revista
CONTROVÉRSIA
ISSN 1808-5253
UNISINOS

TOP-DOWN AND BOTTOM-UP CONSTRAINTS IN MECHANISTIC INQUIRY

DOI: https://doi.org/10.4013/con.2023.193.06

Matheus Diesel Werberich

Doutorando em Neurociência e Filosofia pela Washington University in Saint Louis

matheusdiesel01@gmail.com

https://orcid.org/0000-0002-9297-9089

ABSTRACT:

Mechanisms play a crucial role in scientific research across various disciplines, and philosophers of science have devoted significant effort into understanding their ontology and epistemology. This paper examines the relationship between mechanisms and phenomena, highlighting the inherent dependence of mechanistic delineation on the characterization of phenomena. By acknowledging that characterizing phenomena is influenced by pragmatic considerations and research interests, the paper argues that mechanistic inquiry is inherently shaped by researchers' perspectives. This dependence raises concerns about the possibility of a realist view of mechanisms. To address these concerns, the paper explores how top-down constraints, rooted in researchers' interests and pragmatic concerns, can be balanced by bottom-up constraints derived from empirical considerations. In conclusion, I argue that the interplay between these constraints forms an empirical and realist counterweight to the perspectival nature of top-down constraints.

KEY-WORDS:

Mechanisms. Philosophy of Science. Realism. Antirealism. Pragmatism.

RESTRIÇÕES TOP-DOWN E BOTTOM-UP NA PESQUISA MECANICISTA

RESUMO:

Mecanismos desempenham um papel crucial na pesquisa científica em várias disciplinas, e filósofos da ciência têm se dedicado para entenderem sua ontologia e epistemologia. Este artigo examina a relação entre mecanismos e fenômenos, destacando a dependência inerente do delineamento mecanicista na caracterização de fenômenos. Ao reconhecer que a caracterização dos fenômenos é influenciada por considerações pragmáticas e interesses de pesquisa, o artigo argumenta que a pesquisa mecanicista é inerentemente moldada pelas perspectivas dos pesquisadores. Essa dependência levanta preocupações sobre a possibilidade de uma visão realista dos mecanismos. Para comportar essas preocupações, o artigo explora como as restrições *top-down*, enraizadas nos interesses e preocupações pragmáticas, podem ser balanceadas por restrições *bottom-up* derivadas de considerações empíricas. Concluindo, eu argumento que a interação entre essas restrições forma um contrapeso empírico e realista da perspectiva natural de restrições *top-down*.

PALAVRAS-CHAVE:

Mecanismos. Filosofia da Ciência. Realismo. Antirealismo. Pragmatismo.

Introduction[1]

Mechanisms are ubiquitous in the empirical sciences, especially in the biological and cognitive sciences. Scientists often describe themselves as uncovering mechanisms for a variety of phenomena - from protein synthesis to global economic recessions. It should then come as no surprise that philosophers of science are particularly interested in the ontology and epistemology of mechanisms and the methodologies we use to discover them (GLENNAN, 2017; MACHAMER; DARDEN; CRAVER, 2000). Mechanistic philosophy is primarily concerned with questions such as how mechanisms explain, how do scientists separate mechanisms from their surrounding environment, whether mechanisms are real or just explanatory tools for empirical research, among others.

While mechanistic philosophers tend to disagree on some of these issues, they for the most part agree that individuating mechanisms are highly dependent on what we take its phenomenon to be. In short, mechanisms are always mechanisms *for* some phenomenon (DARDEN, 2008; GLENNAN, 2017; MACHAMER; DARDEN; CRAVER, 2000). In this context, "phenomenon" indicates the event that researchers want to explain; from an initial description of the *explanandum,* researchers try to figure out which elements and interactions within that system are relevant for the phenomenon. The complete description of these elements and the interactions between them is a mechanistic explanation for that phenomenon. So constructed, the phenomenon is nothing more than the behavior of the mechanism.

The dependence of mechanistic delineation upon the characterization of phenomena can entail some problems for a realist position for mechanisms. Considering that characterizing phenomena is dependent on a research project, and that such are dependent on some pragmatic considerations, then mechanistic inquiry is also dependent on pragmatic interests of researchers. As such, mechanistic inquiry will vary significantly between research projects and, hence, may not constitute a completely interest-free framework

---

[1] This paper is based on part of my Master's dissertation work, conducted at the Federal University of Santa Maria (UFSM), Brazil. I thank participants of the UFSM Philosophy of Memory Lab (Mem-Lab) for their comments in previous versions of this article. This study was financied in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Finance Code 001.

for the empirical sciences, as well as inevitably leading to an anti-realist view of mechanisms. In this article, I present some defenses of mechanistic inquiry from the threat of antirealism.

The paper is structured as follows. In section 1, I give an overview of the new mechanistic literature, especially as it pertains to the relation between phenomenon delineation and mechanistic inquiry. In section 2, I show how researchers' interests and pragmatic concerns are intrinsically at play when delineating phenomena. How the phenomenon is constructed grounds what Bechtel & Richardson (2010) call *top-down constraints*. Doubts about the realism of mechanism stem primarily from such constraints and their intrinsic pragmatic considerations. As a response to these doubts, I argue in section 3 that constraints at the physical level also limit the range of possible descriptions of phenomena, thus amounting to an empirical and realist counterweight to the perspectivalist nature of top-down constraints. Such physical considerations are called *bottom-up constraints* (BECHTEL; RICHARDSON, 2010). In section 4, I analyze two examples of how top-down and bottom-up constraints are employed in actual scientific practice. As an example of top-down constraints, I overview how abstract computational accounts of cognitive processes guide researchers' inquiry into their underlying mechanisms (DEWHURST, 2018; SHAGRIR; BECHTEL, 2017). As an example of bottom-up constraints, I discuss how Biderman & Shohamy (2021)'s work on memory and decision making influenced how the phenomenon of decision making is delineated - instead of having the final behavioral decision as a terminating condition, Biderman & Shohamy (2021) suggest that the mechanism for deliberation is still active well after the behavioral output. Finally, in section 5, I summarize the previous discussions and explore some implications of the present framework to the task of delineating mechanisms.

