhang EE, Kay SA (2010) Clocks not winding down: unravelling circadian networks. Nat Rev Molecul Cell Biol 11(11):764–776