2. Mechanisms

In the dawn of philosophy of science, there was a significant emphasis on how to unify scientific knowledge in a large, overarching framework. Exemplified by Carnap's *Der logische Aufbau der Welt*, a main goal of philosophers at the time was to find a way to translate all true scientific statements into a purely empirical language, composed of sense-data and certain logical structures. Still influenced by the lofty goals of Carnap and others, Oppenheim & Putnam (1958) conceived a hierarchy of scientific enterprises, where "higher" disciplines (such as psychology) could be reduced to "lower" disciplines (such as chemistry and physics). In this framework, the reduction relation carries most of science's explanatory weight: to explain a phenomenon means to reduce it to more fundamental entities and describe their interactions with the laws of physics.

Throughout the twentieth century, however, the reductionist framework in the philosophy of science has steadily declined in popularity. One reason for such decline is how scientific practices in biology, psychology, and other "higher" sciences are notoriously difficult to understand as reductions of a target *explanandum*. Moreover, such sciences often lack the rigid structure necessary for a reductionist explanation: their laws often have exceptions, their classifications are cross-cutting in different ways, and their explanatory tools are not primarily mathematical. Instead, the life sciences seemed more interested in investigating the *causal* relations that produce their *explananda*. Thus, Salmon (1984a), Woodward (1984), and Cummins (1985), just to name a few, shifted their focus to how causal and constitutive relations are prevalent in the life sciences. Not only had reductionism fallen out of flavor, the idea of physics as the ultimate model for the scientific enterprise was also replaced by an emphasis on how fuzzy boundaries are prevalent in "higher" sciences, as well as how individualistic explanations can be.

The transition from physics and reductionism to biology and causal relations paved the way for a mechanistic renaissance in the late 1990s and early 2000s. Instead of focusing on how scientific explanations present law-like generalizations, philosophers of science began to pay attention to how explanations often are about how one particular system works in virtue of its underlying components. In this framework, to explain a system means to decompose it into smaller parts, whose interactions and overall organization make up the behavior of the system as a whole (MACHAMER; DARDEN; CRAVER, 2000). Such decomposition amounts to describing the underlying *mechanism* for that system. In the words of Bechtel & Abrahamsen (2005):

> A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena (p. 423).

This description of mechanisms makes clear that, for any mechanistic explanation, there are three elements that are individually necessary and jointly sufficient:

1. A description of the phenomenon to be explained;
2. A description of which components are responsible for the phenomenon;
3. An account of how these components are coordinated in such a way to be responsible for the phenomenon.

There are several challenges researchers must face when tackling these aspects of mechanistic explanation. Firstly, phenomena are often not clearly delineated from their surrounding environment. Consider, for example, the phenomenon of a beating heart: not only is it connected with several different parts of the body, but it is also affected by our breathing rate, stress levels, muscular movements, and so on. This obviates the need for scientists to separate the *explanandum* from other contextual factors that, nevertheless, are causally interactive with the phenomenon. Scientists thus need to separate what they want to explain from the "busy and buzzing confusion that constitutes the causal structure of the world" (CRAVER, 2013, p. 140).

Secondly, and relatedly, among all the causal factors observed interacting with the phenomenon, not all of them will be relevant for its mechanism. For instance, the heart beating is affected by whether we are moving, but bodily movement is not a necessary condition for the heart to beat. Separating what is *actually* relevant for a phenomenon from the myriad of background conditions and spurious effects is a necessary step towards an accurate mechanistic explanation. For this task, we need clear criteria on which components and operations we should include in a mechanistic description, and which we can just abstract away. Thirdly, after component individuation, we need an account of how they work together in such a way that they are responsible for the phenomenon in question.

These three elements of mechanistic explanations highlight a major feature of this framework: mechanisms are always *for* a given phenomenon (CRAVER, 2015; DARDEN, 2008; GLENNAN, 2017). This platitude indicates how characterizing mechanisms necessarily involves describing a set of activities that, together, make up the phenomenon. In this regard, a heart that doesn't beat, a clock that doesn't move, or a neuron that doesn't fire are not mechanisms precisely because they are not active in any meaningful way - i.e., they don't produce any phenomena. Mechanisms require phenomena, and phenomena are activities that the mechanism does. The initial characterization of the phenomenon, together with underlying assumptions on how it works, are initial and crucial elements that guide researchers' inquiry on how possibly the phenomenon is underlined. Such guidance is what are sometimes called *top-down* constraints, which are the topic of the following section.

3. Top-down constraints

As emphasized in the previous section, a necessary part of mechanistic explanations is properly characterizing the phenomenon to be accounted for. Such a step is necessary, albeit not solely sufficient, given the fact that there are no mechanisms which are not active in some way. And it is this activity that characterizes the *explanandum* phenomenon - i.e., what the mechanism *does* and what it is *for*.

Given that the end goal of mechanistic inquiry is to achieve a description of how a mechanism is responsible for some phenomenon, an investigation cannot start with mechanisms themselves. They are out there to be empirically discovered and not gratuitously stipulated from the armchair. In this regard, we must start from some preliminary characterizations, assumptions, and observations of the target phenomenon.

As we'll see in what follows, these elements constrain the range of possible mechanisms we are able to come up with. They are usually called *top-down constraints* (BECHTEL; RICHARDSON, 2010).

There are multiple ways of delineating a phenomenon and, hence, of establishing top-down constraints. We can characterize phenomena either as a particular causal role in a larger system (CUMMINS, 1985; STICH, 1985); as an etiological function that promotes the survival and fit of a given system (CRAVER, 2013; MILLIKAN, 1984); or as a computational function that manipulates inputs based on given set of rules and equations (DEWHURST, 2018; KAPLAN; CRAVER, 2011; SHAGRIR; BECHTEL, 2017). While this is not an exhaustive list of all possible ways of delineating a phenomena (see, e.g., Glennan (2017), chapter 5, for a tentative taxonomy of phenomena descriptions), they represent some of the most influential ways to fix *explananda* in both the life and mind sciences. These ways of characterizing phenomena should not be understood as mutually exclusive, but rather as possibly compatible ways of delineating a system (KÄSTNER; HAUEIS, 2021). In this regard, such plurality of phenomena descriptions is to be expected when we consider that the causal structure of the world is not readily demarcated for us to discern, thereby leaving open the possibility that there is more than one possible way of carving a phenomenon apart from the rest of its environment.

Regarding causal role and etiological accounts of phenomena, there is a crucial and somewhat obvious way in which the general environment of the system is a crucial element in delineating *explananda*. On causal role descriptions, the phenomenon is characterized in function of its interactions with the environment: for example, the behavior of the heart to pump blood can only be determined if we analyze how it interacts with other neighboring parts of the system, such as veins, arteries, and blood. Meanwhile, on etiological descriptions, the phenomenon's function in sustaining a given system is dependent on contextual constraints: for example, the etiological function of the heart is to aid in the gas exchange between cells and the respiratory system, thus maintaining the organism alive. This function can only be fulfilled if the organism is within an environment that allows for such gas exchange.

Moreover, some philosophers convincingly argue that computational descriptions are also crucially dependent on the features of the surrounding environment. Harbecke & Shagrir (2019), for instance, claim that computational accounts are strongly dependent on the context of explanation in the sense that, if the system were to be placed in a different environment, receiving different types of inputs, it would have to perform a different computational function to get the same output. Such computational contextualism is exemplified in Shagrir & Bechtel (2017)'s analysis on Marr (1981)'s account of edge detection in the human retina: the mathematical function proposed by Marr was only possible by "the observation that in our perceived environment sharp changes in light reflectance occur along physical edges such as boundaries of objects" (SHAGRIR; BECHTEL, 2017, p. 200).

In short, causal role, etiological, and computational descriptions of phenomena require consideration of the surrounding environment to fix mechanistic *explananda*. From this, we can extract some very important features of phenomena delineation and top-down constraints in general. Firstly, phenomena delineation is an empirical matter. We need empirical evidence to make these characterizations, since contextual features are necessary to take into account. Secondly, and relatedly, it is likely that our delineation of the phenomenon will change as research progresses: as we gather more data, we may discover some further relations and properties that necessitate a change in our general view of the phenomenon. Thirdly, for re-occurrent phenomena, multiple instances of the same phenomena will interact with different environmental elements and, hence, are bound to have some different properties. Researchers then need a way to say that the same phenomenon is happening in these different circumstances. One way of guaranteeing that is to appeal to how a certain pattern is repeated along these instances.

3.1 Patterns

The concept of *pattern* is relatively recent in the history of philosophy. Philosophers have increasingly taken interest in it after the publication of Dennett's seminal paper *Real Patterns* (1991). In his article, Dennett proposes a response to the problem of whether beliefs exist as real mental entities, or if they are only conceptual constructs of an outdated understanding of human psychology, soon to be replaced by neurocognitive descriptions. His answer claims that beliefs, along with other propositional attitudes, are specific patterns of thought and behavior which, despite having multiple different realizations in the brain, are useful abstractions for explaining and predicting behavior.

Beyond the application of patterns to the problem about propositional attitudes, Dennett provides a formal account of what a pattern is. He claims that patterns are a way of compressing information in such a way that unnecessary details are left out and only the most useful pieces of information survive. For example, imagine that we want a computer to produce a digital copy of Van Gogh's *Starry Night*. A very inefficient way of instructing the computer would be to make a list of all the pixels it has to fill in, each with a particular color. The computer would get the job done, and its reproduction of the *Starry Night* would be extremely accurate, but that would be tremendously time and energy consuming: a program that instructs how to fill in precisely each pixel would be needlessly long. Another, vastly more efficient way would be to describe a mathematical function on how the stars in the painting are arranged, another on how the village is layed out, etc. This description would be significantly shorter than the previous one, even though it may not be as complete as the former. Once the computer finishes processing these functions, there will probably be spots in the picture that are not identical to the original painting, but it is nevertheless recognizable as similar to Van Gogh's famous work. In this case, the most efficient description is the pattern of the *Starry Night*: it conveys only the relevant information to be close enough to the painting, while some details will be abstracted away.

This example illustrates how the same object or system can be described by more than one pattern. Which pattern we ultimately decide to be the most adequate depends on our pragmatic concerns. If we want a perfect copy of the *Starry Night*, the pixel-list would be better; if we don't need such precision, a set of functions roughly describing what the arrangement of elements is would be the most appropriate.

Moreover, the example above highlights two important aspects of patterns: (1) they are real entities in the world, in so far as they can be empirically investigated and, often, mathematically described; and (2) they are perspective-dependent, in so far as they are relative to the particular methods and frameworks of researchers. Haugeland (1998, p. 274) distinguishes these two aspects of pattern in terms of, respectively, (1) "orderly and non-random arrangement" and (2) "candidate for recognition". These characteristics are individually necessary, and jointly sufficient, for a given arrangement to constitute a pattern. In particular, the necessity of (2) indicates that an "unobservable pattern" is a misnomer: if there is a pattern in a given system, it should be observable given the adequate methodological tools.

Applying the idea of patterns to mechanistic inquiry, several philosophers have argued that phenomena description is just a way to figure out which patterns there are in a given system (LEE; DEWHURST, 2021). In this regard, considering how overlapping and chaotic causal relations tend to be in the life and mind sciences, delineating a phenomenon involves figuring out which relations and elements tend to persist despite the surrounding noise. Once this pattern is made explicit, researchers are able to make some predictions about the behavior of the system, since patterns are intrinsically non-random regularities that endure as the surrounding context changes (cf. DENNETT, 1991).

It is important to note that recognizing a certain pattern in a causal system is not the same as uncovering the mechanism that produces that pattern. Following Dennett (1991), to understand a pattern we don't need to describe details on how it was produced. Going back to the *Starry Night* example, the pattern of stars arrangement can be described and studied without reference to how the pattern is manifested empirically (either through paint on a canvas, or pixels on a computer screen). This feature is mirrored in phenomena delineation since, at this stage in mechanistic inquiry, we also still don't know how the phenomenon is implemented; we are just able to describe it in a higher level of abstraction. As such, when we are delineating phenomena, we are describing higher-level characteristics via a pattern meanwhile suspending judgement on the finer details of implementation.

By taking seriously Haugeland (1998)'s idea of pattern as a "candidate for recognition", delineation of phenomena is also dependent on perspective and pragmatic interests. Given that recognizing a pattern is dependent on some perspective, delineating phenomena from their surrounding environment is also dependent on researchers' perspectives and methodologies. For example, Kästner & Haueis (2021, p. 1649) discuss how a "pattern recognition practice" is composed of multiple epistemic activities, such as characterizing a system via operational definitions, or creating schematic models of interactions within that system. Among a community of researchers, these activities need to be coordinated in such a way that they do not become mutually exclusive, but capable of integration into a more definite and repeatable pattern.

As such, not only does pattern recognition depend on the researchers' perspective, but also it is a highly empirical endeavor. It requires both theoretical modelling of known data, as well as experimental approaches that test these models. Given the interaction between empirical methods and how phenomena are delineated, it is likely that the initial characterization of a phenomenon will be improved as research develops. Bechtel & Richardson (2010) highlight this feature of mechanistic inquiry. For instance, they note how the classical Mendelian view of genes as autonomous determinants of phenotypes had to be significantly altered once deviations from this characterization were discovered (BECHTEL; RICHARDSON, 2010 chapter 8). Such findings not only lead to a greater understanding of the underlying mechanism of genetic determination, but also significantly changed the general research question of genetics: now the phenomenon is couched in probabilistic terms, with much greater appreciation on how genes are not the solely determinants of phenotypes, but instead produce biochemical changes that may lead up to a change of characteristic.

Let's take stock. As we saw in section 1, mechanisms are always characterized in reference to a certain phenomenon. More succinctly, mechanisms are always *for* some phenomenon (GLENNAN, 2017; MACHAMER; DARDEN; CRAVER, 2000). This entails that uncovering mechanisms necessitates an adequate description of the phenomenon under investigation. Such descriptions involve the delineation of the phenomenon from its environment and, as such, are empirical tasks that aim to figure out which pattern exists in a given system. Moreover, if phenomenon delineation is about figuring out patterns, and if patterns are intrinsically perspective dependent, then characterizing phenomena must also be so dependent. This brings forth an important problem for the epistemic adequacy of mechanistic explanations. If phenomena delineation is necessary for mechanistic inquiry and is perspective-laden, then how can we maintain that mechanistic explanations posit real entities? Aren't we thus forced to accept an anti-realist position on mechanisms? In the next section, I argue that the concept of *organizational emergence* provides a useful explication on how empirical evidence constraints mechanistic inquiry and phenomena delineation. Such criteria are called *bottom-up constraints*, and they provide sufficient realist grounds for mechanistic inquiry to block such objection.

4. Bottom-up constraints

Reference to mechanisms is ubiquitous in the natural sciences. From forming galaxies to synthesizing proteins, mechanisms are often held to be responsible for these and other phenomena. In this vein, mechanistic explanations always start with a characterization of some phenomenon, which, as described in the previous section, is a feature that might bring serious doubts on a realist account of mechanistic inquiry. If phenomena delineation, the critique goes, is a type of pattern recognition, and if patterns are always dependent on some perspective, then mechanistic explanations are grounded in a perspective-laden activity, thus forcing us to adopt an anti-realist position about mechanisms. Such is the *anti-realist objection*, which is formulated as follows:

- P1. Mechanistic inquiry explains a phenomenon by describing how a mechanism is responsible for it;
- P2. If (P1), then delineating mechanisms is dependent on how its phenomenon is characterized;
- P3. Characterizing phenomena is a matter of pattern recognition;
- P4. Detecting patterns inherently depends on the perspective of a subject;
- C1. *Therefore*, delineating phenomena inherently depends on the perspective of researchers;
- P5. If (P2) and (C1), then delineating mechanisms inherently depends on the perspective of researchers;
- P6. If (P5), then mechanisms don't exist independently on some perspective - they are best seen as useful explanatory tools for some research project;
- C2. *Therefore*, anti-realism about mechanisms is true.

In this section, I defend that this argument does not hold for mechanistic inquiry. Based on the conception of *organizational emergence*, I argue that P6 does not follow: even if mechanistic inquiry depends on the perspective of researchers, it does not necessarily follow that mechanisms are only explanatory tools. Instead, I defend that it is precisely because mechanisms have firm basis on empirical evidence that they have their explanatory weight. Such evidence, in my view, constrains in a *bottom-up* way the functioning of the mechanism and the production of the phenomenon.

My argument starts with the observation that investigators often move between levels of abstraction to make sure that their description of upper-levels is consistent with what is happening at lower-levels (BECHTEL; RICHARDSON, 2010). The way in which physical descriptions limit the range of higher-level phenomena is called *bottom-up constraint*. As such, bottom-up constraints limit the range of relevant and useful perspectives for the phenomenon. Investigating a system entails figuring out which patterns are applicable to it or not. In this sense, patterns may be perspective dependent, but only partially so: they also depend on which empirical evidence there is available.

To make sense of researchers moving between levels, we need a more detailed account of how mechanistic levels are organized, how they relate to each other and could possibly be integrated. Craver (2015) and Povich & Craver (2018) argue that levels in a mechanism are related to one another via *organizational emergence*. They present their account by contrasting it with a reductionist view of levels, according to which higher-level phenomena are nothing more than the sum of lower-level parts. For example, a reductionist would claim that the temperature of a room simply consists in the average kinetic energy of all molecules in that room. In this case, there is nothing over and above a certain temperature than the movement of molecules; we can therefore say that temperature is only an aggregate function of molecules' energy. Following Craver and Povich, an aggregation of parts does not necessitate any particular organization of such parts. For instance, if ψ is an aggregate of elements in set $X = [x_1, x_2, \ldots, x_n]$, ψ is reducible to X in such a way that the organization of x_i is irrelevant for forming ψ. In sum, if the higher-level is only an effect of the culmination of lower-level elements, then only the latter carry any explanatory weight. Reductionism is thus incompatible with the idea that higher-level descriptions are explanatory and, therefore, that there are significant top-down constraints in mechanistic inquiry.

Let us assume, for the sake of argument, that reductionism is true about levels of mechanisms. As such, any phenomenon, ψ, is reducible to the components $[x_1, x_2, \ldots, x_n]$ in the sense that ψ, as well as its properties, are nothing over and above the collection of x_i, and any organization of x_i will be sufficient for ψ. If such is the case, then a very effective way of studying ψ would be to analyze the properties of each x_i individually, given that the interactions between each x_i do not matter. The end goal of this reductive approach is to ultimately derive ψ from the aggregate of x_i (POVICH; CRAVER, 2018, p. 191).

However, to understand the mechanistic relation between ψ and x_i as reductive is significantly at odds with the standard characterization of mechanisms: that is, the behavior of the phenomenon, ψ, is determined by the properties of x_i *and* their coordinated interactions. In this characterization, ψ is literally more than the simple aggregate of x_i, for the interactions between each element x_i is necessary for there to be ψ. Consequently, while it might be useful to study each x_i in isolation to understand some features of the mechanism for ψ, it is not necessary nor sufficient to do so in a proper mechanistic inquiry.

In this scenario, if the reductionist model for mechanisms fails, what should we replace it with? A standard and promising approach is organizational emergence. It consists in the thesis that the phenomenon of a mechanism is dependent not only on the individual parts of the system, but also on how these parts interact with each other. The phenomenon ψ thus emerges from the collection of x_i only if its members are organized and interactive in very particular ways. For example, imagine the audience of a soccer match doing the “stadium wave”: successive groups of people stand up, raise their arms, and sit down. From a distance, the coordinated movement of these people forms a wave that can “travel” across the stadium. The wave cannot be reduced to the movement of a particular person, for the very fact that it requires an orchestrated behavior means that it can only be observed once we abstract away from the particularities of each person to the movement of the whole. In this case, the height, speed, and position of a particular member is not relevant for explaining how the wave is formed. We can only explain it from a relatively higher level of abstraction.

Such characterization of organizational emergence fits well with the notion of pattern, discussed in the previous section. The pattern cannot be reduced to the behavior of individual parts because it requires a more distant standpoint to be picked out from (DENNETT, 1991; HAUGELAND, 1998). If we focus too much on individual parts, we are bound to miss out how the collection of parts is orchestrated as to produce the pattern. For example, some neuroscientists argue that single-neuron analysis and focusing on one particular brain region at a time is not enough to capture how the enormously complex collection of these elements make up cognitive processes (cf. ANDERSON, 2014; PESSOA, 2022). In this regard, understanding cognitive processes requires a level of granularity that is able to encompass how complex the interaction between brain networks really is. If we focus on single-neuron or even single-area analysis, we don’t see the entire network in action and, hence, miss the neurocognitive mechanism we want to capture.

Going back to the anti-realist objection against the pattern account of mechanistic inquiry, organizational emergence provides us with an interesting way of answering that critique. If organizational emergence is true about phenomena and their mechanisms, then to explain phenomenon ψ we need to know how its component parts, x_i, are organized. That we cannot do without methodical empirical work. If ψ is strongly constrained by the organization and physical properties of x_i, then, were we to discover that they cannot support ψ in the way we currently characterize it, then we must change our conception of ψ to accommodate what we know about x_i.

Hence, mechanistic inquiry is not completely perspectival - it necessitates bottom-up constraints. Still, if both bottom-up and top-down constraints are necessary for a proper mechanistic explanation, how do they relate with each other? Does one take precedence over the other? How do they interact in actual scientific practice? These are the topics of the next section.

5. Multi-level constraints in mechanistic inquiry

From the previous sections, we have seen how mechanistic inquiry must be *multi-level constrained*: a mechanistic account of a phenomenon must be consistent with our best description of the *explanandum* and with the physical structure of the component parts. In this framework, if an account does not meet either of both constraints, then it does not carry sufficient explanatory weight.

There is an interesting parallel between multi-level constraints and the discussion on whether mechanistic explanations are *ontic* or *epistemic*. Salmon (1984b) introduced these distinctions as a way to more clearly differentiate his account with Hempel (1965)'s deductive-nomological view of scientific explanation. Following Salmon, Hempel's account is *epistemic* in so far as it characterizes explanations as arguments, wherein the conclusion amounts to the *explanandum*, and the premises, the *explanantia*. Meanwhile, *ontic* accounts of explanations take them as "exhibitions of the ways in which what is to be explained fits into natural patterns or regularities" (SALMON, 1984b, p. 293). In other words, ontic explanations highlight how the *explanandum* phenomenon is inserted in the causal structure of the world.

In the context of the mechanistic literature, Salmon's terminology has been slightly modified to be more inclusive. In this framing, epistemic accounts of mechanistic explanations are more concerned with viewing them as a human activity: explanations thus describe mechanisms as a way to improve knowledge about empirical phenomena (BECHTEL, 2008). Alternatively, ontic accounts understand the mechanisms themselves as explanatory, for they are the basis upon which phenomena are inserted into the causal structure of the world (CRAVER, 2007). While most mechanist philosophers tend to agree with the ontic view, there are still substantial debates between proponents of the epistemic and ontic accounts (cf. HALINA, 2018).

Illari (2013) defends that such characterization of the debate is changing as mechanistic philosophy develops. According to her, new mechanists are increasingly moving away from the traditional understanding of the debate (i.e., as one about what explanations are), towards a framework that asks which explanatory aspects should be given priority. In this vein, epistemic accounts argue that the important constraints on mechanistic inquiry are methodological and, to some degree, relative to the psychological underpinnings of scientific practice; meanwhile ontic views defend that empirical and causal criteria should take precedence over epistemic ones (ILLARI, 2013).

Given such a normative twist on the ontic-epistemic debate, it becomes clear how, in reality, both positions are not mutually exclusive. In fact, Illari (2013) herself shows how both ontic and epistemic theorists can accept each other's views, which diminishes the importance of their disagreement about what criteria are more fundamental. According to Illari:

> It seems that the most sensible conclusion to draw is that neither aim of mechanistic explanation is prior to the other. Ontic and epistemic constraints are both ineliminable, as both aims must be met, to generate a successful mechanistic explanation:
>
> - Describe the (causal) structure of the world: to be distinctively mechanistic, describe the entities and activities and the organization by which they produce the phenomenon or phenomena;
> - Build a model of the activities, entities and their organization that scientists can understand, model, manipulate and communicate, so that it is suitable for the ongoing process of knowledge-gathering in the sciences (ILLARI, 2013, p. 250).

Kästner & Haueis (2021) similarly argue that such debate constitutes a false dilemma: mechanistic explanations must attend to *both* ontic and epistemic criteria. Their premises, however, are more metaphysically flavored. According to them, ontic and epistemic criteria are equally important because they constitute fundamental, yet completely distinct, ways of picking out patterns in phenomena and mechanisms. Following them:

> It is not enough for a pattern recognition practice to simply characterize a phenomenon, i.e., what is salient from above. Mechanistic explanation also requires researchers to specify the elements of a pattern, i.e., what makes the pattern persist from below. To achieve this goal, researchers must introduce various *epistemic operations* that track the entities and activities constituting the pattern. The selection of such epistemic operations is ontically constrained: scientists must tailor them to the particular spatiotemporal characteristics of the entities and activities they are supposed to track (KÄSTNER; HAUEIS, 2021, p. 1649, emphasis from the original).

Ontic, or bottom-up, constraints are widespread in the empirical sciences. As an example of such criteria, consider the recent research on the relation between memory and decision-making. Starting from a commonsense understanding, decision-making consists in a process of evaluating the pros and cons of a set of alternatives in order to eventually make a choice. In this scenario, the process of decision-making has very distinct starting and terminating conditions: we begin with a range of alternatives and, ideally, stop with only one. However, some psychological studies suggest that decision-making does not quite end after an alternative is picked out. In an experiment conducted by Biderman & Shohamy (2021), participants were presented with a pair of paintings, A and B, and had to choose which one would generate a higher profit in an auction. After the participant made their decision (say, painting A) and learned the consequent outcome (say, made a profit), they were asked what they thought the value for painting B was. The majority of participants answered that painting B was less profitable than painting A, lending support for Biderman & Shohamy (2021)'s hypothesis that there is an underlying memory component for decision making, which updates the values of unchosen alternatives by association with chosen ones. This indicates that, even after the decision has been made, the system that produced that decision is still active in considering other alternatives. The experiment thus entails something about a component, in this case memory, of the decision-making mechanism. From data about this component, we have to change how we think about the phenomenon: in this case, how it is still active well after what we first thought it was the terminating point.

Importantly, changing the description of a phenomenon motivated by evidence of lower-level functioning only necessarily applies to a scientific understanding of these phenomena. With regards to changing any commonsense description of a cognitive state, the multi-level framework presented here is neutral. It may be the case that, for some phenomena, there is a real need to change how common sense makes reference to them, but that is not necessarily the case. What is relevant for the present purposes is how the way in which we scientifically understand cognitive phenomena is constrained by lower levels of explanation.

Changing how we conceive of starting and terminating conditions, however, is not the only task of mechanistic inquiry. Top-down and epistemic constraints are also at play when delineating the mechanisms for a given phenomenon. These constraints work by narrowing down the range of possible mechanisms for a phenomenon: from a particular description of a behavior, only some configurations of components and operations are able to produce a phenomenon that matches that description. Once these configurations are properly characterized, scientists can conduct experiments that aim to identify which configuration is actually implemented in the system they are investigating.

One example of top-down constraints is Marr (1981)'s computational analysis of information processing in the human retina. Famously, Marr divided explanations of cognitive phenomena into three levels:

1. *Computational*: characterizes the problem a system has to solve, as well as the mathematical function that resolves such problem;
1. *Algorithmic*: describes how the mathematical function from the computational level can be executed in a series of steps;
2. *Implementational*: details how the algorithmic steps above are implemented in physical hardware.

Note how the Marrian levels form a hierarchy, wherein the lower-levels stand for explanations of how upper-level elements are realized. As such, the algorithmic and implementational levels are finer-grained descriptions of how the system behaves, and a mathematical, coraser-grained, description of that behavior is given by the computational level. Hence, given that the *explanandum* phenomenon is the overall behavior of the mechanism (GLENNAN, 2017; MACHAMER; DARDEN; CRAVER, 2000), the computational level is nothing over and above than a mathematical description of the phenomenon (cf. SHAGRIR; BECHTEL, 2017).

Such quantitative description of the phenomenon is greatly relevant for mechanistic inquiry. From such description, scientists have a more clearly defined target to guide their mechanistic inquiry. The mathematical description of the phenomenon then becomes a parameter against which any mechanistic description must be upheld. As such, if a proposed mechanism does not properly implement the function associated with the phenomenon, it must be discarded as a likely explanation.

Still, there are important caveats with this view. Firstly, scientists may revise their phenomenon description to accommodate evidence on the physical characteristics of the mechanism, instead of searching for mechanisms elsewhere and saving their phenomenon description. This is what happened with the classical computational theory of cognition, due to McCulloch & Pitts (1943): once we learned that the brain is not suited to serial processing of information (or, at least, not to the same extent as a digital computer), the description of cognition as computation had to be altered to fit the brain’s aptitude for parallel and distributed processing (cf. CHURCHLAND; SEJNOWSKI, 1992; PICCININI, 2020).

Secondly, a computational description of a phenomenon falls short of a complete mechanistic explanation. Considering how full mechanistic explanations have to account for how the coordinated operations of components yield the phenomenon, any account that does not mention the physical implementation of a process is not properly mechanistic (PICCININI; CRAVER, 2011) In this sense, describing the phenomenon quantitatively and remaining neutral about its physical implementation is only an incomplete mechanistic account - even though it remains as a very important stage in mechanistic inquiry.

In sum, both top-down and bottom-up constraints are individually necessary, and jointly sufficient, for a proper mechanistic explanation. As such, these explanations are *multi-level constrained*, a property which highlights not only how difficult it is to achieve an adequate mechanistic account, but also how valuable these explanations are. They reflect our epistemic criteria and necessities, as well as the causal and constitutive structure of the world.

6. Summary and conclusions

In the past few decades, the near omnipresence of mechanisms in the sciences has been thoroughly analyzed by philosophers of science. Such blooming research topic yielded impactful insights for our understanding of the art of explanations. In this paper, we looked at how mechanistic explanations must be developed in such a way that both meet our epistemic demands and reflect the causal structure of the world. These constraints are called, respectively, *top-down* and *bottom-up*, and together they make sure that mechanistic accounts are adequate for multiple levels of explanation.

Regarding top-down constraints, they consist in researchers' considerations on what the *explanandum* phenomenon is and how it causally interacts with the environment. Given that mechanisms are always mechanisms *for* a given phenomenon, how we understand the latter greatly influences how we go about discovering the former. In this sense, any mechanistic inquiry must start with attempts at delineating the phenomenon in question. This separation of the phenomenon from its environment consists in an abstract understanding of the system under scrutiny, thus limiting the range of possible mechanisms that might be responsible for the system's behavior.

The fact that delineating the phenomenon from its surroundings is a major step during mechanistic inquiry indicates that it is not a task that can be accomplished a priori. As Bechtel & Richardson (2010) note, characterizing the phenomenon is an empirical task precisely because of the fact that we have picked it out from the causal confusion that is the material world. Moreover, considering that these mechanistic phenomena are repeatable across a range of environments (MACHAMER; DARDEN; CRAVER, 2000), phenomenon delineation has to be able to detect what remains the same across these tokens. Kästner & Haueis (2021) argue that such task fundamentally consists in recognizing patterns, where they are understood both as a non-random organization of elements and a candidate for recognition (DENNETT, 1991; HAUGELAND, 1998).

One possible criticism of the patterns account is that it might leave us no choice but to accept anti-realism about mechanisms. If, the critique goes, delineating phenomena is a necessary task for mechanistic inquiry and amounts to a perspective-laden process of pattern recognition, then the discovery of mechanisms is also perspective-laden and may not reflect any real structure in nature. As such, it seems that the inherently perspectival character of patterns is incompatible with a realist position on mechanisms.

Such counterargument can be blocked by attending to how empirical constraints also have an effect in mechanistic inquiry. These criteria are called *bottom-up constraints*, and they serve as the realist and empirical counterweight to the top-down theory-laden constraints. In this regard, mechanistic inquiry must also attend to whether the physical and organizational structure of the (proposed) components can support the characterization of phenomena that come from the top down. In the same way that cognition cannot be thought of as a serial process of computation because the brain's architecture cannot support this type of processing, an abstract description of a phenomenon has to be abandoned if the underlying hardware is not capable of producing the same behavior.

Bibliography

ANDERSON, M. L. **After Phrenology: Neural Reuse and the Interactive Brain**. Cambridge, Massachusetts: The MIT Press, 2014.

BECHTEL, W. **Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience**. New York: Routledge, 2008.

BECHTEL, W.; ABRAHAMSEN, A. Explanation: A mechanist alternative. **Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences**, Mechanisms in biology. v. 36, n. 2, p. 421–441, jun. 2005.

BECHTEL, W.; RICHARDSON, R. C. **Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research**. MIT Press ed. ed. Cambridge, Mass: MIT Press, 2010.

BIDERMAN, N.; SHOHAMY, D. Memory and decision making interact to shape the value of unchosen options. **Nature Communications**, v. 12, n. 1, p. 4648, Jul. 2021.

CHURCHLAND, P. S.; SEJNOWSKI, T. J. **The computational brain**. Cambridge, Mass.: MIT Press, 1992.

CRAVER, C. F. **Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience**. New York: Oxford University Press, 2007.

______. Functions and Mechanisms: A Perspectivalist View. In: HUNEMAN, P. (Ed.). **Functions: Selection and Mechanisms**. Synthese Library. Dordrecht; New York: Springer, 2013.

______ Levels. In: METZINGER, T.; WINDT, J. M. (Eds.). **Open MIND**. [s.l.] MIND Group, 2015.

CUMMINS, R. **The nature of psychological explanation**. Cambridge, Mass.: MIT Press, 1985.

DARDEN, L. Thinking Again about Biological Mechanisms. **Philosophy of Science**, v. 75, n. 5, p. 958–969, Dez. 2008.

DENNETT, D. C. Real Patterns. **The Journal of Philosophy**, v. 88, n. 1, p. 27, 1991.

DEWHURST, J. Computing Mechanisms Without Proper Functions. **Minds and Machines**, v. 28, n. 3, p. 569–588, set. 2018.

GLENNAN, S. **The New Mechanical Philosophy**. First edition ed. Oxford: Oxford University Press, 2017.

HALINA, M. Mechanistic Explanation and Its Limits. In: GLENNAN, S.; ILLARI, P. M. (Eds.). **The Routledge Handbook of Mechanisms and Mechanical Philosophy**. Routledge Handbooks in Philosophy. London; New York: Routledge, Taylor & Francis Group, 2018. p. 213–237.

HARBECKE, J.; SHAGRIR, O. The role of the environment in computational explanations. **European Journal for Philosophy of Science**, v. 9, n. 3, p. 37, set. 2019.

HAUGELAND, J. Pattern and Being. In: **Having Thought. Essays in the Metaphysics of Mind**. Cambridge, Mass: Harvard University Press, 1998. p. 267–290.

HEMPEL, C. G. **Aspects of Scientific Explanation and Other Essays in the Philosophy of Science**. New York, NY: The Free Press, 1965.

ILLARI, P. Mechanistic Explanation: Integrating the Ontic and Epistemic. **Erkenntnis**, v. 78, n. S2, p. 237–255, dez. 2013.

KAPLAN, D. M.; CRAVER, C. F. The Explanatory Force of Dynamical and Mathematical Models in Neuroscience: A Mechanistic Perspective. **Philosophy of Science**, v. 78, n. 4, p. 601–627, out. 2011.

KÄSTNER, L.; HAUEIS, P. Discovering Patterns: On the Norms of Mechanistic Inquiry. **Erkenntnis**, v. 86, n. 6, p. 1635–1660, dez. 2021.

LEE, J.; DEWHURST, J. The Mechanistic Stance. **European Journal for Philosophy of Science**, v. 11, n. 1, p. 20, jan. 2021.

MACHAMER, P.; DARDEN, L.; CRAVER, C. F. Thinking about Mechanisms. **Philosophy of Science**, v. 67, n. 1, p. 1–25, mar. 2000.

MARR, D. **Vision: A computational investigation into the human representation and processing of visual information**. Cambridge, Mass: MIT Press, 1981.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The Bulletin of Mathematical Biophysics**, v. 5, n. 4, p. 115–133, dez. 1943.

MILLIKAN, R. G. **Language, Thought, and Other Biological Categories: New Foundations for Realism**. Cambridge, Mass: MIT Press, 1984.

OPPENHEIM, P.; PUTNAM, H. The Unity of Science as a Working Hypothesis. In: FEIGL, H.; MAXWELL, G. (Eds.). **Concepts, Theories and the Mind-Body Problem**. Minnesota Studies in the Philosophy of Science, v. 2. Minneapolis: University of Minnesota Press, 1958. p. 3–36.

PESSOA, L. **The Entangled Brain: How Perception, Cognition, and Emotion Are Woven Together**. Cambridge, Massachusetts: The MIT Press, 2022.

PICCININI, G. **Neurocognitive mechanisms: Explaining biological cognition**. First edition ed. Oxford: Oxford University Press, 2020.

PICCININI, G.; CRAVER, C. Integrating psychology and neuroscience: Functional analyses as mechanism sketches. **Synthese**, v. 183, n. 3, p. 283–311, dez. 2011.

POVICH, M.; CRAVER, C. Mechanistic Levels, Reduction, and Emergence. In: GLENNAN, S.; ILLARI, P. M. (Eds.). **The Routledge Handbook of Mechanisms and Mechanical Philosophy**. Routledge Handbooks in Philosophy. London; New York: Routledge, Taylor & Francis Group, 2018.

SALMON, W. C. **Scientific Explanation and the Causal Structure of the World**. Princeton, N.J: Princeton University Press, 1984a.

______. Scientific Explanation: Three Basic Conceptions. **PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association**, v. 1984, n. 2, p. 293–305, 1984b.

SHAGRIR, O.; BECHTEL, W. Marr's Computational Level and Delineating Phenomena. In: KAPLAN, D. M. (Ed.). **Explanation and Integration in Mind and Brain Science**. First edition ed. Oxford, United Kingdom: Oxford University Press, 2017.

STICH, S. P. **From Folk Psychology to Cognitive Science: The Case against Belief**. 5th print ed. Cambridge, Mass.: MIT Press, 1985.

WOODWARD, J. A Theory of Singular Causal Explanation. **Erkenntnis (1975-)**, v. 21, n. 3, p. 231–262, 1984.

Recebido em: 12/09/2023

Aceito em: 19/11/2023